

Chapter 31

X-CM: Extending Entity Relationship Model for Conceptual Modeling in XML Databases

Swee-Mei Chin

Multimedia University, Malaysia

Su-Cheng Haw

Multimedia University, Malaysia

Fang-Fang Chua

Multimedia University, Malaysia

ABSTRACT

The emergence of XML as the de facto for data exchange in the World Wide Web and the increase popularity of XML in the business application have urge momentum research on way to generate a well-formed XML document to store and maintain it in the databases. Thus, a good schema such as XML schema is undeniable needed in order to define the syntax and structure of the XML instance to ensure data integrity. Nevertheless, schemas serve as logical model rather than conceptual model where semantics of the underlying document are hardly expressed. As such, in this paper, the authors proposed X-CM, a new conceptual modeling for XML as the mechanism to model components of XML conceptually and to express the underlying semantic explicitly. First, the authors reviewed the semantics and structure of conceptual modeling of existing approaches. Then, the authors proposed their X-CM modeling construct and implemented X-CM in a university based scenario. Lastly, the authors summarized the evaluation result and comments provided by the XML database experts and evaluators based on the authors proposed model.

INTRODUCTION

With the increasing usage of XML in web applications, the urgency to ensure that the XML data must be properly integrated is essential. As

the result, a conceptual model to describe the structure, syntax and the constraints of the XML instance is very much desirable.

As to-date, Document Type Definition (DTD) and XML Schema Definition (XSD) are the most

DOI: 10.4018/978-1-4666-9562-7.ch031

widely used schemas. However, their textual structures are not easy to be understood especially when the collection of document is large and complex. As the consequence, they are not suitable to be used as the source for clarification on the semantic of XML data. Besides, the schemas represent the logical model rather than the conceptual model and thus the semantics underlying these documents are difficult to be expressed.

It is well known that graphical modeling is the most efficient way for information explore and learning. This clearly justify why conceptual model has been widely utilized in system design and modeling. Conceptual model helps define the real world objects and relationships, by capturing the concepts and semantics by putting all the objects together for better representation. For example, Entity-Relationship (ER) modeling and Unified Modeling Language (UML) is successfully used conceptual model in traditional databases such as relational databases and object-oriented databases respectively. In contrast, although XML is rapidly becoming the de facto standard in data exchange, there is no suitable mechanism to intuitively model the components of XML conceptually.

Although some conceptual modeling for XML databases have been proposed such as extending ER or UML, unfortunately, these models are not suitable as conceptual model for XML databases since they lack of the ability to express the semantics and special characteristics of XML such as hierarchy, ordering and schema-less content. These have encouraged us to propose a new conceptual model to support the special features of XML.

The main contributions of this paper are:

1. Identify and compare the advantages and limitations of each modeling technique for XML databases;
2. Propose a new conceptual model named X-CM which has the ability to capture XML semantics precisely and accurately;
3. Evaluate the usability of the proposed model.

Our paper is structured as follows. Section 2 reviews on the related works. Section 3 is the core of our paper which lists the set of requirements for XML conceptual modeling and hence, describes our proposed conceptual model. Section 4 illustrates how our modeling technique could be applied in a University-based scenario. In addition, we show how other conceptual models which lacks of some semantic expressions are constructed using the same scenario. Section 5 summarizes the comments from the XML experts. Section 6 describes and discusses on the experimental for evaluation on our proposed model. Section 7 analyzes on experimental results and finally, Section 8 concludes the paper.

RELATED WORKS

The review on the related works could be broadly categorized into four main groups, i.e., semantic network-based, extending ER, extending UML and hierarchical-based.

Feng et al. (Feng et al., 2002) proposed the semantic network-based conceptual model which comprises of two design levels, i.e., the semantic level and the schema level. The semantic level is constructed based on semantic network and contains four major components, i.e., the set of atomic and complex node, directed edge, label and constraint. The schema level is the generated transformation of the semantic network into XSD. Different types of semantic relationships such as generalization, aggregation, association and of-property relationship are introduced. Apart from this, their model also supports cardinality, strong/weak adhesion and ordering.

ER modeling has been very successful used in the relational database domain, with a set of simple graphical modeling constructs like rectangles and diamonds to describe the data objects and the relationships between them. Nevertheless, it is not able to model semi-structured data effectively. As

24 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/x-cm/142641

Related Content

Infrastructures of Knowledge Sharing Countrywide

Leila Nemati-Anaraki (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 375-394).

www.irma-international.org/chapter/infrastructures-of-knowledge-sharing-countrywide/142629

Innovative Approaches for Efficiently Warehousing Complex Data from the Web

Fadila Bentayeb, Nora Maïz, Hadj Mahboubi, Cécile Favre, Sabine Loudcher, Nouria Harbi, Omar Boussaïd and Jérôme Darmont (2012). *Business Intelligence Applications and the Web: Models, Systems and Technologies* (pp. 26-52).

www.irma-international.org/chapter/innovative-approaches-efficiently-warehousing-complex/58410

An Analysis of the Use of Predictive Modeling with Business Intelligence Systems for Exploration of Precious Metals Using Biogeochemical Data

Thomas A. Woolman and John C. Yi (2013). *International Journal of Business Intelligence Research* (pp. 39-53).

www.irma-international.org/article/analysis-use-predictive-modeling-business/78275

A Conceptual Model of Metadata's Role in BI Success

Neil Foshay, Andrew Taylor and Avinandan Mukherjee (2014). *Information Quality and Governance for Business Intelligence* (pp. 1-19).

www.irma-international.org/chapter/a-conceptual-model-of-metadatas-role-in-bi-success/96142

Enterprise Reconfiguration Dynamics and Business Alignment

G. D. Putnik, M. Manuela Cunham, B. Conceição Cortes and P. Silva Ávila (2007). *Adaptive Technologies and Business Integration: Social, Managerial and Organizational Dimensions* (pp. 1-32).

www.irma-international.org/chapter/enterprise-reconfiguration-dynamics-business-alignment/4227