Gregory Knapp

CS 312

Dr. Grimsman

10/19/2021

Project 3 Network Routing Solver

Time Complexity Analysis of Project 3 – Breakdown by Implementation

The time complexity for each function found in NetworkRoutingSolver.py of the Project 3 code is as follows and pulled from the comments for each function in the code.

Time Complexity for Array Queue Implementations:

Delete_min - lines 31 - 36

```
# delete_min with unsorted array. With the unsorted array implementation,
delete_min iterates through every
# node in the array to find out which has the shortest distance in the dist
array, saves the index, and
# removes and returns the node that was popped from the queue. Due to the
linear nature of arrays, there is
# not a better way to find the min without checking every value, so the
implementation of this function
# is time complexity O(n)
```

Insert - lines 59 - 61

```
# Insert with unsorted array - This function simply uses the Python
list.append() function to add the next node
# to the array. Because append adds the node to the end of the array, and the
array is unsorted meaning that no
# work post appending needs to be done, the time complexity of insert() is
O(1) time.
```

Decrease_key – lines 66 - 68

```
# decrease_key - Decrease key does not do anything in the array
implementation of the priority queue, so there
# is no time complexity to report for this function. It is called in the
algorithm no matter which queue type is
# used, but is simply passed when the array implementation is in use.
```

Make queue – lines 72 – 75

```
# Make queue - The initial make queue function takes all of the nodes in the
network and adds them to the
# unsorted array using the insert() function. Insert runs at O(1) time, but
because every node must be added to the
# array through a for loop, we call insert() n times where n = number of
nodes in network. This makes the time
# complexity of make queue = O(n)
```

Size - lines 81-82

```
# A helper function to get the size of the array currently. Len(object) in python returns the integer of the # size of the list with O(1) time, so the time complexity of this is O(1)
```

Time Complexity for Heap Queue Implementations

Delete_min - lines 94 - 100

```
# delete_min function using binary heap implementation. This is is one of the
two core functions of the binary
# heap implementation. Because the node with the shortest distance is bubbled
up to the top with decrease_key,
# finding the minimum value is very fast, O(1) since we know it will be at
index 0. However, delete_min's overall
# complexity is not O(1). Because the parent node is removed, we must re-
build the tree using the last node in
# the array and bubble that node downwards until the heap is fixed. This
bubbling takes approximately O(log n)
# since it does not iterate through every node in the tree, just two per
iteration, resulting in a final
# time complexity of O(log n)
```

Insert - lines 193 - 196

```
# The insert function is similar to inserting on the array implementation,
but does need to do some work after
# appending the node. If the node added is smaller than previous nodes added,
it needs to bubble upwards until the
# heap order is restored. For this, insert is not O(1) time, but O(log n), as
it relies on the other queue
# function decrease_key to perform the bubbling up, which has a complexity of
O(log n)
```

Decrease_key - lines 207 - 212

```
# Decrease_key is the other core function of the heap priority queue.
Whenever a distance value is updated, either
# when the node is initially inserted to the tree, or in the main Dijkstra's
algorithm when distances are counted,
# decrease_key is used to update the value in the heap and fix the heap as
necessary. Similar to delete_min, except
```

the values are traveling up the tree, not down, since the key is being
decreased. The bubbling action is the same
where in the worst case, a node can travel from the very lowest level of
the heap to the very first parent node.
This worst case scenario dictates the run time of the function, making it
O(log n) time.

Make queue - lines 259 - 262

Make_queue works the same as in the array implementation of the queue,
however, the time complexity is slightly
different. It iterates through every node in the array, a total of n times,
which is the same. However,
the insert function has a time complexity of O(log n) because it relies on
decrease_key, which is different from
the O(1) insert of the array implementation, making make_queue here O(n log
n) time.

Size - line 267

Size() is the same as the array queue version, just returning the number of nodes in the heap array at O(1) time.

Time Complexity for main Dijkstra algorithm functions

GetShortestPath - lines 280 - 285

Get shortest path is the function used to iterate through the prev loop and save the path and edges to be
displayed on the GUI. As mentioned in the comments inside the function, it is all O(1) functions except for the
while loop that continues until the final node is found. Since the shortest path will always be a small fraction
of the entire set of nodes (unless the graph is a line of points) then the time complexity is not quite O(n)
since the while loop will never have to iterate through every node, so I would more accurately call it
O(s), where s is the number of nodes in the shortest path between source and dest.

$Compute Shortest Path-line\ 330-356$

```
# computeShortestPaths is the code for the main Dijksta's algorithm. The
variable passed in from the GUI determines
# which queue implementation to use, then runs the same algorithm for both.
All the same functions are called on
# either queue type, with different results time complexity wise. (e.g.
decrease_key is called for the array queue,
# but does nothing within the implementation).
#
# As has been discussed in class, the complexity of Dijkstra's algorithm can
be simplified to the following:
# O(|V| (cost to insert + cost to delete min) + |E| (cost to decrease key))
```

Empirical Analysis of Project 3 Code

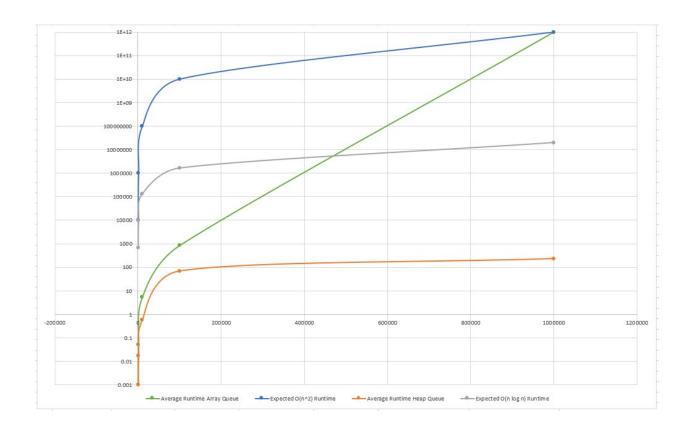
The first table is the run time of the Array Implementation (table is split into two screenshots to fit the word document).

Array Implementatio	n						
N # of Points	Test 1	Test 2	2	Test 3		Test 4	
1	0.001	l	0.001002		0.001		0.001001
10	0.052012	2	0.048504	0.0	52013		0.050013
100	00 5.485227	7	5.48367	5.3	97827		5.193173
1000	00 833.707143	3	824.090715	823.4	61168	8	395.829108
10000	00 1.00E+12	2	1.00E+12	1.0	0E+12		1.00E+12
				-			
Test 5 A	verage Runtime		Expected O(n^2)		Seeds Used		
0.001	0.00	10006		10000	55, 80	, 99, 25,3	
0.051012	0.05	07108		1000000	60, 91	, 4, 19, 71	
5.412509	5.39	44812		100000000	88, 7,	15, 29, 33	
891.377459	853.69	31186		10000000000	44, 13	, 28, 79, 105	
1.00E+12	1.0	0E+12		1E+12	N/A		

The second table is the Heap Implementation runtime and expected times table.

Heap Implementat	ion							
N # of Points		Test 1	Test	2	Test 3		Test 4	
	100	0.001		0.001001		0.001		0.001
	1000	0.017004		0.017022	0.0	17004		0.017004
10	0000	0.588142		0.591133	0.5	89143		0.585142
10	0000	75.920164		64.835658	65.02	77701		75.440055
100	0000	229.064501		235.012054	228	71036		237.89456
Test 5	Aver	age Runtime		Expected O(n	log n)			
0.001		0.001	.0002		664.385619	55, 80	, 99, 25,3	
0.017004		0.017	0076		9965.784285	60, 91	, 4, 19, 71	
0.585693		0.587	8506		132877.1238	88, 7,	15, 29, 33	
66.045932		69.4539	1582		1660964.047	44, 13	, 28, 79, 105	
230.32601		232.20	1497		19931568.57	39, 8,	96. 67. 103	

Lastly, the graph made from the two tables is as follows. The graph plots the average runtime for each test result for each implementation, along with the expected runtime for both implementations, resulting in 4 series being shown.



Empirical Analysis - Conclusions

As is seen in the graph above, at all levels of input, the heap queue outperforms the array queue implementation. The margins are extremely small at the lower levels but become much more significant at 10,000 points and especially 100,000 points. The tests were not run for the array queue for 1,000,000 points, but it is safe to expect that there would be an even greater margin of difference between the two implementations.

There are some quirks with the graph that should be addressed.

Firstly, both the array and heap queues outperformed the expected runtimes for each of the expected times. This happened in my previous process, and it is safe to say that this "time gain" is due to the speed of the computer in handling many of the simple calculations that take place in the algorithm. I don't believe this disproves the time complexity of either of the algorithms, because the way the code is set up, the array queue uses a O(n) function across every node in the queue, resulting in the O(n^2) expected time. That I am certainly confident in, which leads me to believe that the difference between expected and average is just a result of hardware.

The other quirk that came up during my testing was the sudden jump between the 10,000 point and 100,000 point tests for the heap queue implementation. The array queue results do not concern me, as they seem to fit the expected pattern of a dramatic increase in runtime with each factor of 10 added to the input. Dr. Ventura mentioned on slack having a run time of about 5 seconds for 10,000 array queue

test, which matches my output, so I am confident in the array queue results being fair. The heap queue however jumps from a sub 1 second time for the tests up to 10,000 points, then suddenly goes up to 65 to 75 seconds per test. This is still fitting of the expectation that the heap queue will severely outperform the array queue, but the jump seems a little sudden.

I looked around at slack and talked to other students and saw a very wide variety of results, ranging from like mine to significantly faster or slower, making it hard to be sure about what the deal was. I double checked my implementation and noticed I was using dictionaries to store the index of the heap array as well as for the distance array, so I switched those out for normal python arrays, but did not see any significant change in result.

When I would run the 100,000 or 1,000,000 point tests, even just for the heap queue, the GUI window would immediately stop responding and load an error. If I left it alone for a while, it would eventually run to completion and give me the result, but this leads me to believe that maybe it is the unresponsiveness of the GUI that is resulting in the strange run times for the heap. This is just speculation, and I should in the future maybe try the test without the GUI and make my own script to generate points and solve for the array.

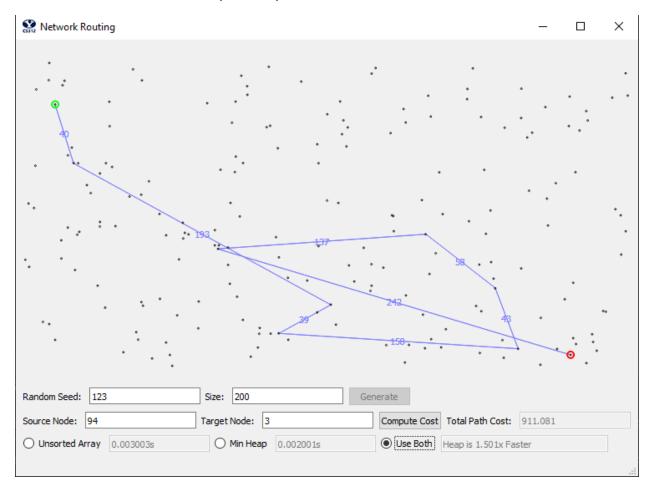
The strange result was worth mentioning, and I am also prone to leaving this up to a hardware issue as well that may have affected performance on the very large tests. I have been noticing my computer having difficulty running high memory usage processes, if the RAM shows more than 50% memory usage, things will start crashing and failing. I'm hoping to take my computer into a repair shop and see if there are any hardware issues, especially before the next project so that I can get a better result, if the hardware was an issue in the first place.

Working Screenshots

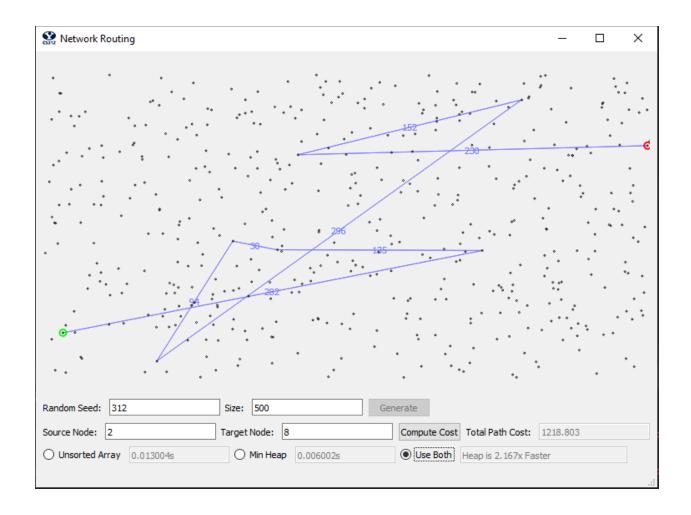
Screenshot #1 - Random seed 42, Size 20, Node 7 to Node 1



Screenshot #2 – Random seed 123, Size 200, Node 94 to Node 3



Screenshot #3 – Random seed 312, Size 500, Node 2 to Node 8



Source Code Files

NetworkRoutingSolver.py

```
def delete min(self, dist):
   def size(self):
class ArrayQueue(Queue):
```

```
self.nodes.remove(nodetoremove)
def make queue(self, network nodes, dist):
```

```
class HeapQueue(Queue):
        if self.size() == 0:
```

```
if right child index < self.size():</pre>
            parent swapped = False
                     self.array indices[parent node.node id] =
left child index
```

```
parent swapped = True # Update boolean value - 0(1)
self.decrease key(node, dist)
```

```
if child index == 0: # If statement comparison - O(1)
   self.nodes[parent index] = child node # Set value in array -
    self.array indices[child node.node id] = parent index # Set
```

```
def size(self):
def initializeNetwork(self, network):
def getShortestPath( self, destIndex):
```

```
path edges.append((edge to add.src.loc, edge to add.dest.loc,
```

```
queue = HeapQueue()
queue = ArrayQueue()
```

```
for node in node array:
    dist.append((float('inf'))) # Initialize value in array - O(1)
while queue.size() != 0:
```

```
Queue N/A, Heap Queue O(log n)

t2 = time.time()
  return t2 - t1
```

CS312Graph.py

```
self.neighbors.append( CS312GraphEdge(self,neighborNode,weight) )
        s.append(n.neighbors)
def getNodes( self ):
```

```
from CS312Graph import *
   def decrease key(self, node, dist):
   def insert(self, node, index, dist):
class ArrayQueue(Queue):
```

```
nodetoremove = self.nodes[minindex]
    self.nodes.remove(nodetoremove)
def decrease key(self, node, dist):
```

```
class HeapQueue(Queue):
       self.nodes.remove(heap end node)
       if self.size() == 0:
```

```
if left child index < self.size():</pre>
    left child node = self.nodes[left child index] # Get value
```

```
self.nodes[parent index] = right child node # Set value
```

```
# This worst case scenario dictates the run time of the function, making
def decrease key(self, node, dist):
       child node = self.nodes[child index] # Get value from heap array
        child dist = dist[child node.node id] # Get value from dist
            self.array indices[parent node] = child index # Set value in
```

```
def getShortestPath( self, destIndex):
```

```
def computeShortestPaths(self, srcIndex, use heap=False):
    t1 = time.time()
        queue = HeapQueue()
        queue = ArrayQueue()
```

```
t2 = time.time()
```

PYQT_VER = 'PYQT5'