

将关系数据库转换为XML文档及反之的系统与方法



信息和通信

计算机/人工智能/数据处理和信息技术

机会

互联网应用和电子商务的激增，加剧了对跨不同系统无缝数据交换和集成的需求。一个重大挑战源于两种主流数据格式的共存：关系数据库（针对传统数据处理和内部操作进行了优化）和XML（可扩展标记语言，已成为网络上数据交换的标准）。这两种格式之间的传统转换方法通常存在局限。许多现有的支持XML的数据库系统只能将少数关系表转换为XML文档，无法转换整个关系数据库或同步复制的XML数据库。此外，这些方法在转换过程中经常无法保留底层的数据语义和约束（如函数依赖、基数、参与度）。这导致了繁琐、易出错的过程，阻碍了真正集成信息高速公路的实现。因此，迫切需要一种强大、自动化的系统，能够实现高效的双向转换，同时保持数据完整性、提高互联网计算性能，并促进如B2B数据传输和遗留系统互操作性等应用。

技术

本专利提出了一种用于关系数据库和XML文档之间双向转换的全面系统和方法。其核心创新在于一个能保留数据语义的模式转换过程。该方法涉及将关系模式反向工程为扩展实体关系（EER）模型，以恢复数据依赖和约束。然后，该EER模型被映射到XML模式，特别是文档类型定义（DTD）。关键的是，本发明引入了XML树模型作为一个概念模式，它以图表形式表示DTD，并将原始关系约束（例如，弱实体、参与度、基数、聚合、泛化）保留在XML文档的层次结构中。对于数据转换，关系元组根据这些映射的约束被加载到XML元素实例中。该过程还支持创建复制的、信息容量等效的数据库。它包括通过将SQL更新事务（插入、更新、删除）转换为文档对象模型（DOM）操作，实现关系数据库和XML数据库之间同步或异步更新的机制，确保两种格式间的数据一致性。

优势

- 在转换后的XML文档中保留了原始关系数据库的数据语义和约束（例如，函数依赖、基数）。
- 能够将整个关系数据库完整地转换为XML文档，反之亦然，不同于功能有限的传统方法。
- 通过允许内部关系数据并行处理和外部XML数据传输，提高了互联网计算性能。
- 在系统故障时，便于从其对应的关系数据库自动恢复XML数据库。
- 允许用户在使用底层系统时使用熟悉的查询语言（如SQL）。
- 支持复制的关系数据库和XML数据库之间的同步或异步更新，保持数据一致性。
- 提供了一个用户友好的概念模型（XML树模型），用于理解XML数据库结构。

备注

IDF: 94

IP状态

专利已授权



技术成熟度等级 (TRL) ?

4

发明人

方士骜教授

查询: kto@cityu.edu.hk

Follow-on

Develop
ConceptProof
Concept

Build Value

应用

- 企业数据集成：共享业务数据，并实现使用不同数据格式的不兼容系统之间的互操作性。
- 电子商务和B2B应用：通过互联网以标准化XML文档形式传输关系数据，用于交易和数据交换。
- 遗留系统现代化：将遗留的关系数据暴露给使用XML的现代应用程序和网络服务。
- 内容管理与聚合：将关系数据发布为XML文档，用于内容分发和聚合。
- 数据库迁移与复制：为不同的操作需求，创建和维护同步的、等效的关系和XML格式数据库。
- 网络服务与数据交换中间件：作为面向服务体系结构中转换数据格式的中间件的核心组件。

