iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** I **Month of publication:** January 2022

DOI: https://doi.org/10.22214/ijraset.2022.39822

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue I Jan 2022- Available at www.ijraset.com

Optimization Techniques to Solve Travelling Salesman Problem Using Machine Learning Algorithms

Prince Nathan. S[1], Onkar Saudagar[2], Rutika Shinde[3]
[1]*PG Student, Department of Data Science and Analytics, National Institute of Electronics and Information Technology, Chennai, India*
[2]*Student, Pune Institute of Computer Technology, Department of Information Technology Engineering, Pune, India*
[3]*Student, Sinhgad Institute of Technology and Science, Department of Information Technology, Pune, India*

Abstract: ***Travelling Salesmen problem is a very popular problem in the world of computer programming. It deals with the optimization of algorithms and an ever changing scenario as it gets more and more complex as the number of variables goes on increasing. The solutions which exist for this problem are optimal for a small and definite number of cases. One cannot take into consideration of the various factors which are included when this specific problem is tried to be solved for the real world where things change continuously.***
There is a need to adapt to these changes and find optimized solutions as the application goes on. The ability to adapt to any kind of data, whether static or ever-changing, understand and solve it is a quality that is shown by Machine Learning algorithms. As advances in Machine Learning take place, there has been quite a good amount of research for how to solve NP-hard problems using Machine Learning. This report is a survey to understand what types of machine algorithms can be used to solve with TSP.
Different types of approaches like Ant Colony Optimization and Q-learning are explored and compared.
Ant Colony Optimization uses the concept of ants following pheromone levels which lets them know where the most amount of food is. This is widely used for TSP problems where the path is with the most pheromone is chosen.
Q-Learning is supposed to use the concept of awarding an agent when taking the right action for a state it is in and compounding those specific rewards.
This is very much based on the exploiting concept where the agent keeps on learning on its own to maximize its own reward. This can be used for TSP where an agent will be rewarded for having a short path and will be rewarded more if the path chosen is the shortest.
Keywords: LINEAR REGRESSION, LASSO REGRESSION, RIDGE REGRESSION, DECISION TREE REGRESSOR, MACHINE LEARNING, HYPERPARAMETER TUNING, DATA ANALYSIS

I. INTRODUCTION

In today's day and age, a lot of online websites have started delivery from town to town or city based. It leads to many variations of traveling salesman's problem, vehicle routing problem, job-shop scheduling problem, and more variants. These problems are classified as Combinatorial Optimization Problems. They are commonly found in domains like transportation, operations research, logistics, and telecommunications. The overall goal is to find an optimal solution, which can be modeled as a sequence of actions/decisions to maximize/minimize the objective function for a specific problem. However, as the scale of the problem grows, the time to find the optimal solution increases exponentially.
TSP is a non-deterministic polynomial-time hard (NP-hard) computational problem with no-known polynomial time solutions yet. There are still many questions open to be answered for the worst-case analysis of exact algorithms for NP-hard problems. Today, due to more research in the field of machine learning, we can compare some well-established machine learning algorithms such as Ant Colony Optimization and Q-learning.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

Sr	Paper	Summary	Limitations	Publication Year
1	Reinforcement Learning for Combinatorial Optimization[1],	Explore the recent advancements of applying RL frameworks to hardcombinatorial problems, which options to consider and theproblems generated by it	Does not discuss how to deal with parameter tuning	2020
2	Deep Reinforcement Learning for Travelling Salesman Problem with Time Rejection Windows[2]	Discusses RL algorithms to solve TSPWTRW with 100-1000 times faster results than tabu searches.	Discusses only with one agent and one variable which is time rejection.[3]	2020
3	A Comparative Study of Machine Learning Heuristic Algorithms to Solvethe Traveling Salesman Problem[3]	Gives a comparison between different algorithms like NN, Genetic Algorithm, Ant Colony Optimization and Q-Learning	Deals with the base case of TSP withoutintroducing any variables like time and fuel	2010
4	A comparative analysis of the traveling salesman problem: Exact andmachine learning techniques[4]	Compares Exact, Heuristic and machine learning algorithms performances on the traveling salesman problem	Does not have enough information on themachine learning algorithm used	2019
5	Reinforcement learning for the traveling salesman problem with refueling[5]	Discusses using SARSA and Q-Learning for solving TSP with refueling. Also discusses RL parameter tuning for this exact matter.	Does not consider fuel stations as nodesthat represent more close to the real life situation.	2021

III. TSP AND ALGORITHMS

A. Introduction

Traveling Salesman Problem (TSP) is a classical and most widely studied problem inCombinatorial Optimization. It has been studied intensively in both Operations Research and Computer Science since the 1950s as a result of which a large number of techniques were developed to solve this problem. Much of the work on TSP is not motivated by direct applications, but rather by the fact that it provides an ideal platform for the study of general methods that can be applied to a wide range of Discrete Optimization Problems. Indeed, numerous direct applications of TSP bring life to the research area and help to direct future work.

The idea of the problem is to find the shortest route of salesman starting from a given city, visiting n cities only once and finally arriving at origin city.

B. Different Approaches to solve TSP

There are two types of approaches to solving TSP.

One way is to use exact solvers, which implies getting the optimal solution every time by using methods like Interior Point, Branch-and-Bound, and Branch- and-cute. But these come at cost of huge run times, space requirements and are only viable for a fewer number of nodes.

Another way is to use non-exact solvers. These methods offer potentially non-optimal but typically faster solutions. The way to get better results for solving TSP is to make these types of algorithms more optimized and being fast at the same time to get the ideal solution.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue I Jan 2022- Available at www.ijraset.com

C. State of Art Algorithms

State of the Art heuristic approaches to solve TSP are Ant Colony Optimization and Q-Learning which will be discussed below.

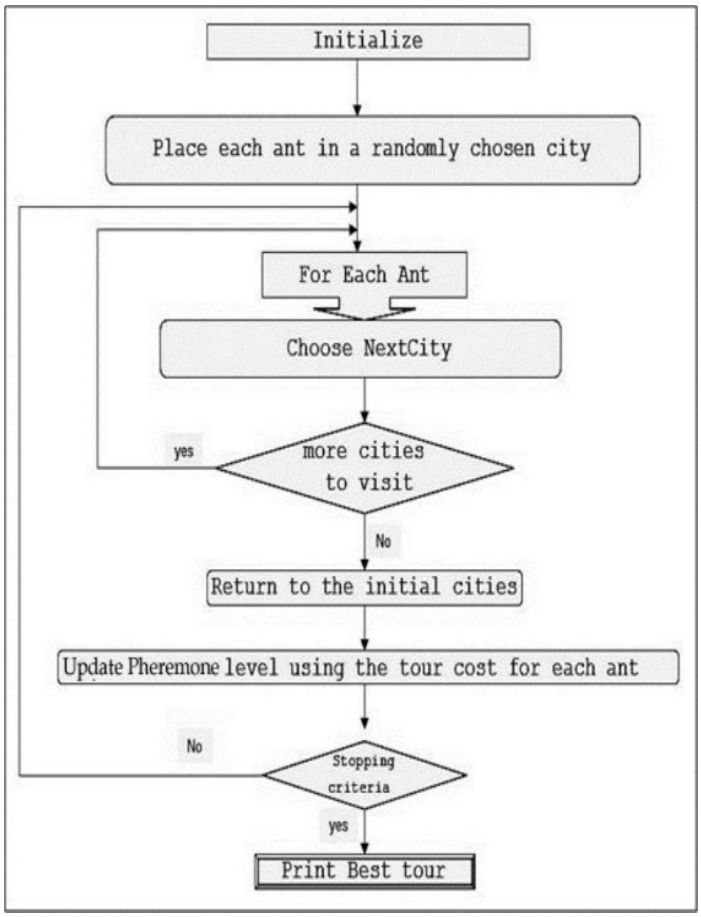


Figure 1: Flowchart of implementation of ACO for TSP

D. Ant Colony Optimization

Ant Colony Optimization is a heuristic algorithm that is inspired by examining ants and their movement. Ants cooperate to find resources by laying down a chemical on the ground called a pheromone for other ants to follow. The more amount of pheromone, the high probability the ants will follow that path. In this fashion, shorter routes to food sources get more amount of pheromone.

By using this phenomenon, researchers have devised an algorithm for optimizing different problems. One of the problems where it is applied is TSP. This algorithmuses the nearest-neighbor search and diffluence strategy.

1) *Using ACO for TSP*: The algorithm for implementing ACO for solving TSP is shown below in form of a flowchart. The ants are placed at random cities and are made to move to the next city byusing a formula that uses the concepts of probability, the pheromone levels, andheuristic value. It donates the probability of ant k that is in city r to go in city s.

$$p_k(r,s) = \frac{[\tau(r,s)]^{\alpha}[\eta(r,s)]^{\beta}}{\sum_{u \in U_k(r)}[\tau(r,s)]^{\alpha}[\eta(r,s)]^{\beta}} \quad if\ q > q_0.$$

α denotes the importance of pheromone for the ant and β denotes the importance of heuristic values.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

2) *Drawbacks*: This algorithm takes a lot of time for searching and is prone to be trapped in the local minima. Though ACO tries to prevent it by updating local rules, there are other ways to boost it's performance using machine learning which are being researched.

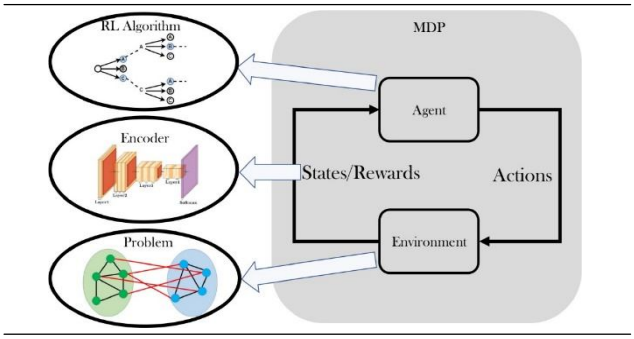


Figure 2: RL pipeline which uses an encoder a graph neural network to represent the nodes as vectors.

E. Q-Learning

Q-Learning is a reinforcement learning algorithm that is used on agents for predicting an action for a state the agent is in. This algorithm specializes in maximizing the rewards the agent can get in the current and successive states.

1) *Using Q-Learning for TSP*: The agent is placed at a random city and it takes action according to the Q-value function giving out different Q values. The Q- values are stored in a Q-Table which can be used to find the shortest path. We use a ϵ and a learning rate γ. By using the ϵ-greedy approach, if the ϵ has the greatervalue, a city is randomly selected, otherwise using the nearest neighbor. This gives animmediate reward to the agent which makes it adjust its Q-values and checks if the tour length was the shortest from the beginning. If it is found, the rewards areupdated by themselves, and the agents are given a reward to the ones who belong to that shortest tour. Then the global minimum tour is constructed by using the Q-values. RL needs a setup of environment enabled with an encoder/decoder system totake necessary actions according to its states. The diagram of an example pipelineis given above [Fig 2].

```
Set the parameters: α, γ and ϵ ;
Initialize the matrix Q(s, a) = 0 ;
Initialize TSPWR variables and constants ;
Observe the state s0: initial city;
repeat
    repeat
        Select the action a (destination city) using
          ϵ-greedy method ;
        Take the action a;
        Calculation of the fuel level in the tank:
          level;
        % Calculation of the truck cost:;
        if level < 0 then
            Reset the tank level: level = 0;
            Calculation of the truck cost: g_ij;
            z_ij = 1 ;
        else
            The truck cost is zero;
            z_ij = 0;
        end
        % Calculation of the refueling cost:;
        if (level < level-ref) and ( a =! s0) then
            Calculation litres amount for refueling:
              l_j;
            Calculation of the refueling cost: c_j l_j;
            Maximum tank level: level = L_max ;
        else
            The refueling cost is zero: c_j l_j = 0;
        end
        Updates the total cost (route): Eq. (4) ;
        Updates the distance travelled on the route;
        Receive immediate reward: Eqs. (13) to (17);
        Observe the new state s' (new city);
        Update Q(s,a): Eq. (3);
        s = s';
    until complete the route;
until the stopping criterion is satisfied;
```

Figure 3: Application of Q-Learning to solve TSPWR

This algorithm is really useful when applying for different variations of the problem like TSP with refueling(TSPWR). The algorithm from the Ottoni's paper[]is given above[Fig 3].

2) Benefits

a) Easier to implement various variations like Fueling and time limitations.
b) Self-learning and can be used to handle larger scale problems.
c) Can be used to improve itself over time in a longer duration.

3) Drawbacks

a) Tuning RL parameters are quite hard.
b) Simulating and understanding how the algorithm is behaving to optimize iscan get very complex.
c) It needs a lot of testing for real world applications against very quick anddrastic changes.

F. Experimental results

After discussing the working of the algorithm, here are a few experimental results of the same. The algorithms are compared with other popular methods to solve TSP like Nearest Neighbour and Genetic Algorithm and the Optimal time has been given as well. The instances are from several TSPLIB instances for whichthe solutions are already given. The results are presented as the best out of 15 trials. As one can see, Q-Learning comes quite close to the optimal values and can be improved with additional research and work. As the number of cities increases, ACO and QL remains the nearest ones to the optimal value.

Instance	OPT	NN	GA	ACO	QL
Burma14	30.88	31.88	30.88	30.88	30.88
Eil51	426.00	513.61	432.71	451.50	437.97
St70	675.00	761.69	725.51	697.15	698.10
Berlin52	7542	8182	7739	7549	7657
Pr76	108159	130921	110605	117116	117345
KroA100	21282	24698	22546	22671	22128
D198	15780	18295	17539	15888	16431

IV. APPLICATIONS

There are many applications for TSP and by using these algorithms, we can solveTSP with different variations. Here are a few applications:

A. Maps application considering time rejection and fuel.
B. Delivery Applications like food delivery or parcels where food hotness anddelivery time are a few variables to be considered.
C. Logistics applications.
D. Planning and scheduling.
E. Manufacturing microchips where the robot has to go from one point to anotherto solder components.
F. Astronomy where telescopes are used, where the telescope has to go fromone point to another in minimum time and different variables.

V. CONCLUSION

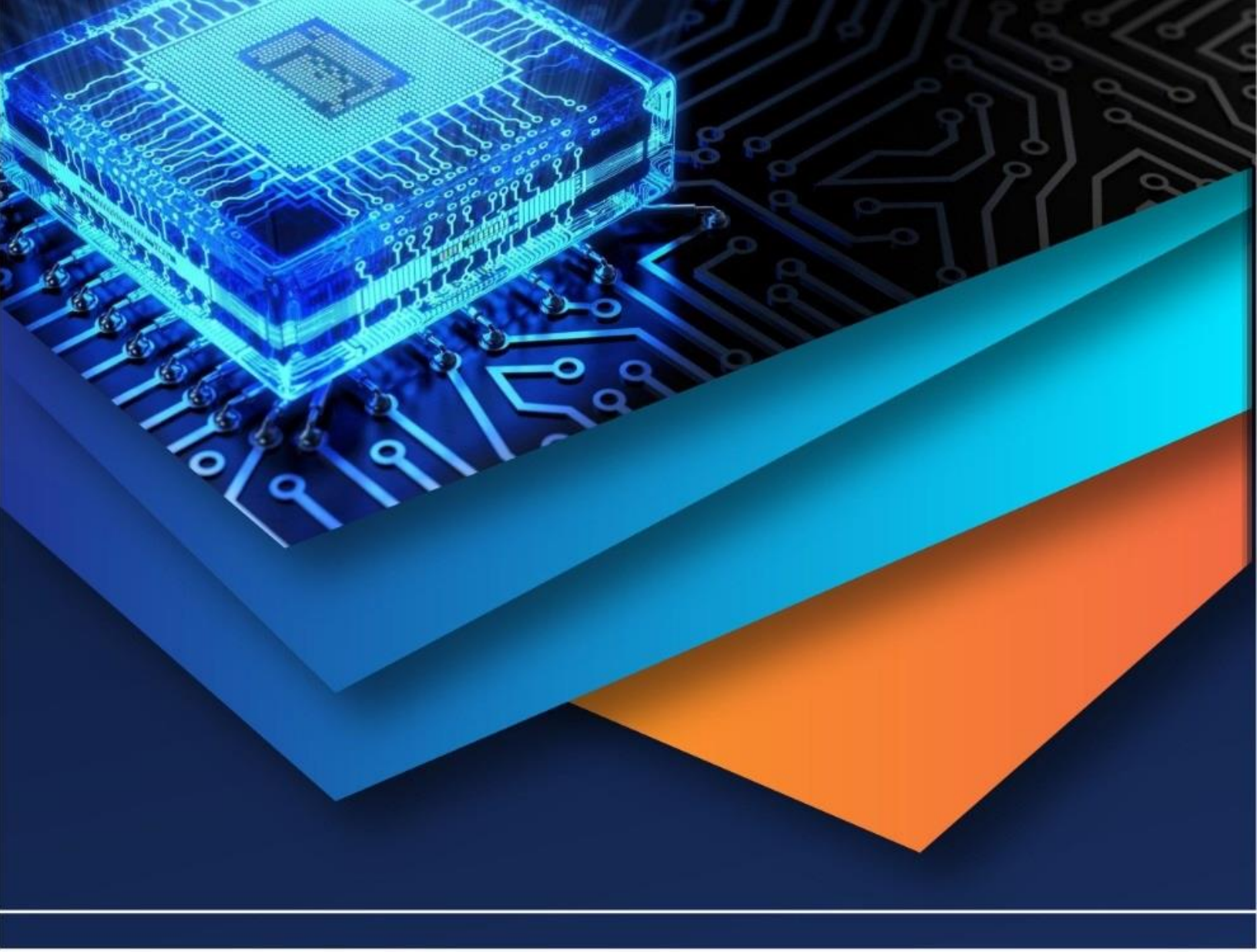
Machine learning can help us solve different complex problems included Combinatorial Optimization Problems. The algorithms derived are very useful where heuristic results. that is, results which are fast and near-optimal solutions are sufficient. As we studied through two most promising heuristic algorithms, we can notice that both of them are quite promising and can be improved if more research is done for the same. ACO can really benefit from machine learning as seen in Yuan Sun's paper [6]. Q-Learning is relatively newer, is still going through a vigorous study by machine learning scientists, and takes a lot of time and trial and error to fine-tune. This can be improved much more in the future to get better results. For the future, more variables can be introduced like Fuel + Time or traffic to seewhether these algorithms can perform well there as well.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] Nina Mazyavkina, Reinforcement Learning for Combinatorial Optimization.
[2] Rongkai Zhang, Deep Reinforcement Learning for Travelling Salesman Problem with Time Rejection Windows.DOI: 10.1109/IJCNN48605.2020.9207026
[3] Jeremiah Ishaya,A comparative analysis of the traveling salesman problem: Exact and machine learning techniques.DOI:10.30538/psrp-odam2019.0020
[4] André L. C. Ottoni,Erivelton G. Nepomuceno,Marcos S. de Oliveira,Daniela C. ,R. de Oliveira, Reinforcement learning for the traveling salesman problemwith refueling.DOI:10.1007/s40747-021-00444-4
[5] Yuan Sun, Sheng Wang, Yunzhuang Shen, Xiaodong Li, Andreas T. Ernst, Michael Kirley, Boosting Ant Colony Optimization via Solution Prediction and Machine Learning

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)