Chapter 5.15
Context-Aware Query Processing in Ad-Hoc Environments of Peers

Nikolaos Folinas
University of Ioannina, Greece

Panos Vassiliadis
University of Ioannina, Greece

Evaggelia Pitoura
University of Ioannina, Greece

Evangelos Papapetrou
University of Ioannina, Greece

Apostolos Zarras
University of Ioannina, Greece

ABSTRACT

In this article, we deal with context-aware query processing in ad-hoc peer-to-peer networks. Each peer in such an environment has a database over which users execute queries. This database involves (a) relations which are locally stored and (b) virtual relations, all the tuples of which are collected from peers that are present in the network at the time when a query is posed. The objective of our work is to perform query processing in such an environment and, to this end, we start with a formal definition of the system model. Next, we formally define SQLP, an extension of SQL that covers the termination of queries, the failure of individual peers and the semantic characteristics of the peers of such a network. Moreover, we present a query execution algorithm as well as the formal definition of all the operators that take place in a query execution plan.

INTRODUCTION

Nowadays, the synergy between network and database management systems provides opportunities for the integration and querying of various heterogeneous sources of information, spread over an ad hoc network of peers. *The fundamental topic of this article is the context-aware processing of queries in ad hoc networks of peers through Web services.* We assume the existence of a set of peers who communicate with each other, thus forming a time varying ad hoc network of peers. For reasons of interoperability, we also assume that these peers use Web services for their interactions. Each peer has a database where (a) data can be locally stored, or (b) descriptions of data are present, in a form that allows their collection from the appropriate peers and their subsequent querying with traditional database mechanisms. The querying and/or collection of these data is dependent on the state of the peer network and on the knowledge that the peer has about this state; therefore, each time a query is posed, its processing must be adapted to this state. *In other words*, *the state of the peer posing a query and*, *most importantly*, *the state of its surrounding network constitutes the context under which the query is processed*.

Assume the case where several kinds of vehicles are driving in a highway. Each vehicle is a part of a global pervasive computing environment where computations can be performed, data can be exchanged between computing devices of the environment and information is interactively requested and presented to the users. Cars interact with each other through Web services, providing dynamically changing information regarding the vehicle's location, velocity and fuel deposit. Moreover, each vehicle comprises services that offer static information concerning its type and technical characteristics. On the highway, there exist exits to parking areas, which may include facilities such as gas stations, fast food restaurants, medical help, and shopping centers. Each one of these facilities also comprises Web services, which range from simple ones, reporting the existence of the facility, to more complex ones providing information regarding, for instance, the price lists, the availability of certain goods or the number of patients waiting for medical help. The users of the facilities of the pervasive environment, for example, the drivers of the vehicles, can obtain information by posing queries to global information space of the environment. For instance, they may be interested in obtaining information like the closest gas station with a price of gasoline under 2€/gallon, the closest Italian restaurant, or notifications for the average speed of all the cars ahead.

To facilitate the smooth operation of peers within the aforementioned environment, specific technical challenges must be addressed. A significant problem is the fact that traditional query processing must be reconsidered to adapt to the particularities of our computing environment. In this article, we are specifically interested in the problem of formally defining a declarative query language that enables the posing of queries over an ad hoc network of peers as well as the introduction of a mechanism for the transformation of declarative database queries to query execution plans.

First, we start with the theoretic formulation of the problem. We construct a directed graph of peers, where each node corresponds to a peer and each edge to the physical connection among two peers. The graph of peers is time varying, since nodes and edges are added or invalidated as time passes. Apart from the possibility of communication, which dictates the structure of the graph, peers are further organized in *communities*, based on their semantic similarity, or *classes*, based on the interface of Web services they support. All our deliberations are based on the *principle of local scope*, that dictates that no peer has a global knowledge of the entire graph, and therefore, all its decisions must be made depending solely on the knowledge that this node has at a given time