

Efficient skyline computation over an incomplete database with changing states and structures

ABSTRACT

Skyline query has been studied extensively and a significant number of skyline algorithms have been proposed, mostly attempt to resolve the optimisation problem that is mainly associated with reduction in the processing time of skyline computations. While databases change their states and/or structures throughout their lifetime to reflect the current and latest information of the applications, the skyline set derived before changes are made towards the initial state of a database is no longer valid in the new state/structure of the database. The domination relationships between objects identified in the initial state might no longer hold in the new state. Nonetheless, computing the skylines over the entire new state/structure of the database is inefficient, as not all pairwise comparisons between the objects are necessary to be performed. In tackling the above issue, this paper proposes a solution, named Δ Skyline, which aims at avoiding unnecessary skyline computations when a database changes its state and structure due to a data definition operation(s) (add or remove a dimension(s)). This is achieved by identifying and retaining the prominent dominance relationships when pairwise comparisons are performed; which are then utilised in the process of computing a new skyline set. Δ Skyline consists of two optimisation components, namely: $\Delta +$ Skyline which derives a new skyline set when a new dimension(s) is added to a database and $\Delta -$ Skyline which derives a new skyline set when an existing dimension(s) is removed from a database. To make our solution more useful, it is applied on a database with incomplete data. Extensive experiments have been conducted to evaluate the performance and prove the efficiency of our proposed solution.

Keyword: Multi-criteria decision making; Skyline queries; Incomplete database; Dynamic database; Pairwise comparisons