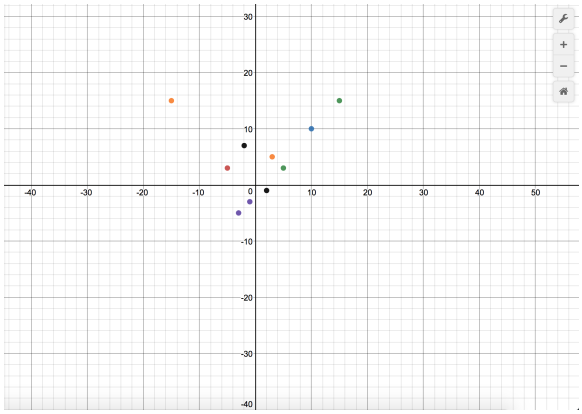
An Analysis of the Nearest Neighbor Problem

Bailey Yu, 861271666

April 29, 2018

1 Introduction

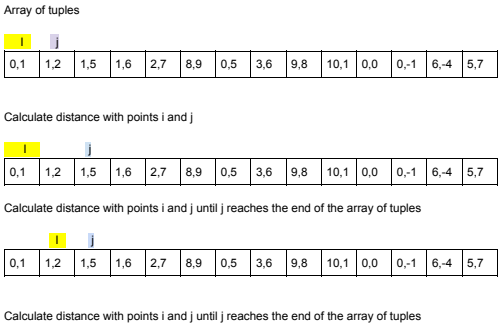
We are trying to find a way to find the pair of nearest coordinates given a finite amount of coordinates. We will be using both the brute force method and the divide and conquer method. The brute force method is a naive solution where we compare each individual coordinate to every other coordinate, which will drastically increase in runtime with each additional input. The divide and conquer method however, will allow us to recursively split and solve for the coordinates, thus reducing the runtime.



2 Algorithms

2.1 Brute Force Algorithm

The brute force algorithm works similarily to selection sort in the sense it uses a nested for loop and that it will take each coordinate and generate the distance between itself and the other coordinates. It will compare the current minimum distance to the current distance and if it is less, then it will replace the current minimum distance.



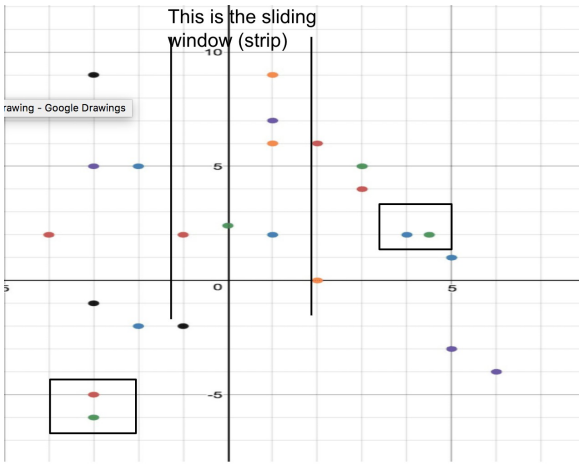
Run-time

The theoretical runtime for this is O(n^2).

2.2 Divide and Conquer Algorithm

The Divide and conquer algorithm we came up with uses two recursive statements that splits the sorted coordinates left and right until we reach base case. Then we find the min distance and compare it to the min distance of the middle split. Eventually when all the calls returns to the original function call, we will be left the the minimum distance and the cooresponding coordinates.

Size	BruteForce	Divide and Conquer
10	0.0000627s	0.0001317s
100	0.0028316s	0.000642s
1000	0.29277s	0.00867s
10000	30.7s	0.1012s
100000	3070s	1.1915s
1000000	307000s	16.712s



Run-time

The theoretical runtime of this algorithm is O(nlogn).

3 Results

Brute force will definitely take longer than divide and conquer is because comparing runtimes, O(n^2) vs O(nlogn), O(nlogn) will be considerably faster. The brute force runtime, as depicted by this chart, increases exponentially with each 10x increase to size, thus when the input is 100,000 or above, the runtime is extremely unrealistic.