Route Optimization

Reinforcement Learning

TABLE OF CONTENTS

01

PROJECT

Description and goals

Vehicle Routing Problem (VRP)

- A number of customers located at different places desire a certain number of goods that can be delivered from a single depot.
- Determine the # of routes and paths in each route that will maximize the total distance travelled by all the vehicles in supplying all the demands.

Travel Salesman Problem (TSP)

A special scenario under VRP. The goal of this project is to find the optimal route to visit a set of destinations and returning back to the starting point

02

DATASET

Source and characteristics

DATASET

- The database was randomly generated, creating different destination locations, and then grouping them into subgroups according to the number of cities that the salesman must travel on a trip.
- The size of the training database is 1,000,000 samples of trips with 20 random destinations each. The validation dataset contains 1,000 samples of trips.

Trip	LOCATION
Location 1	(x1, y1)
(...)	(...)
Location 20	(x20, y20)

03

MODELS

Algorithms and models used

CLASSICAL HEURISTICS APPROACH

Sweep Algorithm is a 2-phase heuristic algorithm.

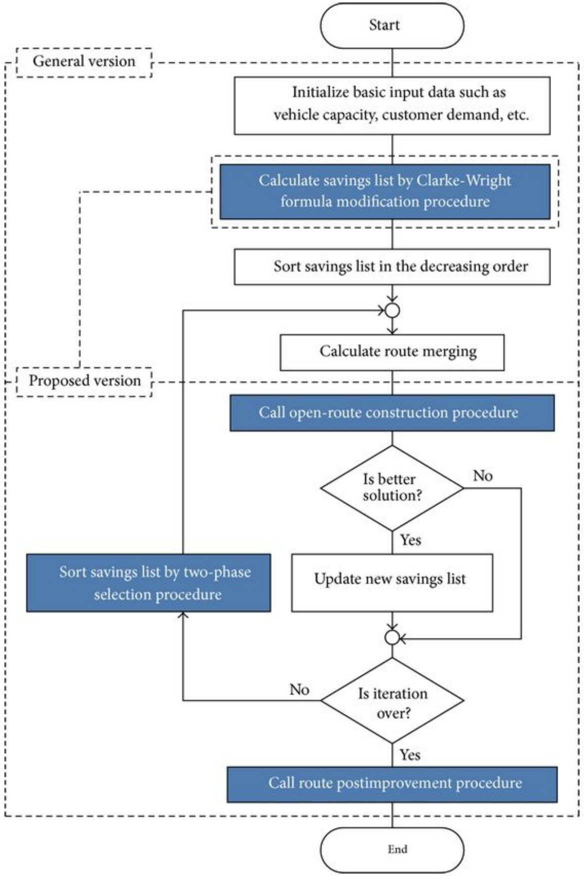
It first creates clusters of routes and then solves each cluster as a TSP problem.

It follows the steps:

- Step 1. Route Initialization:
  - Choose an unused vehicle k.
- Step 2. Route Construction:
  - Starting from the unrouted vertex having the smallest angle, assign vertices to the vehicle k as long as its capacity or the maximal route length is not exceeded.
  - If unrouted vertices remain go to Step 1.
- Step 3. Route Optimization:
  - Optimize each vehicle route separately by solving the corresponding TSP.

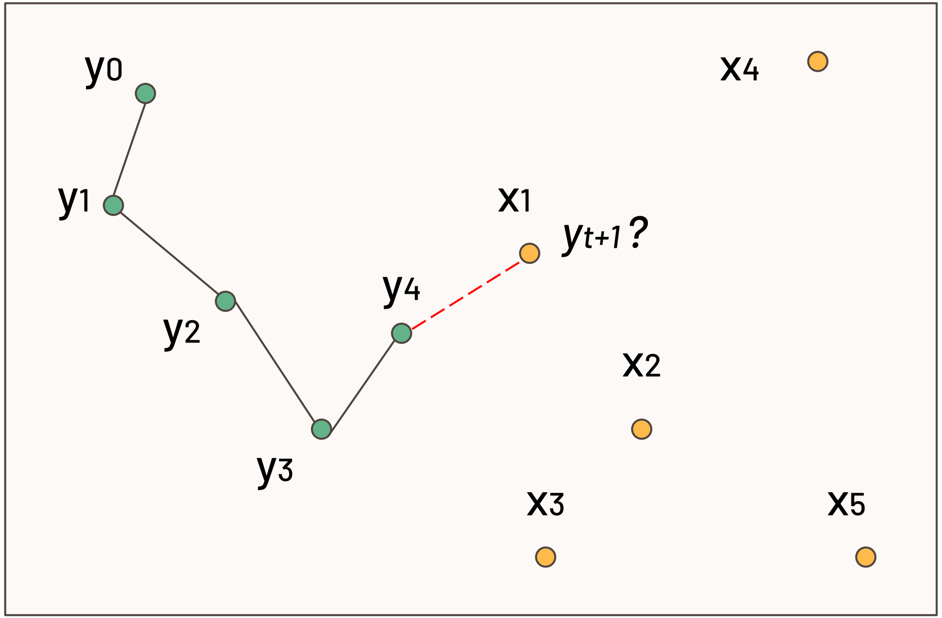
CLASSICAL HEURISTICS APPROACH

Savings Algorithm by Clark & Wright is another widely used heuristic algorithm.



DEEP RL APPROACH

- We made a meta-heuristic algorithm in which we use neural networks, we decided to select the heuristic that best fits the problem.



$Yt : \{y_0, y_1, y_2, y_3, y_4\}$

$Xt: \{ x_1, x_2, x_3, x_4, x_5 \}$

$yt+1$: next location

DEEP RL APPROACH

- We are interested in finding a stochastic **policy π** which generates the **sequence Y** in a way that **minimizes a loss objective** while satisfying the problem constraints.
- We use the probability chain rule to decompose the **probability** of generating **sequence Y**:

$$P(Y|X_0) = \prod_{t=0}^{T} \pi(y_{t+1}|Y_t, X_t)$$

y_{t+1} : next location
Y_t: decoded sequence up to time t
X_t: available inputs to become *y_{t+1}*

- With a **recursive update** of the problem representation with the state **transition function f**:

$$X_{t+1} = f(y_{t+1}, X_t)$$

- Each component π in is computed by the attention mechanism:

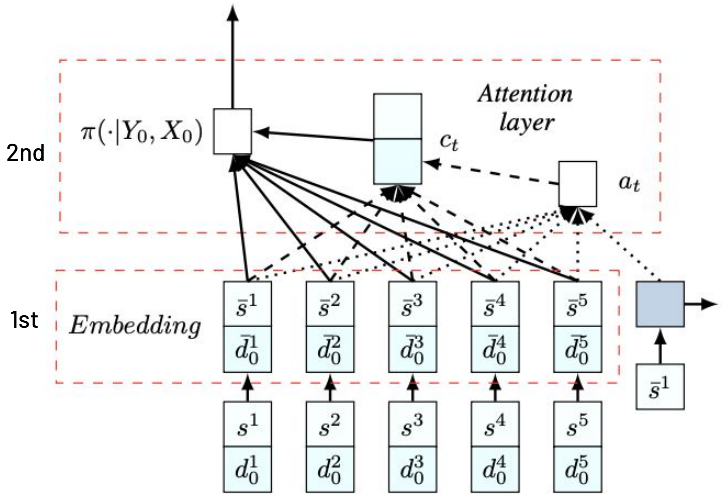
$$\pi(\cdot|Y_t, X_t) = \text{softmax}(g(h_t, X_t))$$

g: attention mechanism
h_t : vector decoded steps y0, · · · , yt
X_t: available inputs

DEEP RL APPROACH

Architecture:

- **First component:** Embedding layer that maps the inputs to a high-dimensional vector space.
- **Second component:** RNN decoder that points to the input at every decoding step. Then, the RNN hidden state and embedded input produce a probability distribution over the next input using the attention mechanism.



04

RESULTS

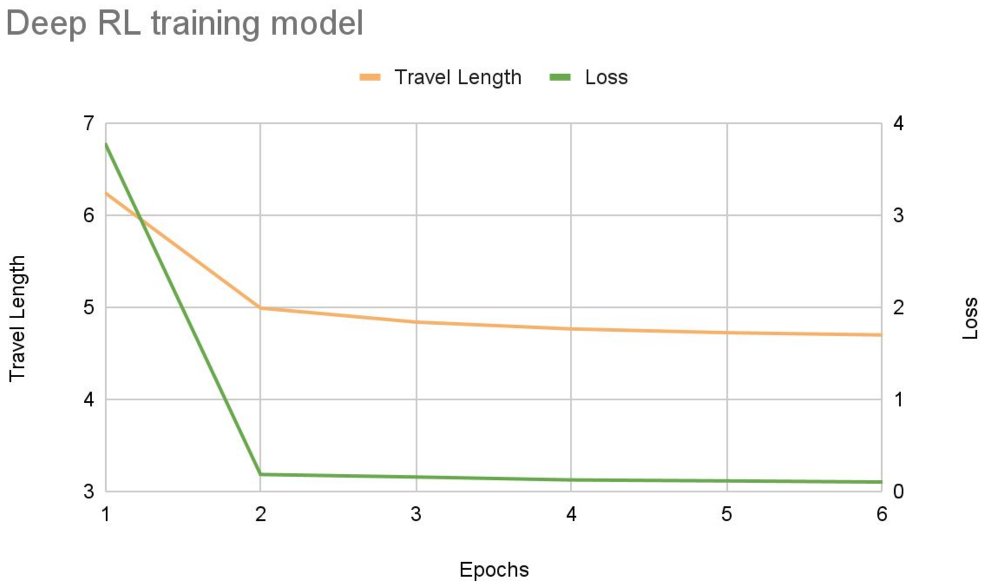
Main results and findings

CLASSICAL HEURISTICS APPROACH

Heuristic Algorithm	Average Tour Length
Sweep [1]	5.35
Clark & Wright [2]	3.93

[1]: GM74-SwRI: Gillett, B. E. and Miller, L. R. (1974). A heuristic algorithm for the vehicle-dispatch problem. Operations Research, 22(2):340-349.
[2]: CW64-PS:Clarke, G. and Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. Operations Research, 12(4):568-581.
[3]: Library: VeRyPy - https://github.com/yorak/VeRyPy

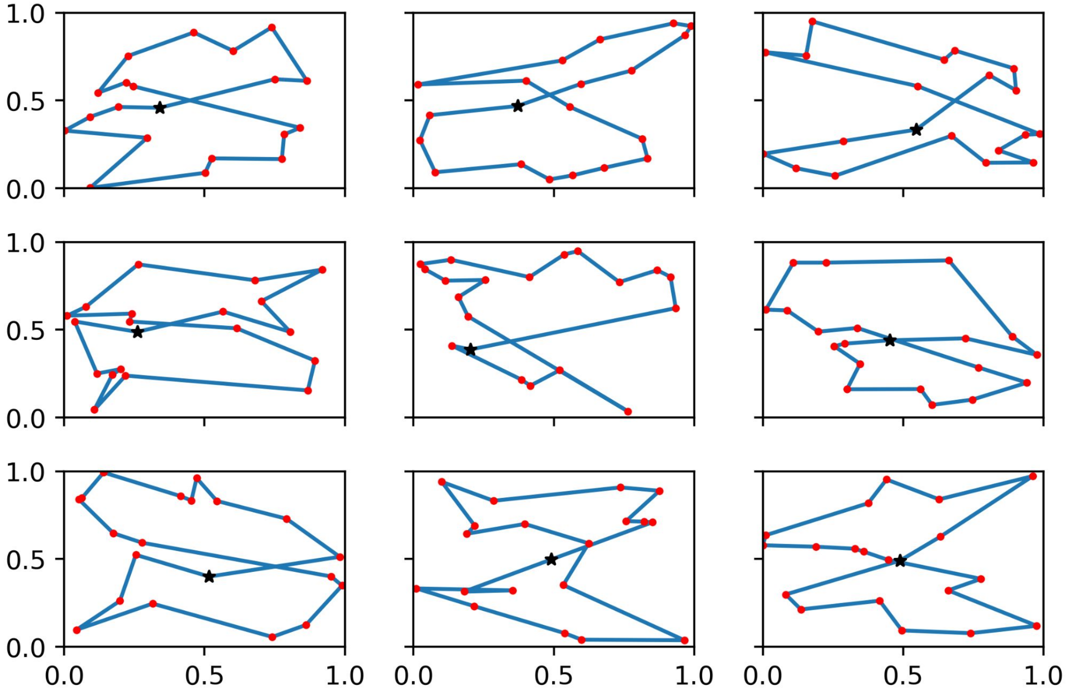
DEEP RL APPROACH



- The average travel length of the path drops sharply after the first epoch, and then slowly decreases, reaching less than 5
- The loss after the 5th epoch is very close to 0, meaning that the path length is very close to the optimal one

DEEP RL APPROACH

Final result: Average tour length **4.66**



FINDINGS

- One of the main problems we had with the deep RL approach was that of computation power. We could only run a few epochs before our RAM ran out of space. However we got decent results compared to the state of art.

State of Art final tour length	Our Model final tour length
3.97	4.66

- Results for classic heuristic are comparable to those from the Deep RL model, Clark and Wright is actually giving better results. We can reach those scores by training for more epochs.
- This experiment was done for a simple problem of 20 cities. But when the problem becomes more complex we are expecting that the RL model will perform better.

THANKS