Week 10: Spatial Index

03.11.2020

Objectives

In this tutorial, we will use simple sample data to illustrate the construction and usage of a few spatial index structures covered in the lecture.

Question 1: Point Data: Kd-Tree and QuadTree

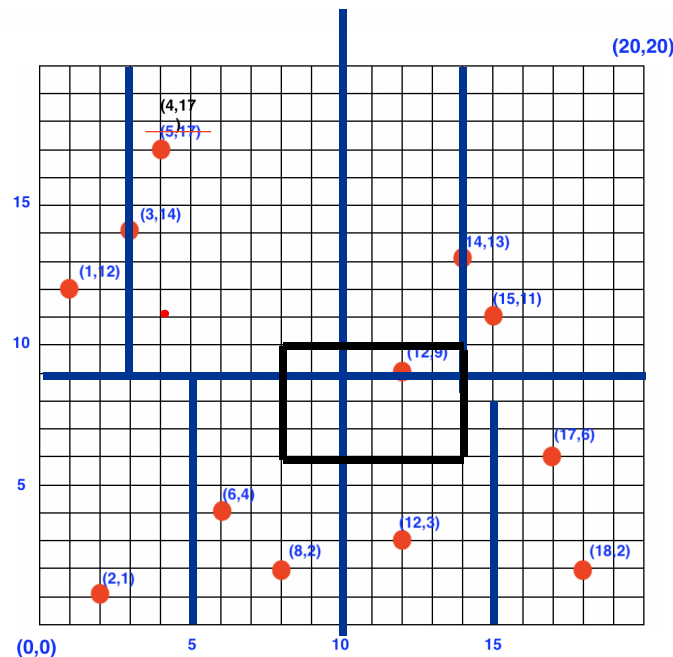Fig 1 shows a set of data points in a 2d space with y-coordinate range is [0,10] and y-coordinate range is [0,10].

Figure 1: Sample Data Points

a) Assume two records can fit in one block. Place the sample in a kd-tree. Assume the tree's root level is split along x-coordinate and each level is split along an alternate coordinate. At each level, pick a dividing value that partitions both the space and data as evenly as possible. It is OK to pick a dividing line with the same value as a point's

coordinate. In that case, the point is considered as belonging to the region smaller or equal to the dividing value.

b) A query looks for all points in a rectangle area with $8 \leq x \leq 14$ and $6 \leq y \leq 10$, which nodes/link will be examined/followed to find candidate points.

c) A query looks for the closest neighbour of point $(4, 11)$, which nodes/link will be examined/followed to find candidate point.

d) Place the sample in a quad tree, also assuming two records can fit in one block.

Question 2: Point Data: R-Tree

Fig 2 shows a set of data points and the R-Tree regions covering all the points. Fig 3 shows the corresponding R-Tree. Consider a query that looks for all points whose x-coordinate is greater than 6 and y coordinate between 2 and 5. Which nodes are examined in answering this query?
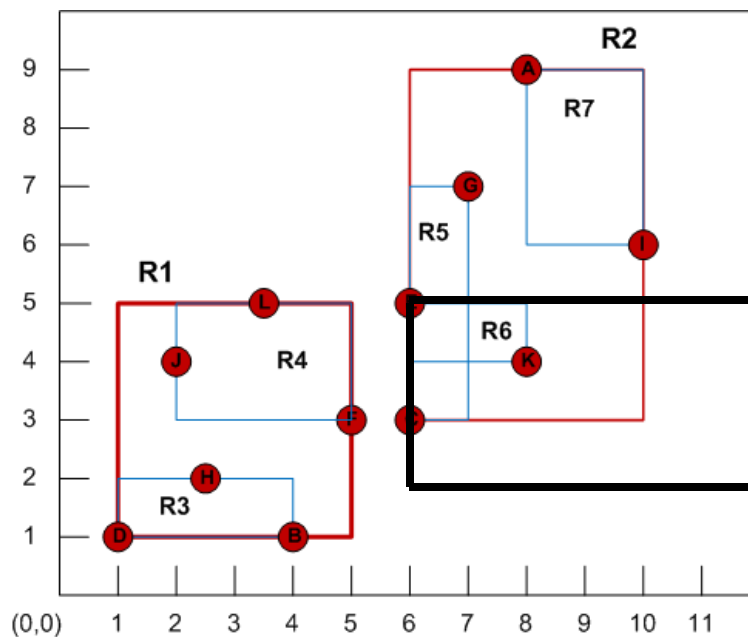

Figure 2: R-Tree Sample Data Points

R6 and R5 are intersecting with query region therefore the nodes examined are C, G, E, and K
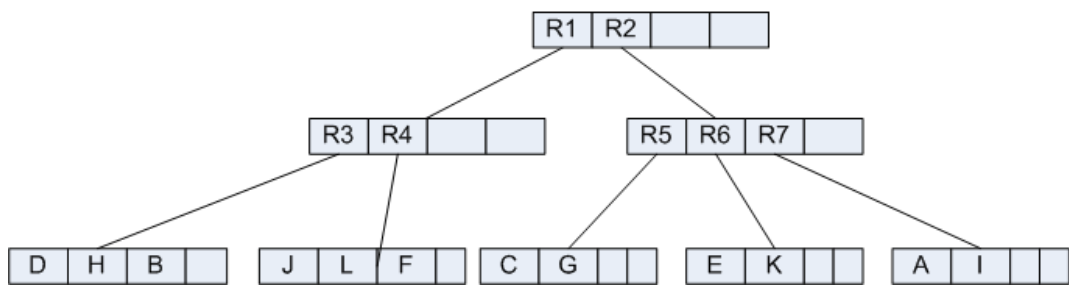
2

Figure 3: R-Tree