

A Model for Query Optimization Based on Multi-mobile-agent

YANG Jin-cai¹, LIU Xiao-jiao², ZHAO Sen, HU Jin-zhu

Department of Computer Science, Central China Normal University, Wuhan, China. 430079

1. jcyang@mail.ccnu.edu.cn, 2. betty_3007@163.com

Abstract: Because of the mobile database system with mobility, frequent disconnection, asymmetry of network and limited resource, the traditional query optimization has no longer applicable. Based on analyses on the characteristics of mobile database and mobile agent, a multi-mobile-agent technology is introduced into the mobile database and a model of query optimization based on multi-mobile-agent is proposed. Cooperation of multi-mobile-agent, query decomposition and data broadcast can reduce the times of query and network flow, improve the efficiency of query.

Key words: mobile database; multi-mobile-agent; data broadcast; query optimization

一个基于多移动 Agent 的查询优化模型

杨进才¹, 刘小姣², 赵森, 胡金柱

华中师范大学计算机科学系, 武汉 中国, 430079

1. jcyang@mail.ccnu.edu.cn, 2. betty_3007@163.com

【摘要】: 由于移动数据库系统具有移动性、频繁断接性、网络通信非对称性以及资源有限性等特点, 传统的查询优化方法已不再适用。在分析了移动数据库和移动 Agent 各自特点的基础上, 将多移动 Agent 技术引入到移动数据库中, 提出了一个基于多移动 Agent 的查询优化模型。该模型中多移动 Agent 的协作、查询分解、数据缓存、数据广播技术的应用, 能减少查询次数和网络流量, 提高数据库的查询效率。

【关键词】: 移动数据库; 多移动 Agent; 数据广播; 查询优化

1 引言

查询一直都在数据库技术中占有重要地位, 且由于移动计算^[1]中的无线通信、移动性和移动设备有限的计算和存储能力等特点, 使得查询优化在移动数据库中显得尤为重要, 然而传统的查询优化技术不再适用。目前, 弱连接条件下的移动查询技术^[2]研究主要有两个方面, 一是基于复制、缓存技术的查询, 二是采用适应性中间件技术以及代理技术。

移动 Agent^[3]是具有移动特性的智能 Agent, 它可以自主地在网络上从一台主机移动到另一台主机, 并代表用户完成指定的任务, 如检索、过滤和收集信息, 甚至可以代表用户进行商业活动。MA (Mobile Agent) 技术是分布式技术和 Agent 技术想结合的产物, 在移动计算环境中, 它具有其明显的优势^[4]:

(1) MA 技术能较大地减轻网络上的数据流量。通过将服务请求 Agent 移动到目的主机, 使该 MA 直接访问该主机上的资源, 与源主机只有较少的交互, 从而避免了大量数据的网络传送, 降低了系统对网络带宽的依赖, 这也同时缩短了通信时延, 提高了服务响应速度。

(2) MA 能以异步的方式自主地运行。可以将要完成的任务嵌入到 MA 中, 并通过网络将其派出去, 然后就可以断开源主机与目标主机的连接。此后, MA 就独立于最初生成它的进程, 可以异步自主地运行。源主

机可以在随后适当的时候再与目标主机连接并接收运行的结果信息。

(3) MA 可以根据服务器和网络的负载动态决定移动目标, 有利于负载均衡。而且, MA 的智能路由减少了用户浏览或搜寻时的判断。

(4) 在进行任务处理时, 可通过动态创建多个 Agent 并行工作提高效率并降低对任务的相应时间。

(5) 能够克服网络隐患, 在不可靠的网络中也能提供稳定的服务。

在移动计算环境下, 由于移动客户机无法准确预计应用程序在断接时需要哪些数据, 而且移动客户机的存储能力很有限, 因此用传统技术实现数据密集型访问是很困难的。另外, 从服务器到移动客户机的下行通信带宽一般要远大于从移动客户机到服务器的上行通信带宽, 而且移动客户机从服务器接收数据的开销也小于发送数据的开销, 因此在大部分场合, 即使是处于断接状态的移动客户机也可以选择接收从服务器发送的下行广播信息。于是, 针对网络的这种非对称性, 广播技术把大量用户频繁访问的数据^[4]组织在共享无线信道上, 通过对空中数据及其索引的合理组织, 以周期性的广播方式提供给移动客户机访问, 使移动用户能够以最小的代价有选择地接收数据, 从而能够提高查询效率及移动数据库的可伸缩性。

文中提出的模型就是将多移动 Agent 技术应用到移动数据库的查询处理中, 移动 Agent 由源计算机产

基金项目: 湖北省科技攻关项目 (2007AA101C49)。

生后，它有一定的智能性，即使在移动用户断接的情况下，也能携带执行代码和数据自主的在异构网络中漫游，逐个访问相关服务器，直到完成用户提交的任务。同时对热点数据进行记录，存储，建立一个查询结果信息库，当新的查询请求到达时，可先到此信息库中查找是否有结果记录，如果有则直接使用此结果。另外在此系统中还设置有广播子系统，它以一定的周期，将出现频率较高的查询语句以及结果按照一定的广播算法对移动终端用户进行广播。

2 基于多移动 Agent 的查询优化模型

2.1 主要思想

(1) 查询分解：按照“分解后的子查询是简单查询且能被并行执行，对应的 Agent 只需到一个工作站点”的粒度来分解查询。(2) 多 Agent 协作^[5]：多个移动 Agent 之间能通过按照约定的协议进行通信、协商与协作来共同完成单个 Agent 所不能完成的任务。(3) 数据存储广播^[6]：将查询频率较高的结果储存起来，可以直接或者作为中间结果使用，同时以固定的广播周期，按照一定的算法将数据进行广播。

2.2 数据结构定义

(1) 查询语句表

查询语句表主要是用来记录查询语句信息以及统计查询语句出现的频率。逻辑结构如下：

msg_id	frequency	query_msg	new_time
msg_id: 查询语句的 id 号，由系统生成。			
frequency: 查询语句出现的次数。			
query_msg: 查询请求的内容、结果等信息。			
new_time: 查询语句最后的更新时间。			

(2) 任务 Agent 记录表

任务 Agent 记录表主要记录的是 Agent 管理器在分配任务时，不同的任务 Agent 对应的子查询。逻辑结构如下：

sub_query	mob_agent
-----------	-----------

sub_query: 由查询分解后，生成的子查询语句的 id 号，由系统生成。

mob_agent: 生成的每个任务 Agent 的 id 号，由系统生成。

(3) 任务 Agent 状态表

任务 Agent 状态表记录的是每个任务 Agent 当前的状态，忙或是空闲，它是通过一个布尔值来反映的。逻辑结构如下：

mob_agent	sta_bool
-----------	----------

sta_bool: 标识任务 Agent 的当前状态。当值为 0 表示空闲，当值为 1 则表示当前正在执行任务。

2.3 模型结构

2.3.1 模型定义

一个基于多移动 Agent 的查询优化模型的形式化定义如下：

定义：一个基于多移动 Agent 的查询优化模型= $\langle Q, A, T \rangle$ 其中： Q 表示用户请求的查询集合 $Q = \{Q_1, Q_2, Q_3, \dots, Q_m\}$ ，每个查询 Q_i 可以分解成若干个子查询 $Q_i = \{S_{i1}, S_{i2}, \dots, S_{in}\}$ ， $i \in [1, m]$ ， $n \geq 1$ ，每个 S_{ij} 都有唯一的标识符 sub_query，由系统生成的。 A 表示负责完成这些查询的 Agent 管理器集合 $A = \{A_1, A_2, A_3, \dots, A_m\}$ ，每个 Agent 管理器的管理内容包括：

(1) 任务 Agent 集合 $A_i = \{A_{i1}, A_{i2}, \dots, A_{in}\}$ ， $i \in [1, m]$ ， $n \geq 1$ ，每个 A_{ij} 都有唯一的标识符 mob_agent，由系统自动生成，每个 A_{ij} 可以执行 Q_i 中的一个或者多个子查询。

(2) 任务 Agent 记录表存放着 Agent 管理器在分配任务时，不同的任务 Agent 对应的子查询。

(3) 任务 Agent 状态表记录着每个任务 Agent 当前的状态，忙或是空闲。

(4) 计数器 counter (Q_i)，其初始值为 Q_i 中所有的任务 Agent 个数。当有任务 Agent 完成子查询返回结果时，counter (Q_i) 的值就减 1，直到 counter (Q_i) = 0，表示 Agent 管理器收到所有任务 Agent 返回的结果。

2.3.2 系统结构

基于多移动 Agent 的查询优化模型的系统结构图如图 1 所示：

2.3.3 功能描述

(1) 用户接口

用户接口是用户和移动终端交互的窗口，可以通过界面 Agent 来实现，主要是接收用户的查询请求，对用户身份的合法性进行验证。

(2) 查询分解器

查询分解器是按照一定的分解算法将查询语句分解为能够并行执行，对应的 Agent 只需到一个工作站点执行的简单查询。

(3) 查询匹配器

查询匹配器对分解后的每个子查询语句在查询语句表查找是否有结果记录，如果有，则直接用此结果，将结果送到广播子系统，不用进行查询操作。

(4) Agent 管理器

在收到经过查询匹配后有结果的子查询和没有结果记录的子查询，Agent 管理器会为没有查到结果的子查询分别生成一个 MA (Mobile Agent)，并将子查询任务分配给它们。计数器 counter 赋初值为生成 MA 总的个数，同时，在任务 Agent 记录表中录入 MA 与子查询的相关信息，

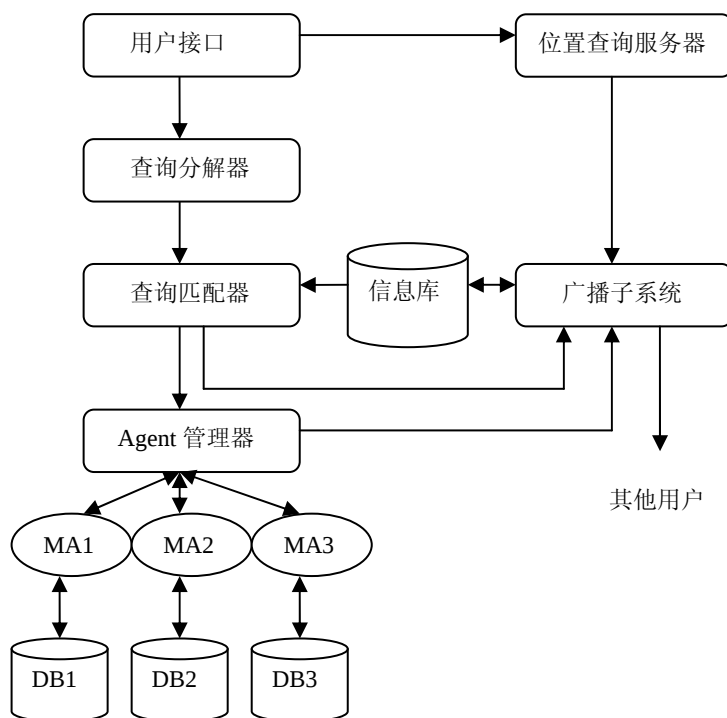


图1 基于多移动 Agent 的查询优化模型

如：哪个子查询由哪个 MA 去执行以及工作站点和查询任务的内容，参数等。还要将状态表中相应的 MA 的状态值改为 1（表的初始值为 0）以表示当前的 MA 在执行任务。

在任务执行过程中，MA 顺利完成任务，成功返回结果后，Agent 管理器接收结果，counter 的值减 1，同时将对应的状态值改为 0。如果 MA 遇到阻塞等问题不能完成任务时，则向 Agent 管理器发出 help 消息，Agent 管理器则在状态表中查询当前状态值为 0 且离阻塞 MA 最近的一个 MA 去接替它的任务。同时，更新任务 Agent 记录表中的记录，相应的子查询对应的值更新为新指派的 MA 的值，且将其状态值改为 1。直到 counter = 0，表示所有的子查询结果都返回了，可以对生成的所有 MA 撤销，MA 记录表和状态表清空，将所有子查询结果送到广播子系统。

(5) 广播子系统

广播子系统将每个新的查询语句都记录到查询语移动客户机的当前位置。

2.3.4 通信机制

模型中各个 Agent 之间是通过发送消息^[8]来实现通信，相互协调，共同完成任务的。各个消息定义如下：

(1) help 消息：当 MA 遇到阻塞或其他问题不能完成任务时，向 Agent 管理器发送 help 消息，将自己的 id 号告知管理器。

句表中，对表中已存在的查询语句，将其次数进行更新，并进行统计，当查询效率达到一定数值时，系统以文本形式存储这些语句。对返回的子查询结果按照客户的请求整合结果，将此结果也按规则记录到查询语句表中，再从位置服务器得到客户机的当前位置，将最后的结果发送到客户。同时，以一定的广播周期，将查询语句表中出现频率高的查询语句的相关信息，按照一定的广播调度算法对移动终端用户进行广播。

(6) 位置查询服务器

位置查询服务器^[7]主要负责跟踪各移动客户机的位置，并告知广播子系统客户机当前所处位置。每个移动客户机都需要向当地的位置查询服务器（宿主服务器）做永久性的注册，而当一个移动客户机离开该宿主服务器的管辖范围，进入另一个位置查询服务器的辖区时，要向该位置查询服务器登记，同时向宿主服务器报告，这样，位置查询服务器就可以跟踪各个

(2) tell 消息：当某一 MA 向 Agent 管理器发出 help 消息后，Agent 管理器就从 MA 状态表中找到一个目前处于空闲状态且距离阻塞的 MA 最近的一个 Agent，向它发送 tell 消息，告知去接替阻塞的 MA 完成子查询任务，同时携带新任务的内容，工作站点等参数。

2.4 各 Agent 的算法描述

(1) Agent 管理器的算法描述：

生成 MA，指派各 MA 到相应的子查询的工作站点；

counter = MA 的个数；

MA 状态表的值改为 1；

MA 记录表中记录 MA 与子查询的相关信息；

While (counter > 0)

{ 监听各 MA 的状态；

if (有 MA 成功返回)

接收结果，将状态表中的值改为 0，且

counter -1；

if (收到 help 消息)

{ 查询状态表；

if (当前没有状态为 0 的 MA)

等待；

else

按照就近原则选择一个 MA，向它发送

tell 消息；

}

}

(2) MA (移动 Agent) 算法描述：

被 Agent 管理器派往工作站点执行子查询任务；

if (遇到阻塞等问题)

{ 发送 help 消息到 Agent 管理器；

自动挂起；

}

if (收到 Agent 管理器的 tell 消息)

{ 按照消息指定的任务和地点去接替另一个 MA 未完成的任务；

完成子查询任务，成功返回结果；

}

else 完成子查询任务，返回结果到 Agent 管理器；

3 实验结果及分析

本次实验是在局域网的环境下，利用 5 台微机，一台模拟查询服务器，其余的模拟客户端向服务器发出查询请求。由于条件的限制，没有实现移动客户端的频繁断接，客户端是一直与服务器端处于连接状态的。操作系统都是 Windows xp，数据库系统是 SQL Server 2000。在基于 Java 的移动 Agent 系统平台 Aglet Workbench 上进行移动 Agent 的创建、迁移、交互、协作等操作，采用基于通信原语的 KQML (knowledge query and manipulation language) 实现 Agent 的信息交换，移动 Agent 的传输协议为 ATP (Agent Transfer Protocol)。

选取对学生成绩数据的查询作为应用示例，同时采用有 10%重复率的查询语句，在上述模拟环境下进

行了传统的分布式查询与基于移动 Agent 广播技术的查询的仿真对比实验 (如图 2 所示)。图中成功率的值是在选定一段合适长度时间内查询请求成功的语句数与总的查询语句数的百分比，实验结果表明，随着查

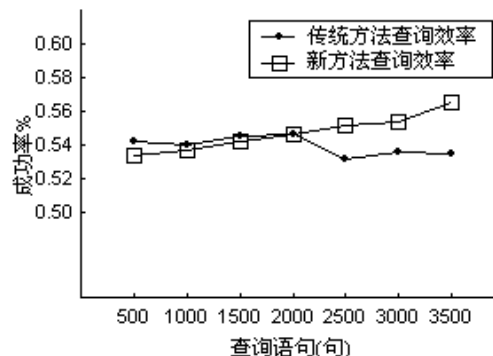


图 2 仿真对比实验结果

询语句的增加，采用基于移动 Agent 广播技术的查询可以提高系统的查询效率。

分析原因可以发现，当系统刚开始运行时，由于存储在查询语句表中的内容较少，此时的查询过程与传统的分布式查询过程一样，每一个查询操作都需要经过查询分解，由任务 Agent 完成子查询，最后再对子查询结果整合、发送给客户端这一过程，所以刚开始的效率还不如前者。不过系统运行一段时间后，随着查询语句表中存储结果的增加，而且还通过广播子系统将高频结果发送给各移动用户，使得一些查询操作可以直接得到结果，从而缩短了查询时间，查询请求成功的语句数也增加了，也提高了系统的效率。

4 结束语

本文所阐述的多 Agent 的移动数据库查询优化模型充分利用了移动 Agent 和数据广播技术，并应用于移动数据库查询系统。移动 Agent 的应用，使得系统可以在网络频繁断接的条件下工作，具有较好的可扩展性。数据广播的应用，通过对热点数据的组织和广播，并利用移动终端的存储能力，查询客户端有可能通过接收广播数据就可以得到查询数据，提高访问的效率，并进一步减少对无线网络的依赖性。当然，本文所设计的模型还不够细致，移动 Agent 的移动性实现技术包括路由选择、动态负载均衡等还有待详细实现，通信机制也还需要进一步的完善和优化，而且热点数据的组织算法和广播调度算法也还有待进一步研究。

References (参考文献)

- [1] ALI S. Analysis on the performance of mobile agents for query retrieval [J]. Information Sciences, 2005, 172(3):28 2-283.
- [2] Cao Jian, Li Minglu. Service Process Configuration Based on

- Multi-Agent Negotiation [J].Chinese Journal of Computers, 2006,29(7):1118-1119.曹健, 李明禄.基于多 Agent 协商的服务流程定制[J].计算机学报,2006,29(7):1118-1119.
- [3] QU Haijun, LU Jukang. Distributed Information Query System Based on Mobile Agent [J].Computer Engineering, 2003,29(7):140-141.瞿海军, 陆菊康. 基于移动 Agent 的分布式信息查询系统[J]. 计算机工程,2003,29(7):140-141.
- [4] Xu Ming,Cao Jiannong. Mobile computing[M].Beijing, Tsinghua University Press,2008,140-155.徐明,曹建农等.移动计算技术[M].北京, 清华大学出版社,2008,140-155.
- [5] Zhang Wendong, Shi Xiaoyan. New method for distributed query processing in distributed database [J]. Computer Engineering and Design, 2007, 28(19):4600-4601.张文东, 石小艳.分布式数据库中查询处理的新方法研究[J].计算机工程与设计,2007,28(19):4600-4601.
- [6] Zhang Qiang.Reserch an query optimization technology of mobile database based on mobile agent[D].Harbin University of Science and Techology,2007,21-36.张强.基于移动 Agent 的移动数据库查询优化技术的研究[D].哈尔滨理工大学,2007,21-36.
- [7] YON D C. An efficient broadcast data clustering method for multipoint queries in wireless information systems [J]. Journal of Systems and Software, 2002, 64(3):173-174.
- [8] AGUSTINUS B W. Research in mobile database query optimization and processing[J]. Mobile Information Systems, 2005, 1(4):230-231.
- [9] POOYA F . A multi-agent architecture for control of AGV systems [J]. Robotics and Computer -Integrated Manufacturing, 2004, 20(9):475-476.