Chapter XIII

Distributed Approach to Continuous Queries with kNN Join Processing in Spatial Telemetric Data Warehouse

Marcin Gorawski
Silesian Technical University, Poland

Wojciech Gębczyk
Silesian Technical University, Poland

ABSTRACT

This chapter describes realization of distributed approach to continuous queries with kNN join processing in the spatial telemetric data warehouse. Due to dispersion of the developed system, new structural members were distinguished: the mobile object simulator, the kNN join processing service, and the query manager. Distributed tasks communicate using JAVA RMI methods. The kNN queries (k Nearest Neighbour) joins every point from one dataset with its k nearest neighbours in the other dataset. In our approach we use the Gorder method, which is a block nested loop join algorithm that exploits sorting, join scheduling, and distance computation filtering to reduce CPU and I/O usage.

INTRODUCTION

With expansion of location-aware technologies such as the GPS (Global Positioning System) and growing popularity and accessibility of the mobile communication, location-aware data management becomes a significant problem in the mobile computing systems. Mobile devices
become much more available with concurrent growth of their computational capabilities. It is expected that future mobile applications will require scalable architecture that will be able to process very large and quickly growing number of mobile objects, and to evaluate compound queries over their locations (Yiu, Papdias, Mamoulis, Tao, 2006).

The paper describes realization of distributed approach to the Spatial Location and Telemetric Data Warehouse (SDW(l/t)), which bases on the Spatial Telemetric Data Warehouse (STDW), which consist of telemetric data containing information about water, gas, heat and electricity consumption (Gorawski, Wróbel, 2005). DSDW(l/t) (Distributed Spatial Location and Telemetric Data Warehouse) is supplied with datasets from Integrated Meter Reading (IMR) data system and by mobile objects location.

Integrated Meter Reading data system enables communication between medium meters and telemetric database system. Using GPRS or SMS technology, measurements from meters located on a wide geographical area are transferred to database, where they are processed and put forward for further analysis.

The SDW(l/t) supports making tactical decisions about size of medium productivity on the base of short-termed consumption predictions. Predictions are calculated basing on data stored in a data warehouse by ETL process.

**DESIGNED APPROACH**

First figure illustrates designed approach architecture. We can observe multiple, concurrently running mobile objects (query points), the Gorder (Chenyi, Hongjun, Beng Chin, Jing 2004) service responsible for processing simultaneous continuous queries over k nearest neighbors, RMI’s SDWServer and the central part of the designed data system - SDW(l/t), which is also referenced as a query manager. Communication between SDW(l/t) and query processing service is maintained with Java’s Remote Method Invocation (RMI) solutions.

Principal goal of the described approach is to distribute the previously designed system over many independent nodes. As a result we expect faster and more efficient processing of similarity join method Gorder. In the previous approach, all components shown in figure 1 were linked together on a single computer. All active processes were using the same CPU. Because of high CPU usage and long evaluation time we decided to distribute the SDW(l/t) into independent services, linked together with Java RMI technology. The most efficient solution assumes moving the Gorder service to separate computer because it causes the highest CPU consumption from all components. Other components may be executed on different computers, or on the same computer; their influence on the CPU usage is insignificant.

The designed system works as follows. First, using SDW(l/t), we have to upload a road map and meters into a database running on Oracle Server. Then we start the SDWServer, the Gorder service and as many mobile objects as we want to evaluate. Every new mobile object is registered in the database. In SDW(l/t) we define new queries for active mobile objects. Queries are also registered in the database. The Gorder service