

1998年 XML 1.0<sup>[1]</sup> 正式成为 W3C 的推荐标准，标志着 XML 的正式面世。从此，XML 获得了迅猛地发展。XML 正逐渐成为互联网数据呈现和交换的新标准，并广泛应用于电子商务、电子数据交换、电子图书馆等领域。如何对 XML 数据进行有效管理（如存储和查询）已成为一个急需解决的问题。大体上，有三个解决途径<sup>[2]</sup>：一是建立专门的数据库管理系统（Native XML DBMS）；二是使用面向对象的数据库管理系统（Object-Oriented DBMS）；三是使用关系数据库管理系统（Relational DBMS）。Xparent<sup>[3]</sup>认为，当要把 XML 数据存储在关系数据库管理系统时，为存储 XML 数据而进行存储模型设计的问题就变成了数据库模式的设计问题。XRei<sup>[4]</sup>把数据库模式设计成两类：结构映射策略（structure-mapping strategy）和模型映射策略（model-mapping strategy）。前者是在诸如 DTD、XML Schema 等模式文件的帮助下，不预先定义关系模式，而是实行为 XML 到关系的模式转换。后者的思路是，先建立一个数据模型，在没有模式文件的帮助下，根据模型的特点和查询的需求，预先定义一个关系数据库模式。

本文提出了一种新的关系模式下的存储策略：XEDGE。它以归在模型映射策略的门下。与其他模型映射策略类的策略相比，有如下优势：

- (1) 它建立在 XML 数据通用参考模型（XDGRM）的基础上。
- (2) 扩充了节点类型和边类型，所有的混合元素都有文本节点和文本边。从而使得内部节点都统一起来，更简洁，利于表达化。
- (3) 有路径信息冗余，显示化路径结构信息，从而有了

ML 数据的全局信息。

- (4) 它是一个五表关系模式：元素边表、属性边表、值边表、外部路径表、内部路径表。
- (5) 试验结果表明，它的查询较优。

## XDRGM

XEDGE 策略需要一个 XML 数据模型的支持。为此，笔者设计了一个 XML 数据模型，即 XDGRM，全称是 XML 数据通用参考模型（XML Data General Reference Model）。它的主要特征如下：

- (2) 与 OEM、XPath 等模型相比，具有更好的通用性。
- (3) 直接支持 XML 数据，支持 XML 数据的有序特性，持有混合内容的元素、支持文档间的链接等。
- (4) 是一个以边为中心的 XML 数据图模型，XML 信息集中在边上。
- (5) 显示化节点之间的关联信息，如元素与元素之间、元素与属性之间的关联、元素与文本之间的关联、文本元素与另一个文档中的元素的关联等。
- (6) 显示化路径信息。
- (7) 面向关系模式。

为了更好地说明 XDGRM，笔者使用了一个 XML 实例，见图 1 所示。

```
<db>
  <article key="conf/WLDDB/2002XMLP082801">
    <author>LiuJianhua</author>
    <cite href="books/acm/Wilson95">
      reference book from Jennifer
    </cite>
  </article>
  <book key="books/acm/Wilson95">
    <author>Jennifer Wilson</author>
  </book>
</db>
```

图 1 一个 XML 实例

图 2 是图 1 中 XML 实例的 XDGRM 图表示。



图 3 不同策略下的查询结果

最后,做了一个可伸缩性试验,分别在 CL2DM 1、CL2CL3DM 3 进行查询,查询结果如图 4。结果表明,XEDG 策略具有良好的可伸缩性。



其覆盖区域中都将产生弱信号区和盲区。直放站具有性价比高，施工简单等特点，在我国移动网络上有大量的应用。对一些远地区和用户数不多的盲区，由于架设模拟或数字基站成本高，基础设施也较复杂，所以一般采用直放站。直放站的使用大提高了运营商网络的处理能力和覆盖范围，同时也不可避免地带来一些问题，具体如下：

由于直放站要填补基站覆盖不到的区域，因此直放站设备布区域较广，位置大多比较偏远。直放站在整个网络中是孤的，一旦出现故障，无法及时获知设备的告警信息，也无法及时对故障进行处理，因此给少数用户地区和信号盲区的移动通信设备维护和管理工作带来很大的困难，对整个网络的质量也产生一定程度的影响。

为了解决上述问题，本文设计出一种新的直放站方案。该方案使直放站能与中心服务站直接无线通讯，并且具有自动检测连接状态和自动拨号功能，防止不确定因素干扰而使网络连断。本设计方案采用基于 S3C2410 的 Windows CE.NET 的嵌入式系统来实现。

关系模式、元素边关系模式、属性边关系模式和值边关系模式。它们是建立在核心边类型的基础上的。通过试验对比，说明 EDGE 策略是有效的，且远优于边表策略。

但是，该策略仍然存在诸多问题：

(1) 如何支持 XML 数据的更新，如何重构从关系数据库返回的结果，这方面没有做深入的研究和相应的试验对比。

(2) 在查询方面，XML 查询到 SQL 查询的转换还处于手工阶段。

上述几个问题都具有很大的挑战性，今后应重点展开研究，使得本策略更为实用和高效。

参考文献：

[1] World Wide Web Consortium. Extensible Markup Language (XML)[EB/OL]. <http://www.w3.org/XML/>,2006.2.25.

[2] A Deutsch,M F Fernandez,D Suciu. Storing Semistructured Data with STORED [C] In Proc of SIGMOD Conference, 1999:431~442

[3] H Jiang,H Lu,W Wang et al Path materialization revisited;



图 1 基于嵌入式技术的无线上网直放站组成

基于 S3C2410 的 Windows CE.NET 的嵌入式系统的上网直放站主要由以下几部分组成：

(1) 直放站。用于延伸移动通信网络。通过串口 1 与嵌入式系统通信。

(2) USB 接口。利用 USB 的热插拔性能，通过 U 盘可本地实现软件的升级和数据更替。

(3) CDMA-Modem 模块。通过串口 2 与嵌入式系统连接，当自动拨号程序启动时，对其拨号，以登录网络。

(4) 基于 S3C2410 的 Windows CE.NET 的嵌入式系统。要实现人机交互、数据通信、自动拨号的功能。

An efficient storage model for XML data[C] In Proc of ACM SIGMOD,2002.5(1):85~94

[4] M Yoshikawa,T Amagasa,T Shimura et al XRM: Path-Based Approach to Storage and Retrieval of Documents using Relational Databases [J] ACM Transaction on Internet Technology,2001.1(1):110~141

[5] Yannis Papakonstantinou,Hector Garcia-Molina, and Jen Widom. Object exchange across heterogeneous information sources[C] In Proc of Int Conf on Data Engineering (ICDE), Taipei,Taiwan,1995:251~260

[6] World Wide Web Consortium. XML Path Language (XPL)[EB/OL] <http://www.w3.org/TR/xpath20/>,2005.10.2

[7] Schmidt AR,Waas F,Kersten ML,et al The XML benchmark project[R] Amsterdam: CWI,2001.

[8] 曹亮,王茜,卢菁.XML 数据在关系数据库中存储和检索的研究.东南大学学报(自然科学版),2002.32(2):124~128

[9] 王晓玲,宋金锋,董逸生.基于演化计算的 XML 数据的关系存算机研究与发展,2003.40(7):1110~1116

