Memory-Efficient Data Management for Spatio-Temporal Applications

Keven Richly

Abstract

The wide distribution of location-acquisition technologies means that large volumes of spatio-temporal data are continuously being accumulated. Positioning systems such as GPS enable the tracking of various moving objects' trajectories, which are usually represented by a chronologically ordered sequence of observed locations. The analysis of movement patterns based on detailed positional information creates opportunities for applications that can improve business decisions and processes in a broad spectrum of industries (e.g., transportation, traffic control, or medicine). Due to the large data volumes generated in these applications, the cost-efficient storage of spatio-temporal data is desirable, especially when in-memory database systems are used to achieve interactive performance requirements.

To efficiently utilize the available DRAM capacities, modern database systems support various tuning possibilities to reduce the memory footprint (e.g., data compression) or increase performance (e.g., additional indexes structures). By considering horizontal data partitioning, we can independently apply different tuning options on a fine-grained level. However, the selection of cost and performance-balancing configurations is challenging, due to the vast number of possible setups consisting of mutually dependent individual decisions.

In this thesis, we introduce multiple approaches to improve spatio-temporal data management by automatically optimizing diverse tuning options for the application-specific access patterns and data characteristics. Our contributions are as follows: (1) We introduce a novel approach to determine fine-grained table configurations for spatio-temporal workloads. Our linear programming (LP) approach jointly optimizes the (i) data compression, (ii) ordering, (iii) indexing, and (iv) tiering. We propose different models which address cost dependencies at different levels of accuracy to compute optimized tuning configurations for a given workload, memory budgets, and data characteristics. To yield maintainable and robust configurations, we further extend our LP-based approach to incorporate reconfiguration costs as well as optimizations for multiple potential workload scenarios. (2) To optimize the storage layout of timestamps in columnar databases, we present a heuristic approach for the workload-driven combined selection of a data layout and compression scheme. By considering attribute decomposition strategies, we are able to apply application-specific optimizations that reduce the memory footprint and improve performance. (3) We introduce an approach that leverages past trajectory data to improve the dispatch processes of transportation network companies. Based on location probabilities, we developed risk-averse dispatch strategies that reduce critical delays. (4) Finally, we used the use case of a transportation network company to evaluate our database optimizations on a real-world dataset. We demonstrate that workload-driven fine-grained optimizations allow us to reduce the memory footprint (up to 71% by equal performance) or increase the performance (up to 90% by equal memory size) compared to established rule-based heuristics.

Individually, our contributions provide novel approaches to the current challenges in spatio-temporal data mining and database research. Combining them allows in-memory databases to store and process spatio-temporal data more cost-efficiently.