

一种关系数据库的网上发布和集成方法*

奚建清 夏志忠 刘发贵

(华南理工大学计算机工程与科学系 广州 510641)

E-mail: csjqxi@scut.edu.cn

摘要 该文针对现有数据库信息网上发布和集成方法的缺点,提出了一种新的发布和集成方法,称为 MOAT (movable operations as texts,可移动文本化操作)模型.该方法不是把数据库的操作以一段程序的方式表示并在网上发布,而是表示为一段文本文字来发布.其特点是表示方式简洁直观并具有可修改性,允许用户通过简单的工具,直观地修改或复合已有的数据库访问操作,生成新的数据库访问操作,并且可以根据需要将新的操作在网上再发布.文章还介绍了 MOAT 模型的一种实现框架,最后提出了展望.

关键词 发布,共享,数据库,信息访问,互连网.

中图法分类号 TP311

随着网络技术的发展,各种信息通过网络,特别是 Internet 进行交流、共享. Internet 带来了 WWW 服务器和客户浏览器的概念,共享信息常常以文本方式(主页)从 WWW 服务器传输到客户端的浏览器进行显示.由于主页使用的是文本方式,因此,客户端用户可以进一步地编辑和修改,并能够把新的主页发布到网上.本文的研究目的在于,探讨利用类似的浏览器界面,用文本的方式表达数据库信息的访问操作并发布到网上,使用户能从客户端下载这种文本访问命令,进行数据库访问操作的编辑、修改和再发布.

1 已有数据库访问方法的分析

数据库访问是一个历史较长并且成熟的课题.自关系数据库发展和商品化以来,用户访问数据库信息主要通过以下两种方式^[1,2]来实现.

(1) 通过 SQL 命令界面

这种方式可以让用户通过网络远程或在终端上直接访问数据库,其缺陷是需要用户懂得 SQL 命令的使用以及数据库的联接等操作;如果数据库分布在不同的结点上,则需要更多地了解多数据库集成等方面的知识;用户的操作界面和结果都是字符型的,不能适应现代用户的需要.

(2) 通过应用程序界面

在应用程序中,插入访问数据库的命令或函数,用户通过应用程序访问数据库.这种方法能够提供图形界面,用户操作方便,无需了解数据库的专业知识.这是目前最流行的方法.随着 Java 技术和 Active-X 技术的兴起,应用程序通过 HTML 主页能够从服务器端下载执行,使得访问数据库的用户界面可移动,从而进一步提高了数据库信息的共享性.因此,把数据库访问操作集成在 Java applet、cgi 程序或 VB 程序中,已成为在 WWW 网上发布数据库信息的主要手段之一.

这种方法的一个特点是数据库访问操作需由专业人员通过程序手段来安排,普通用户一般不能改动已经设置好的访问命令,仅仅可以输入一些参数值等,缺乏在已有的数据库访问操作的基础上创建并发布新的数据库

* 本文研究得到广东省自然科学基金(No. B6-109-275)资助.作者奚建清,1962年生,博士,副教授,主要研究领域为数据库,分布信息系统.夏志忠,1972年生,助教,主要研究领域为模式识别,数据库.刘发贵,女,1953年生,讲师,主要研究领域为人工智能,对象技术.

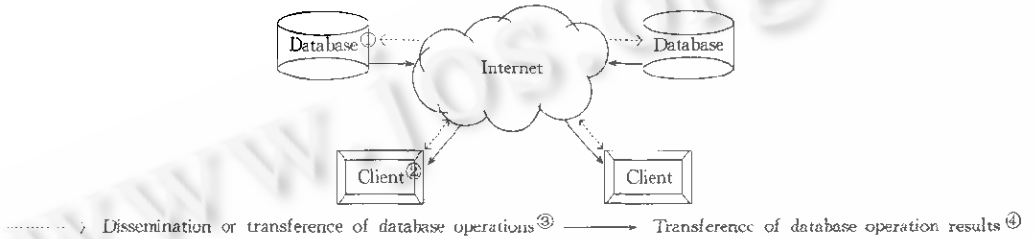
本文通讯联系人:奚建清,广州 510641,华南理工大学计算机工程与科学系

本文 1998-09-21 收到原稿,1999-05-30 收到修改稿

访问的手段。因此，这种类型的数据库信息是通过中央化控制来发布的，这与在信息社会中，信息的多样化、多元化是相悖的。

2 MOAT 模型

本节介绍一种新的数据库网上发布和集成模型 MOAT (movable operations as texts)，其主要思想是不把数据库的操作以一段程序的方式表示在网络上进行传输，而表示为一段文本文字，也就是把数据库的访问命令（主要是 select 命令）作为文字串直接由某一个信息定义者发布到网上，网络用户通过一定的客户端和服务端软件接收到该命令串，如果用户要访问该命令所对应的数据库数据，则由客户端软件执行该命令，并把结果返回。由于发布的数据库访问操作均是文字性的（而不是一个程序），因此具有可操作性。通过客户端软件，用户可以直观地修改或复合已有的数据库访问操作，生成新的数据库访问操作，并且可根据需要把新的操作命令在网上发布。图 1 是 MOAT 模型的简单图示。



①客户端，②数据库，③数据库操作的发布或传送，④数据库操作结果的传送

Fig.1 The transference of operations and results in MOAT model

图1 MOAT模型中操作及其结果的传送

(1) 基本概念

为了清楚地说明 MOAT 的思想，需要定义一些基本概念。

定义 1. 一个关系数据库操作 O 是指对某些数据库表的一个查询、插入、删除、修改操作或一个存储于过程的调用，表示为 $\langle [db\text{-names}], operation \rangle$ ，其中 $db\text{-names}$ 为一组数据库的名字， $operation$ 是上述操作之一。

例如， $\langle [db\text{-server} // HumanResource], "select name, age from person;" \rangle$ 是一个操作，其意义是对数据库服务器 $db\text{-server}$ 上的数据库 $HumanResource$ 进行查询操作。

定义 2. 一个可移动的关系数据库操作 MO 表示为 $\langle drivers, O \rangle$ ，其中 $drivers$ 表示操作中涉及的数据库的 ODBC 驱动程序信息（包括驱动程序名、用户名和口令信息等）。

上述 O 和 MO 均是用文本方式来表示的。以下是 MO 的一个例子：

$\{ [select name, age from person where age > \# \$X][\# \$\text{年龄}]; \sim HumanResource \sim ODBC = db2; db\text{-server}; LI4; LI4123; 20 \}$ 。

这个 MO 的含义是通过 ODBC 对服务器 $db\text{-server}$ 上的 $db2$ 数据库 $HumanResource$ 中的表 $person$ 进行查询操作，以 $LI4$ 的身份和 $LI4123$ 的口令进行查询。

(2) 数据库操作的发布

所谓发布是指某种信息能够在 Internet 网上被用户以简单的方式访问到。目前，访问的方式可以有通过 WWW 浏览器或其他应用程序。这里假定程序 P 表示这样的程序，它能执行任意 MO 并把结果显示给用户。

定义 3. 一个 MO 的发布是指 P 能够得到 MO 的表示，并执行 MO 所定义的操作并把结果显示给用户。

显然，这种定义意味着 P 必须具有数据库连接和执行 MO 中所定义的数据库操作的能力。目前，商用的浏览器 Netscape 和 Internet Explorer 都没有这种功能，因此，它们不适合直接完成 MO 的发布。第 3 节介绍的 Pcwergates 系统建立在 IE4.0 上。它不仅能够直接发布和浏览 HTML 主页，而且能够发布、执行 MO 和显示结果。

(3) MO 的复合

由于 MO 的定义是文本信息,并且能够在客户端得到,因此,可以提供特定的工具,使用户修改、合成各种 MO 的功能,我们把这种能力称为 MO 的复合。

这里仅以查询型的 MO 作为例子来加以说明。所谓查询型的 MO 是指所含的数据库操作是一个查询操作。对这种操作的复合分成以下两类。

(a) 多个 MO 合成一个 MO

由于每个 MO 在逻辑上可以看成是一个关系表,因此,多个表可以进行联结(join)运算,形成新的关系表,这个新表的文本定义成为一个新的 MO。

例如,对于已有的两个查询语句:

```
{[select name, age from person where age > # $X][# $ 年龄]; ~HumanResource ~ODBC=db2; db-server; LI4; LI4123; 20}
```

```
{[select name, salary from income where city = # $X][# $ 城市]; ~COMPANY ~ODBC=ORACLE7.0; SERVER2; CHANG3; GUANGZHOU; 20};
```

它们的连接可以生成新的查询操作:

```
select name, salary, age from income, person where city = # $X and age > # $Y.
```

这个操作的描述加上合适的 ODBC 等信息,就可以组成一个新的可发布的 MO。

当然,除了联结,还可以定义多个 MO 之间的联合、差等集合运算,只要它们有一致的逻辑定义。

(b) 修改已有的 MO

通过对已有的一个 MO 的选择条件、投影字段的选择进行修改,可以形成新的 MO 的文本定义。例如,对于

```
{[select name, salary from income where city = # $X][# $ 城市]; ~COMPANY ~ODBC=ORACLE7.0; SERVER2; CHANG3; GUANGZHOU; 20};
```

可以修改成如下的 MO:

```
{[select name from income where city = # $X][# $ 城市]; ~COMPANY ~ODBC=ORACLE7.0; SERVER2; CHANG3; GUANGZHOU; 20}.
```

MO 的复合可以在客户端进行,形成新的 MO 并对此再发布。因此,MOAT 模型为在网上共享、发布和产生新的数据库操作提供了必要的基础。

3 访问控制

由于数据库信息的保密性,各个 MO 的发布执行必须通过用户权限检查。传统的数据库应用系统中有两种检查方式,一种是检查应用程序执行者的数据库访问权限,这种方式适用于动态 SQL 语句执行方式;另一种是检查应用程序编写者的权限,这种方式适用于静态 SQL 的访问控制。

在三层信息访问系统的体系结构中,数据库常常隐藏在中间层软件,如 WWW 服务器的后面,客户端(浏览器)通过 WWW 服务器访问到数据库的信息。在这种模型中,WWW 服务器也可以设置成为信息访问权限的控制者,这种方式的缺陷是,WWW 服务器可能需要为不同的数据库设置不同的访问方式和控制方式,每次用户访问均需通过 WWW 服务器控制,因此容易成为网络瓶颈。

在 MOAT 模型中,客户端程序 P 负责把数据库服务器上的数据库操作发布到网上,同时 P 还可以负责接受其他用户的数据库访问请求,执行请求中包含的数据库操作,然后把访问结构返回给其他用户, P 同时可以控制用户的访问。由于 P 往往和数据库服务器在一个工作组内或同一台机器上,因此,可以制定关于特定数据库的访问策略,且不会成为控制瓶颈,如图 2 所示。因此,MOAT 模型实际上也是一种三层体系结构,所不同的是中间层就是客户端,对数据库的访问不存在类似于 WWW 服务器的访问和控制瓶颈问题。

4 MOAT 的实现——Powergates 系统

Powergates 是在 Window95 平台上以 HTML 界面方式,支持多种信息访问、浏览和共享的分布信息访问和发布系统。该系统已经被作者在华南理工大学完成实现。Powergates 的主要功能之一是实现了 MOAT 模型,支

撑环境为 MS IE4.0 和 ODBC.

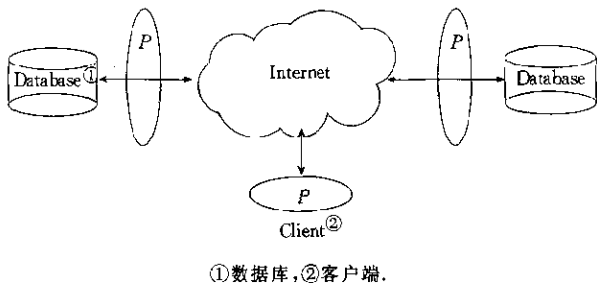


Fig. 2 The 3-tier architecture of MOAT model
图2 MOAT模型的3层结构

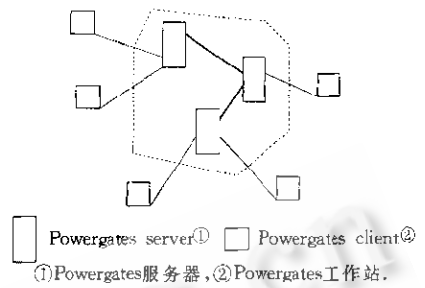


Fig. 3 Architecture of Powergates
图3 Powergates系统结构

4.1 Powergates 的体系结构

Powergates 的主要设计思想是,在信息网络中,任何一个用户既是信息的消费者,又是信息的产生者;信息种类可以是各种文档,也可以是 MO(数据库访问操作),或一个程序的执行结果.对用户而言,上述信息种类的内在区别和详细访问实现并不重要,重要的是有一个统一的信息访问、发布和产生的界面. Powergates 的系统结构如图 3 所示.

在 Powergates 中,各个客户端用户产生或定义各种共享信息,Powergates 服务器收集这些信息的地址(URL)或定义描述,组成信息目录,并向其他用户发布这些共享信息目录.服务器之间同步从不同客户端得到的共享信息,并负责维护这些共享信息目录.在服务器信息目录中,登记各种文档信息的 URL 或 MO 的文本定义.

4.2 MO 的执行

当一个用户通过定义、继承或从服务器中得到一个 MO 后,该 MO 就存储在该用户的本地信息目录内. MO 的执行分为 3 步,第 1 步是提示用户输入必需的信息,例如用户名、口令或一些执行参数;第 2 步是分析 SQL 命令,把分布查询 SQL 命令按结点分解成子 SQL 命令,并把这些命令发送到相应的发布端;第 3 步是在各个相关客户端执行完 SQL 命令后,把结果发回到原查询端,在查询端综合计算后,得到结果并在客户端显示结果.

4.3 MO 的复合

Powergates 提供了方便的操作界面,使得用户可以方便地把任何关系表用查询语句的形式发布到网上.同时,已有的 MO 可以复合成新的 MO,即便其中涉及的关系表可能是在不同的地点和不同的数据库中.

MO 的复合是纯粹的字符串操作,把多个 select 命令串和相应的 ODBC 信息、显示字段名称、用户信息等参数有选择地复合在一起,就可以形成新的查询操作(MO).

4.4 用户信息的发布

对于客户端的任何信息项(包括各种 MO),一旦它被用户标记成共享,Powergates 客户端程序就会把其 URL 或定义信息登记到服务器端.当其他用户需要访问最新信息时,会与服务器相连,新的共享信息被发送到用户端,用户可以访问信息内容. Powergates 客户端的结构如图 4 所示.

5 前景和结论

MOAT 模型的基本思想是能够在网上方便、高效地实现数据库信息的发布和异种数据库信息之间的集成,能够让用户方便地裁剪已有的信息,定制适合不同需要的信息内容,并能方便地发布出去.所基于的基本方法是把信息的定义(描述)文本化,特别是变程序方式的数据库发布为文本化的发布.

MOAT 模型提供的是一种基本方法,其实现可以通过任何一种语言.其中一种方法是利用 XML 语言^[3],把数据库访问命令定义为一种信息资源,扩充客户端浏览器功能或利用 Java 编写访问数据库的应用程序,但这种

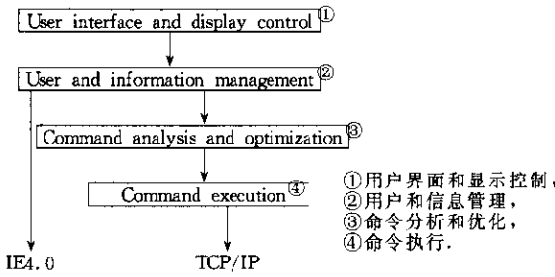


Fig. 4 The structure of Powergates client software
图4 Powergates客户端软件结构

- ① 用户界面和显示控制,
- ② 用户和信息管理,
- ③ 命令分析和优化,
- ④ 命令执行.

方法仍有可能造成 WWW 服务器成为瓶颈。

MOAT 模型使得用户之间的关系更为复杂,因为各个用户既可以是信息的创造者,又可以是信息的消费者,各种共享关系纵横交错,还涉及到数据库的访问权限.因此,必须研究新的合适的用户信息交流模型来规范信息的交流和共享.

Powergates 系统是实现 MOAT 的初步尝试,它证明了 MOAT 模型的可行性.目前,Powergates 已经用于小规模的应用系统中,作为通用的数据库和各种文档信息的共享和访问界面.下一步的工作是把该系统升级

到企业级的网络计算环境中,并要研究新的用户信息交流和共享模型.同时,加强系统的安全性,引入有效的安全通信算法和加密算法.

致谢 本文所讨论的 Powergates 系统已经在华南理工大学实现.温道明、胡海滔、金海华、陈颖棠和英杰等同学为系统的实现付出了很大的努力,完成了关键性的编码工作,在此表示感谢.

参考文献

- 1 Wang Neng-bin. Database System. Beijing: Electronic Industry Publisher, 1995
(王能斌.数据库系统.北京:电子工业出版社,1995)
- 2 Chen Xi-ming, Yang Guo-wei. Java Programming and Application. Chengdu: Sichuan University Press, 1996
(陈锡明,杨国纬. Java 程序设计及应用.成都:四川大学出版社,1996)
- 3 W3C Recommendation 10-February-1998. Extensible markup language (XML) 1.0. <http://www.w3.org/TR/1998/REC-xml-19980210>

A Method for Disseminating and Integrating Relational Database across Networks

XI Jian-qing XIA Zhi-zhong LIU Fa-gui

(Department of Computer Engineering and Science South China University of Technology Guangzhou 510641)

Abstract Current methods for disseminating and accessing database data or information have a few drawbacks. This paper proposes a new way, which is called MOAT (movable operations as texts) model. The model differs from the traditional ones in the representation of the database accessing operation across the network. MOAT represents the operation as a kind of text description while the traditional methods use a segment of programs (such as Java applets) to represent data accessing operations. MOAT model makes the transference of operations more succinct, visualized and modifiable. Users can easily modify or integrate the existing operations in text formats to form new database operations, which can also be disseminated again across the network. This paper also introduces an implementation architecture of MOAT model. Some further work is given finally.

Key words Dissemination, share, database, information accessing, Internet.