Tree Graph Views for a Distributed Pervasive Environment

Tuyêt Trâm Dang Ngoc[1] and Nicolas Travers[2]

[1] ETIS Laboratory - University of Cergy-Pontoise, France
Tuyet-Tram.Dang-Ngoc@u-cergy.fr
[2] PRiSM Laboratory - University of Versailles, France
Nicolas.Travers@prism.uvsq.fr

Abstract. The pervasive Internet and the massive deployment of sensor devices have lead to a huge heterogeneous distributed system connecting millions of data sources and customers together [Fra01]. On the one hand, mediation systems [BGL+99, DNJT05] using XML as an exchange language have been proposed to federate data accross distributed heterogeneous data sources. On the other hand, work [MSFC02, AML05, BGS01, NDK+03] have been done to integrate data from sensors. The challenge is now to integrate data coming from both "classical" data (DBMS, Web sites, XML files) and "dynamic" data (sensors) in the context of an ad-hoc network, and finally, to adapt queries and result to match the client profile.

We propose to use the TGV model [TDNL06, TDNL07a] as a mobile agent to query sources across devices (sources and terminal) in the context of a rescue coordination system. This work is integrated in the PADAWAN project.

Keywords: XQuery evaluation, Tree Graph View (TGV), Pervasive environment, Rescue Coordination.

1 Introduction

The pervasive Internet and the massive deployment of sensor devices have lead to a huge heterogeneous distributed system connecting millions of data sources and customers together.

On one side, data sources are heterogeneous as they can be of different types (relational, text, XML, streaming value, etc.) and can have different update frequencies (from "never" for some text document to "always" for sensors value) and their autonomy (from non-manageable obfuscated black box that just provide values to a full access management on a DBMS). On the other side, there can be different profiles of clients: access permission, terminal capabilities, user preferences, etc.

To deal with distributed heterogeneous and autonomous data sources, mediation systems have been widely studied [Wie92, MFK01, NGT98, BGL+99, DNJT05]. Such mediation systems [Wie92] provide a uniform interface to a multitude data sources using mediators and wrappers to handle respectively distributivity and heterogeneity.

T. Enokido, L. Barolli, and M. Takizawa (Eds.): NBiS 2007, LNCS 4658, pp. 406–415, 2007.

XML [BPSM98], has become the preferred format to represent semi-structured data [Abi97] and an effective way to define any type of data that can be represented as a tree.

Moreover, XQuery [W3C05] has proved to be an expressive and powerful query language to query XML data both on structure and content, and to make transformation on the data. In addition, its query functionalities come from both the database community (filtering, join, selection, aggregation), and the text community (supporting and defining function as text search).

TGV (Tree Graph View) [Tra06, TDNL06, TDNL07a] is a Tree Pattern-based model (such as TPQ [CJLP03] and GTP [AYCLS01]) to model XQuery queries. This model is suitable to our needs since :

- it supports the complexity of the full untyped-XQuery specification: relational and set operator, aggregation, ordering, nested reconstruction, conditional predicate, etc.
- it is designed for a mediation context accessing to distributed autonomous and heterogeneous data sources: its structure identifies data collections and dependencies between them. An annotation model in layers allows to annotate any piece of information (location of the source(s), cost model, etc) that can be useful a evaluation time. Finaly, transformation rules have been defined to optimize and evaluate the TGV to produce the result.

The rest of this paper is organized as follows. To start off with, we motivate the need for a mobile semi-structured model in a pervasive environment in Section 2 and we express issues and related works in Section 3. Further, we recall the TGV model and functionalities in Section 4 and show how it is suitable to our context. We then present some extensions to the TGV model that would make it more suitable to a pervasive environment (Section 5). In the end, we conclude in Section 6 and present future directions of our work.

2 Context

In the global context of our work, different types of data sources and terminal client are dissiminated all over a network consisted of traditional IP routing and addressing (e.g. the Internet), and ad-hoc routing scheme.

2.1 Motivational Scenario

The application scenario is the deployment of a Rescue Coordination Center after or during a disaster (fire, earthquake, flood, etc.) A truck (Figure 1) carries the PADAWAN proxy, some access points (wire, wireless WI-FI, SINK, etc.) and an Internet Access.

Using these gateway access points, the PADAWAN proxy can reach three types of networks:

- *Wireless ad-hoc Sensor Network [Toh01] (ad-hoc WSN)*: Sensors are deployed (eg. from helicopter or embedded in the rescue team equipment) in