Automatic Performance Tuning of Spatial Indexes for Dynamic Workloads in Spatial Databases (PI: Dr. Yiu Man Lung; 2012/13)

Spatial databases are deployed to process incoming workloads on spatial objects in many location-based applications, e.g., inventory management in ports and warehouses, and the management of fleets of vehicles. It is important to tune the database performance because such applications face potentially massive workloads containing: (i) location updates that are issued from objects (e.g., containers) as they move and (ii) spatial queries that are issued by users (e.g., warehouse staff) to find objects in specified regions. A spatial database administrator can build a spatial index on the data objects for processing queries efficiently; however, the index may incur significant overhead on location updates. The conventional rule is that the administrator should build an index when the workload is query-intensive and should drop the index when the workload becomes update-intensive. This project aims at developing automatic performance tuning techniques for spatial indexes. The dynamic nature of spatial workloads poses new challenges to the traditional "build-or-drop index" approach. First, the amount of queries and updates in the workload varies along time and across different regions. This renders it difficult to decide whether to build or drop a spatial index in advance. Second, the problem is complicated by the bursts of updates and queries in the workloads. For example, bursts of updates emerge when containers/goods arrive at a port or warehouse, and bursts of queries appear in certain regions due to shift of interests, sudden events, e.g., accidents. It is expensive to build and drop an index frequently for dealing with these bursts. Third, in road network applications (e.g., fleet management), the workloads also include traffic updates on roads, e.g., congestion, closure, and re-opening of roads, which

affect travel times among different locations on the road network. Unpredictable traffic makes it challenging to tune the performance. Although extensive work has been conducted on spatial indexing, there is only little work on tuning the performance of spatial indexes. We fill this research gap and enable spatial indexes to accommodate dynamic workloads by self-tuning. This research project aims to investigate: (i) techniques for tuning the spatial index size, (ii) tuning techniques for spatial indexes in commercial spatial databases, (iii) tuning techniques for road network indexes, (iv) performance evaluation and prototype tuning tool. Real workloads will be collected for evaluating the execution time and workload throughput of our techniques. This project will benefit location-based applications and will result in index tuning techniques of relevance for commercial spatial databases.