

Efficient Processing of Skyline Queries Using MapReduce

ABSTRACT:

The skyline operator has attracted considerable attention recently due to its broad applications. However, computing a skyline is challenging today since we have to deal with big data. For data-intensive applications, the MapReduce framework has been widely used recently. In this paper, we propose the efficient parallel algorithm SKY-MR+ for processing skyline queries using MapReduce. We first build a quadtree-based histogram for space partitioning by deciding whether to split each leaf node judiciously based on the benefit of splitting in terms of the estimated execution time. In addition, we apply the dominance power filtering method to effectively prune non-skyline points in advance. We next partition data based on the regions divided by the quadtree and compute candidate skyline points for each partition using MapReduce.

INTRODUCTION:

THE skyline operator and its variants have attracted considerable attention recently due to their broad applications such as product recommendations , review evaluations with user ratings , querying wireless sensor networks and graph analysis . However, computing a skyline is challenging today since we have to deal with big data. For data-intensive applications including similarity joins and top-k substring matching, the MapReduce framework has been considered as a de facto standard. Thus, several skyline processing algorithms , using MapReduce have been proposed. MR-GPMRS consists of the partitioning and global skyline phases. The partitioning phase of MR-GPMRS divides the data space into grid partitions and prunes the partitions that cannot contain any skyline point by utilizing the dominance relationships between grid partitions. In the global skyline phase, in every unpruned partition P, the points which are located in other unpruned partitions and may dominate a point in P are first collected and each point in the partition P is compared with the collected points to determine whether it is a global skyline point in parallel.

Technofist,

YES Complex, 19/3&4, 2nd Floor, Dinnur Main Road, R.T.Nagar, Bangalore-560032

Ph:080-40969981, Website:www.technofist.com. E-mail: technofist.projects@gmail.com