



## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

### 19ITB201 - DESIGN AND ANALYSIS OF ALGORITHMS

#### CLOSEST PAIR PROBLEM :

We are given an array of  $n$  points in the plane, and the problem is to find out the closest pair of points in the array. This problem arises in a number of applications. For example, in air-traffic control, you may want to monitor planes that come too close together, since this may indicate a possible collision. Recall the following formula for distance between two points  $p$  and  $q$ .

The Brute force solution is  $O(n^2)$ , compute the distance between each pair and return the smallest. We can calculate the smallest distance in  $O(n \log n)$  time using Divide and Conquer strategy. In this post, a  $O(n \times (\log n)^2)$  approach is discussed. We will be discussing a  $O(n \log n)$  approach in a separate post.

#### Algorithm

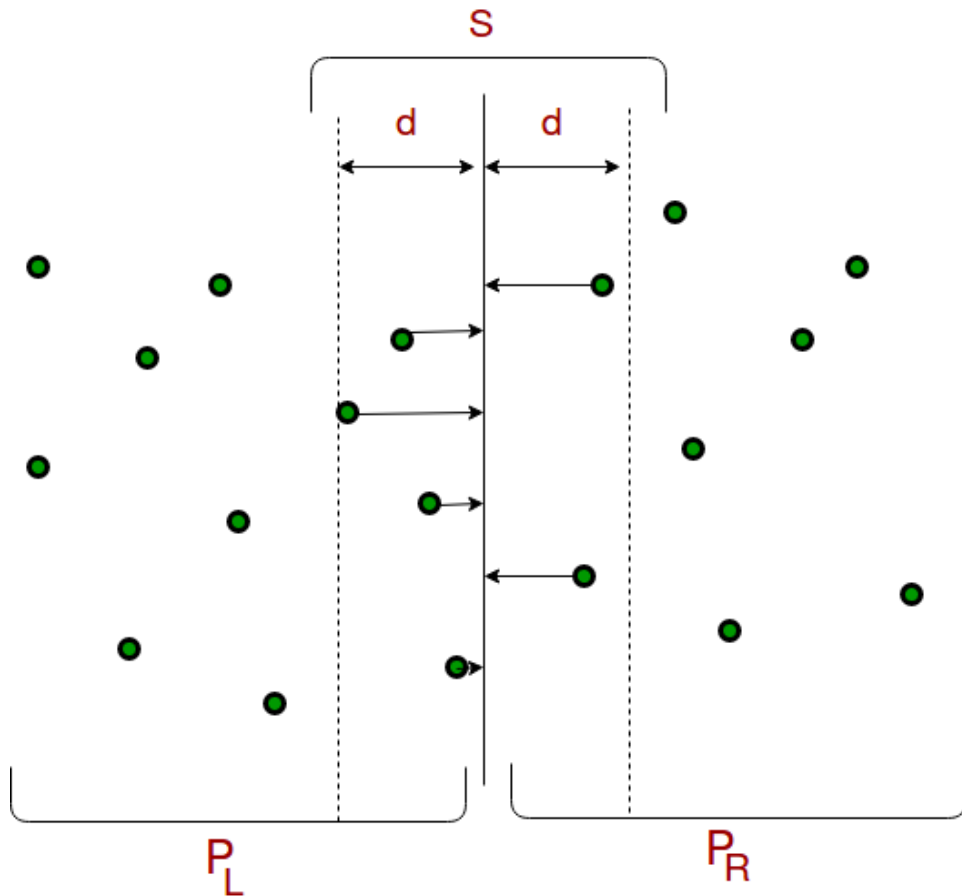
Following are the detailed steps of a  $O(n (\log n)^2)$  algorithm.

*Input:* An array of  $n$  points  $P[]$

*Output:* The smallest distance between two points in the given array.

As a pre-processing step, the input array is sorted according to  $x$  coordinates.

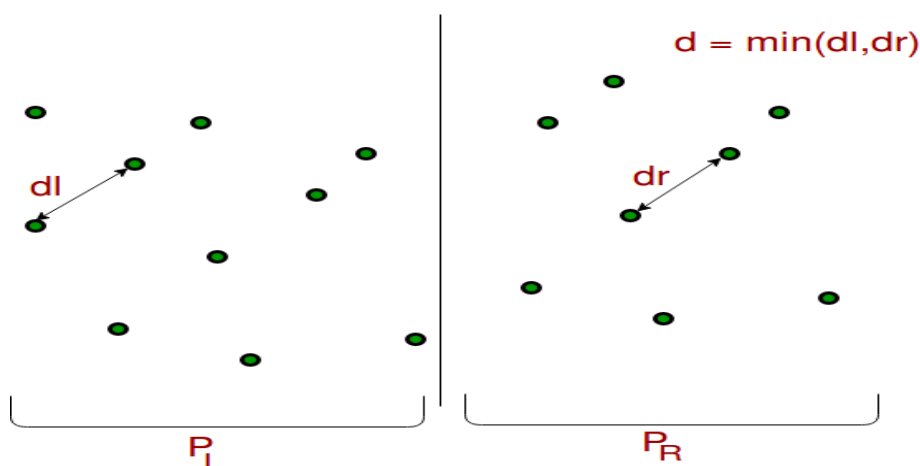
- 1) Find the middle point in the sorted array, we can take  $P[n/2]$  as middle point.
- 2) Divide the given array in two halves. The first subarray contains points from  $P[0]$  to  $P[n/2]$ . The second subarray contains points from  $P[n/2+1]$  to  $P[n-1]$ .
- 3) Recursively find the smallest distances in both subarrays. Let the distances be  $d_l$  and  $d_r$ . Find the minimum of  $d_l$  and  $d_r$ . Let the minimum be  $d$ .



4) Sort the array `strip[]` according to y coordinates. This step is  $O(n \log n)$ . It can be optimized to  $O(n)$  by recursively sorting and merging.

5) Find the smallest distance in `strip[]`. This is tricky. From the first look, it seems to be a  $O(n^2)$  step, but it is actually  $O(n)$ . It can be proved geometrically that for every point in the strip, we only need to check at most 7 points after it (note that `strip` is sorted according to Y coordinate).

6) Finally return the minimum of  $d$  and distance calculated in the above step (step 6)



7) From the above 3 steps, we have an upper bound  $d$  of minimum distance. Now we need to consider the pairs such that one point in pair is from the left half and the other is from the right half. Consider the vertical line passing through  $P[n/2]$  and find all points whose  $x$  coordinate is closer than  $d$  to the middle vertical line. Build an array `strip[]` of all such points.