

Goods transportation problem solving via routing algorithm

Mikhail Shchukin, *CS 820 Graduate Student, University of Regina*^{*}

Aymen Ben Said, *CS 820 Graduate Student, University of Regina*,[†]

and André Lobo Teixeira, *CS 820 Graduate Student, University of Regina*[‡]

Email: ^{*}vladimmi@uregina.ca, [†]abb549@uregina.ca, [‡]teixeira@cs.uregina.ca

Abstract

This project report outlines the ideas behind developing a graph-based heuristic-driven routing algorithm designed for a particular instance of a goods transportation problem with a single good type. The proposed algorithm solves the optimization problem of satisfying the demand of goods on a given undirected transportation graph with minimizing the estimated cost for each traversed segment of the delivery path. The operation of the routing algorithm is discussed and overall evaluation of the proposed problem solving technique is given.

Index Terms

Algorithm, graph, report, logistics.

I. INTRODUCTION

THE transportation problem is one of the well-known and hot topics both in mathematics and economics. It was first conceptualized by Gaspard Monge back in 1781 [1]. Since then, it was analyzed with various approaches, including linear programming and graph-based algorithms. In this project, a graph-based interpretation of the problem is used for reasons of better visualization of potential solutions (i.e. shipping paths) and the graph form being a comfortable environment to design the searching and routing algorithm. This project strives to deal with this problem using graph theory approach and a particular decision-making algorithm in an attempt to acquire an acceptable solution, if it exists, or give partial solution otherwise.

From our point of view, the transportation problem can be represented by a graph with nodes of 3 types: *store*, *warehouse* and a *joint*. The edges between nodes represent roads to traverse with associated *distance* and *time*. The stores and warehouses have a certain associated amount of *demand* and *supply* of goods respectively. Joints act solely as transfer points (major cities, towns, etc.). For simplicity, all goods are assumed to be of the *single type*. The problem is finding the way to provide the stores with goods available in the warehouses, *maximizing* the amount of delivered goods, while *minimizing* the transportation costs. It is assumed that the shipping is being done by a *single agent*, a truck with a certain *maximum carry capacity*, that is traversing a certain path while loading and unloading goods where needed.

In this project, we provide an algorithm that takes a pre-generated transportation problem graph with the described characteristics and returns the following key elements of the proposed solution:

- Truck path with shipping and restocking segments
- Algorithm log with the following details:
 - shipping and restocking segments explained
 - routing decisions made for next segment (if applicable)
 - decision cost (heuristic value)
 - Remaining total amount of demand or supply at the termination of algorithm
- Total runtime of the algorithm

In the following sections we describe the testing environment for such algorithm, the mode of operation of such algorithm and the experimental results.

II. TESTING ENVIRONMENT - GRAPH GENERATOR

First of all, in order to provide a capability of reproducing the experimental results of using the suggested algorithm, it is important to have a testing environment that can produce a set of various transportation problem instances (graphs) of a varying scale to serve as an input to the solving algorithm. The suggested testing environment attempts to pseudorandomly generate a transportation problem graph instance on each run. This allows to create a wide variety of sample cases that are applicable to the real-life instances of goods transportation problem.

The graph generator requires the following parameters to operate with:

- Total number of nodes to generate
- Maximum number of edges to spawn per node

- Total number of stores to generate
- Total number of warehouses to generate
- Types of goods to generate (in scope of this project, this defaults to 1)
- Total supply and total demand per each good type

Given these parameters, the graph generator proceeds to create a requested number of nodes. Each node is created as a joint by default. Then, the algorithm pseudorandomly decides which nodes are converted into stores and warehouses. To keep the graph realistic, each node receives pseudorandomly assigned coordinates of X,Y, where $X,Y \in [0, \text{MAX_MAP_SIZE}]$, where MAX_MAP_SIZE corresponds to the scale of the square space where the transportation problem is defined, it defaults to 1000. Originally, it is assumed that the units of map scale are expressed in *kilometers*.

The generator then distributes goods to be supplied one by one between each warehouse until all supplies are allocated. Similarly, each store keeps incrementing the amount of demanded goods one by one until the demand is equally distributed among all stores. Equal distribution of supply and demand among stores and warehouses is assumed to avoid degenerate cases to happen (i.e. most goods are concentrated in a warehouse too close to the stores, one store demands much more goods than the others, etc.).

After the node generation is done, the algorithm spreads the required amount of edges. Each time an edge is to be created, there is a 50% chance for that to happen. To avoid disconnected graphs, the algorithm forces at least one edge to be spread. The destination node is picked pseudorandomly among the rest of the generated nodes. When the edge is created, the X,Y coordinates of the source and destination are used to obtain the *distance between nodes*. Then, an *average assumed velocity* is picked pseudorandomly within the range of [40,100]. These bounds are assuming that the velocity is expressed in *kilometers per hour*, where lower bound corresponds to speed limits of bigger cities and upper bound resembles the highways. Based on the average assumed velocity of the edge, the time is then calculated as

$$time_in_minutes = (distance_between_nodes / assumed_average_velocity) \times 60$$

with floating point of it truncated. The time and distance are the two key parameters that affect the later decision-making, so they are kept in the graph as the background information.

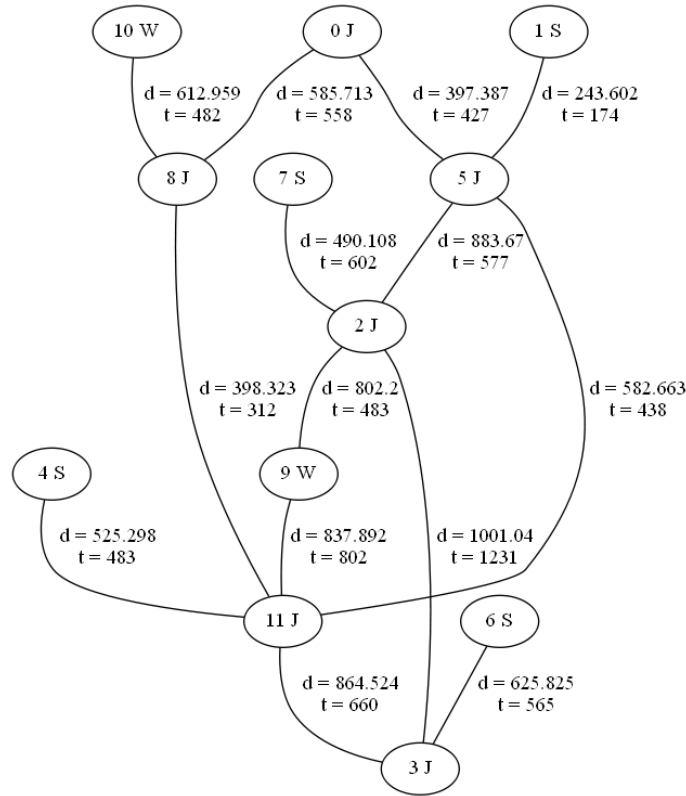


Fig. 1. An example of transportation graph produced by the graph generator

When edges between nodes are created, the graph generator produces a simple text file, that has the problem-related information associated with nodes and edges. Notably, the syntax for such text file is specifically designed to follow the syntax of WebGraphViz, the online graph visualization tool that can produce and display the graph, based on the plain text graph description with all embedded information as a picture [2]. A particular example of such visualization can be seen on Fig. 1.

Evidently, the graph generator provides a fairly connected graph that represents a particular instance of a transportation problem. Each node shows a tuple of $\{id, type\}$. Node 0 is the first generated node and the total number of transportation nodes in this particular example is 12. To generate this example, the graph generator is initialized with 2 edges to spread from each node, 2 warehouses and 4 stores to be placed. The total goods demand is set to 90 and the total supply of goods is set to 100, which is spread equally across stores and warehouses respectively. Note, that in order to reduce the visible complexity of the picture, the visual part does not contain all the problem parameters, such as supply or demand, straight on it. Nevertheless, all crucial problem parameters are still stored in the source file structure, as can be seen on Listing 1 below.

```
graph G {
"0" [label="0 J", type=0, supply=0, demand=0]
"0"--"5"[label=" d = 397.387\n t = 427", distance=397.387, time=427]
"0"--"8"[label=" d = 585.713\n t = 558", distance=585.713, time=558]
"1" [label="1 S", type=1, supply=0, demand=22]
"1"--"5"[label=" d = 243.602\n t = 174", distance=243.602, time=174]
"2" [label="2 J", type=0, supply=0, demand=0]
"2"--"9"[label=" d = 802.2\n t = 483", distance=802.2, time=483]
"3" [label="3 J", type=0, supply=0, demand=0]
"3"--"2"[label=" d = 1001.04\n t = 1231", distance=1001.04, time=1231]
"4" [label="4 S", type=1, supply=0, demand=22]
"4"--"11"[label=" d = 525.298\n t = 483", distance=525.298, time=483]
"5" [label="5 J", type=0, supply=0, demand=0]
"5"--"2"[label=" d = 883.67\n t = 577", distance=883.67, time=577]
"5"--"11"[label=" d = 582.663\n t = 438", distance=582.663, time=438]
"6" [label="6 S", type=1, supply=0, demand=23]
"6"--"3"[label=" d = 625.825\n t = 565", distance=625.825, time=565]
"7" [label="7 S", type=1, supply=0, demand=23]
"7"--"2"[label=" d = 490.108\n t = 602", distance=490.108, time=602]
"8" [label="8 J", type=0, supply=0, demand=0]
"8"--"11"[label=" d = 398.323\n t = 312", distance=398.323, time=312]
"9" [label="9 W", type=2, supply=50, demand=0]
"9"--"11"[label=" d = 837.892\n t = 802", distance=837.892, time=802]
"10" [label="10 W", type=2, supply=50, demand=0]
"10"--"8"[label=" d = 612.959\n t = 482", distance=612.959, time=482]
"11" [label="11 J", type=0, supply=0, demand=0]
"11"--"3"[label=" d = 864.524\n t = 660", distance=864.524, time=660]
}
```

Listing 1. The underlying textual form of the graph

This testing environment combines the commodity of reproducible results, configurable problem scaling and the power of visualization through external tools. In the following sections, we refer to Fig. 1 as a particular case of using this testing environment.

III. THE ROUTING ALGORITHM

Given an instance of the goods transportation graph, the routing algorithm requires the following parameters to attempt at solving the problem:

- Truck initial position (defaults to node 0, which is the first generated node)
- Truck initial load of supplies (defaults to 0, we assume here that the truck always starts empty)
- Truck maximum carry capacity (MAX_CAPACITY as referred below)
- Resupply threshold T (level of current supply below which a truck restocks at a warehouse, not proceeding to stores)

With these parameters set, the truck attempts to follow the least expensive path, which is recorded as a solution. The heuristic cost that is used in determination of the least expensive edge to traverse is defined as follows:

$$edge_cost = time + distance$$

Accordingly, at the beginning of each segment of the traversed path the truck driver decision-making works as follows:

- If the position of the truck is at a warehouse node with a capacity $< T \times \text{MAX_CAPACITY}$: it goes to the nearest warehouse which is the current node and restocks right there.
- If the position of the truck is at a joint node with a capacity $< T \times \text{MAX_CAPACITY}$: it goes to the nearest warehouse to restock with goods.

- If the position of the truck is at a store node with a capacity $< T \times \text{MAX_CAPACITY}$: it goes to the nearest warehouse to restock with goods.
- If the position of the truck is at a warehouse node with a capacity $\geq T \times \text{MAX_CAPACITY}$: it goes to the nearest store to drop goods.
- If the position of the truck is at a joint node with a capacity $\geq T \times \text{MAX_CAPACITY}$: it goes to the nearest store to drop goods.
- If the position of the truck is at a store node with a capacity $\geq T \times \text{MAX_CAPACITY}$: it goes and drops goods at the current node.

Note that for the sake of simplicity, we assume $T = 0.5$ for the particular instance of the graph shown in Fig. 1.

The algorithm treats each node according to their type with 2 parameters: supply and demand. For warehouse nodes, the demand is always 0; for stores, the supply is always 0; for joints, both supply and demand are 0.

The truck keeps going back and forth between the stores and the warehouses until either all warehouses deplete their supply (which indicates an *overconstrained problem* with only a partial solution available) or there remains no store still waiting for goods (which indicates a *successful termination* of the algorithm with all goods shipped to the stores).

IV. CONCLUSION

After observing the fundamental operation principles of the proposed algorithm, it is important to note that such algorithm is capable of functioning with a wide spectrum of transportation problem instances. Under experimental conditions, the given algorithm launched on the demonstrated graph with variable truck capacity produces economically useful and mathematically reasonable results. As seen from Table I, with increasing the truck capacity, the number of produced path segments and the cost of the overall path decreases.

TABLE I
ROUTING ALGORITHM EXPERIMENTAL RESULTS

Truck Capacity	Number of path segments	Overall path cost
10	22 segments	68360.5
15	16 segments	48142.7
20	14 segments	42781.1
22	13 segments	40764.5
23	9 segments	26577.2

However, it is notable that the change is not quite proportional. As seen on Fig. 2 below, changing the truck capacity produces a slope-like response in the cost reduction and in the number of path segments produced.

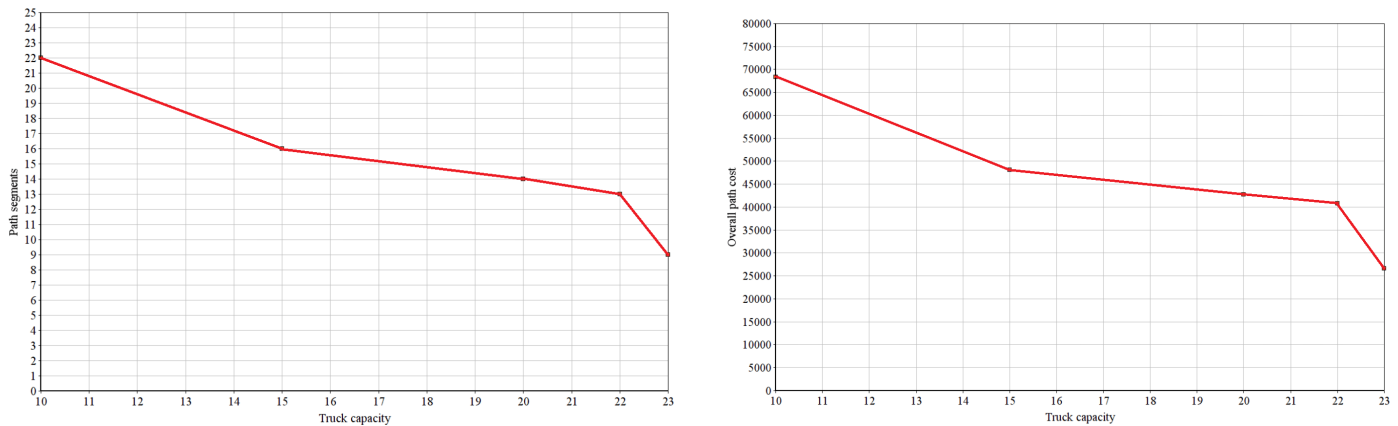


Fig. 2. Left: Truck capacity increase reduces the number of shipping path segments. Right: Truck capacity increase gradually reduces the overall path cost

Given such statistics, it is possible to estimate how much each added unit of capacity corresponds to the decrease in the overall solution cost for this particular instance of the problem. This directly translates into the economic reality of any business involving transportation of some items between certain locations on the map. Such algorithm is thus capable to provide not only a solution to the given problem with set parameters, but also *an ability to estimate the potential alternatives to a given solution with different parameters*.

What if one had to pick a bigger truck? How could extra expenses on deploying another truck be covered by saving on the number of trips between the map locations? What is the most efficient size of the truck, that can solve a given transportation problem on a given map? All these questions can eventually be answered by running the routing algorithm provided here

several times with tweaked parameters. Basically, by solving the given problem with different parameters and combining the resulting data together, this algorithm can be used as a *cost estimation tool*.

There are, however, major improvements potentially available to this algorithm. For example, the cost estimation function could involve more real-life factors affecting any transportation, such as: truck fuel consumption, desirability of paths (major city roads might not allow truck traffic coming through at certain times), required stops (drivers taking breaks and changing shifts) and more. Another improvement to deploy could be allowing multiple types of cargo, which considerably increases the complexity of decision-making and solving the transportation problem.

In the described algorithm, the decision-making heuristic of combined time and distance is only applied to selecting next place to go, so the overall cost is minimized per each segment of the solution path. However, sometimes the transportation problem can involve prioritizing time over cost, cost over time or other decision factor. Thus, wisely balancing the decision-making in the described routing algorithm along with dealing with various shipment types becomes the major design difficulty to overcome.

APPENDIX A TOOLS USED

Appendix two text goes here.

REFERENCES

- [1] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.