



(12)发明专利

(10)授权公告号 CN 106933906 B

(45)授权公告日 2020.05.22

(21)申请号 201511032355.X

(22)申请日 2015.12.31

(65)同一申请的已公布的文献号

申请公布号 CN 106933906 A

(43)申请公布日 2017.07.07

(73)专利权人 北京国双科技有限公司

地址 100083 北京市海淀区北四环中路229
号海泰大厦4层南401号

(72)发明人 洪超 黄健 何恺铎

(74)专利代理机构 北京康信知识产权代理有限
责任公司 11240

代理人 韩建伟 张永明

(51)Int.Cl.

G06F 16/28(2019.01)

(56)对比文件

CN 104462434 A,2015.03.25,

CN 103020301 A,2013.04.03,

CN 104063376 A,2014.09.24,

CN 105095249 A,2015.11.25,

CN 102667774 A,2012.09.12,

US 9081825 B1,2015.07.14,

审查员 张敏

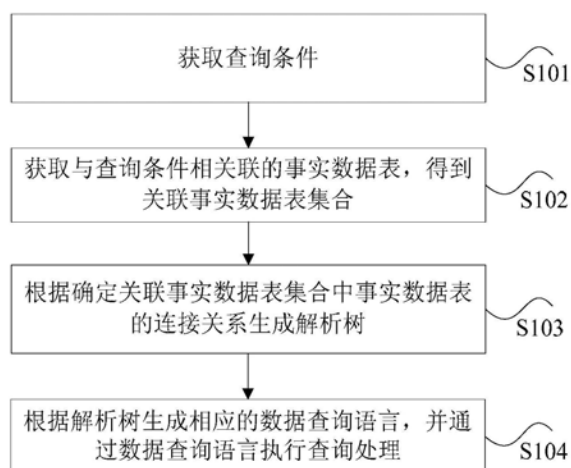
权利要求书2页 说明书9页 附图1页

(54)发明名称

数据多维度的查询方法及装置

(57)摘要

本申请公开了一种数据多维度的查询方法及装置。该方法包括:获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;获取与查询条件相关联的事实数据表,得到关联事实数据表集合;根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理。通过本申请,解决了数据多维度查询效率较低的问题。



1. 一种数据多维度的查询方法,其特征在于,包括:

获取查询条件,其中,所述查询条件包括维度剖析链和待查询指标,所述维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;

获取与所述查询条件相关联的事实数据表,得到关联事实数据表集合;

根据所述关联事实数据表集合中事实数据表的连接关系生成解析树,其中,所述解析树是由所述查询条件和所述连接关系生成并用于反应数据流过程的数据结构;以及

根据所述解析树生成相应的数据查询语言,并通过所述数据查询语言执行查询处理;

其中,获取与所述查询条件相关联的事实数据表,得到关联事实数据表集合包括:

获取与所述查询条件相关联的事实数据表,得到多个关联事实数据表;

获取关联键,其中,所述关联键用于关联所述多个关联事实数据表;以及

通过所述关联键关联所述多个关联事实数据表,得到所述关联事实数据表集合。

2. 根据权利要求1所述的方法,其特征在于,在获取与所述查询条件相关联的事实数据表,得到关联事实数据表集合之后,在根据所述关联事实数据表集合中事实数据表的连接关系生成解析树之前,所述方法还包括:

获取所述关联事实数据表集合中各事实数据表对应的连接键值;以及

根据所述连接键值将所述关联事实数据表集合中的各个事实数据表进行连接。

3. 根据权利要求2所述的方法,其特征在于,根据所述关联事实数据表集合中事实数据表的连接关系生成解析树包括:

确定所述关联事实数据表集合中各个事实数据表之间连接的顺序关系;

按照预设条件调整各个事实数据表之间连接的顺序关系;以及

根据调整后的各个事实数据表之间连接的顺序关系生成所述解析树。

4. 根据权利要求1所述的方法,其特征在于,获取所述查询条件包括:

接收外部输入的查询指令;

根据所述查询指令确定所述维度剖析链和所述待查询指标;以及

将所述维度剖析链和所述待查询指标作为所述查询条件。

5. 一种数据多维度的查询装置,其特征在于,包括:

第一获取单元,用于获取查询条件,其中,所述查询条件包括维度剖析链和待查询指标,所述维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;

第二获取单元,用于获取与所述查询条件相关联的事实数据表,得到关联事实数据表集合;

生成单元,用于根据所述关联事实数据表集合中事实数据表的连接关系生成解析树,其中,所述解析树是由所述查询条件和所述连接关系生成并用于反应数据流过程的数据结构;以及

处理单元,用于根据所述解析树生成相应的数据查询语言,并通过所述数据查询语言执行查询处理;

其中,所述第二获取单元包括:

第一获取模块,用于获取与所述查询条件相关联的事实数据表,得到多个关联事实数据表;

第二获取模块,用于获取关联键,其中,所述关联键用于关联所述多个关联事实数据

表;以及

关联模块,用于通过所述关联键关联所述多个关联事实数据表,得到所述关联事实数据表集合。

6.根据权利要求5所述的装置,其特征在于,所述装置还包括:

第三获取单元,用于获取所述关联事实数据表集合中各事实数据表对应的连接键值;以及

连接单元,用于根据所述连接键值将所述关联事实数据表集合中的各个事实数据表进行连接。

7.根据权利要求6所述的装置,其特征在于,所述生成单元包括:

第一确定模块,用于确定所述关联事实数据表集合中各个事实数据表之间连接的顺序关系;

调整模块,用于按照预设条件调整各个事实数据表之间连接的顺序关系;以及

生成模块,用于根据调整后的各个事实数据表之间连接的顺序关系生成所述解析树。

8.根据权利要求5所述的装置,其特征在于,所述第一获取单元包括:

接收模块,用于接收外部输入的查询指令;

第二确定模块,用于根据所述查询指令确定所述维度剖析链和所述待查询指标;以及

第三确定模块,用于将所述维度剖析链和所述待查询指标作为所述查询条件。

数据多维度的查询方法及装置

技术领域

[0001] 本申请涉及数据处理领域,具体而言,涉及一种数据多维度的查询方法及装置。

背景技术

[0002] 在联机分析处理(On-LineAnalytical Processing,简称OLAP)的环境下,目前常用的查询引擎有Hive,Impala这些结构化的带元数据的数据库,在这些数据库中,通常,将各类事件的数据记录在各个事实数据表中。例如,在“教学管理系统”中,教学管理”数据库包含以下几个事实数据表:“教师”表、“课程”表、“成绩”表、“学生”表、“班级”表和“授课”表等,通过以上事实数据表用来管理教学过程中学生、教师、课程等信息。再例如,互联网进行的会话监测,一般会有会话表(Session),页面浏览表(PageView),站内搜索表(SiteSearch),订单表(Ecommerce),自定义事件表(Event)等很多代表各种业务场景的事实数据表,但都会通过客户的会话标识(SessionID)进行关联起来,形成整个会话的所有实体。当用户需要从多个角度跨越多个事实数据表将数据关联进行查看指标数据和剖析维度时,相关技术中,根据用户查询需求,编写相应代码,从而在数据库中去查询相关的指标数据和相关维度,由于相关技术中缺乏统一将用户查询需求转化为查询的技术方案,导致在分析领域中的数据多维度查询效率较低。

[0003] 针对相关技术中数据多维度查询效率较低的问题,目前尚未提出有效的解决方案。

发明内容

[0004] 本申请的主要目的在于提供一种数据多维度的查询方法及装置,以解决数据多维度查询效率较低的问题。

[0005] 为了实现上述目的,根据本申请的一个方面,提供了一种数据多维度的查询方法。该方法包括:获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;获取与查询条件相关联的事实数据表,得到关联事实数据表集合;根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;以及根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理。

[0006] 进一步地,在获取与查询条件相关联的事实数据表,得到关联事实数据表集合之后,在根据关联事实数据表集合中事实数据表的连接关系生成解析树之前,该方法还包括:获取关联事实数据表集合中各事实数据表对应的连接键值;以及根据连接键值将关联事实数据表集合中的各个事实数据表进行连接。

[0007] 进一步地,根据关联事实数据表集合中事实数据表的连接关系生成解析树包括:确定关联事实数据表集合中各个事实数据表之间连接的顺序关系;按照预设条件调整各个事实数据表之间连接的顺序关系;以及根据调整后的各个事实数据表之间连接的顺序关系生成解析树。

[0008] 进一步地,获取与查询条件相关联的事实数据表,得到关联事实数据表集合包括:获取与查询条件相关联的事实数据表,得到多个关联事实数据表;获取关联键,其中,关联键用于关联多个关联事实数据表;以及通过关联键关联多个关联事实数据表,得到关联事实数据表集合。

[0009] 进一步地,获取查询条件包括:接收外部输入的查询指令;根据查询指令确定维度剖析链和待查询指标;以及将维度剖析链和待查询指标作为查询条件。

[0010] 为了实现上述目的,根据本申请的另一方面,提供了一种数据多维度的查询装置。该装置包括:第一获取单元,用于获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;第二获取单元,用于获取与查询条件相关联的事实数据表,得到关联事实数据表集合;生成单元,用于根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;以及处理单元,用于根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理。

[0011] 进一步地,该装置还包括:第三获取单元,用于获取关联事实数据表集合中各事实数据表对应的连接键值;以及连接单元,用于根据连接键值将关联事实数据表集合中的各个事实数据表进行连接。

[0012] 进一步地,生成单元包括:第一确定模块,用于确定关联事实数据表集合中各个事实数据表之间连接的顺序关系;调整模块,用于按照预设条件调整各个事实数据表之间连接的顺序关系;以及生成模块,用于根据调整后的各个事实数据表之间连接的顺序关系生成解析树。

[0013] 进一步地,第二获取单元包括:第一获取模块,用于获取与查询条件相关联的事实数据表,得到多个关联事实数据表;第二获取模块,用于获取关联键,其中,关联键用于关联多个关联事实数据表;以及关联模块,用于通过关联键关联多个关联事实数据表,得到关联事实数据表集合。

[0014] 进一步地,第一获取单元包括:接收模块,用于接收外部输入的查询指令;第二确定模块,用于根据查询指令确定维度剖析链和待查询指标;以及第三确定模块,用于将维度剖析链和待查询指标作为查询条件。

[0015] 通过本申请,采用以下步骤:获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;获取与查询条件相关联的事实数据表,得到关联事实数据表集合;根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理,解决了数据多维度查询效率较低的问题,进而达到了提升数据多维度查询效率的效果。

附图说明

[0016] 构成本申请的一部分的附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0017] 图1是根据本申请实施例的数据多维度的查询方法的流程图;以及

[0018] 图2是根据本申请实施例的数据多维度的查询装置的示意图。

具体实施方式

[0019] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0020] 为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分的实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本申请保护的范围。

[0021] 需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0022] 为了便于理解,以下对本申请实施例涉及的几个术语进行解释说明:

[0023] 指标:是指可以进行聚合运算的值,例如,综合浏览量是一个指标,使用求和进行聚合;平均停留时间也是一个指标,使用平均进行聚合,其中,聚合运算包括求和、平均、计数等。

[0024] 维度:是指查看指标的角度,例如,浏览器是一个维度,可以从浏览器这个维度查询页面浏览(Page View,简称为PV),从而可以知道用户使用哪些浏览器查看页面,以及使用这些浏览器查看页面的次数;操作系统为另一个维度,也可以从操作系统这个维度查询PV,从而可以知道用户使用哪些操作系统查看页面,以及使用这些操作系统查看页面的次数。

[0025] 成员:是指维度具体对应的对象,例如,浏览器是一个维度,而IE浏览器、Chrome浏览器即为该维度的一个成员。

[0026] 多维剖析:是指可以从多个维度剖析多个指标的值,例如,从操作系统维度查看会话量值后,选择其中几个成员后,继续从浏览器维度查看会话值,即从两个维度剖析会话值。依次类推,可以从多个维度进行自由的剖析。

[0027] 维度表:维度表可以看作是用户来分析数据的窗口,维度表中包含事实数据表中事实记录的特性,有些特性提供描述性信息,有些特性指定如何汇总事实数据表数据,以便为分析者提供有用的信息,维度表包含帮助汇总数据的特性的层次结构。例如,包含产品信息的维度表通常包含将产品分为食品、饮料、非消费品等若干类的层次结构,这些产品中的每一类进一步多次细分,直到各产品达到最低级别。在维度表中,每个表都包含独立于其他维度表的事实特性,例如,客户维度表包含有关客户的数据。维度表中的列字段可以将信息分为不同层次的结构级。

[0028] 事实数据表:数据仓库架构中的中央表,它包含联系事实与维度表的数字度量值和键。事实数据表包含描述业务(如银行事务或产品销售)内特定事件的数据。一般来说,一

个事实数据表都要和一个或多个维度表相关联,用户在利用事实数据表创建多维数据集时,可以使用一个或多个维度表。根据本申请的实施例,提供了一种数据多维度的查询方法。

[0029] 图1是根据本申请实施例的数据多维度的查询方法的流程图。如图1所示,该方法包括以下步骤:

[0030] 步骤S101,获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤。

[0031] 多维剖析时,会传入一个维度剖析链(预设过滤条件)和待查询指标,根据剖析链对多个事实数据表进行过滤,例如,维度剖析链为:操作系统版本(window7)>>浏览器(chrome)>>是否新访问者(是)>>页面查看量(1)>>访问时长。待查询指标是访问量等。

[0032] 获取包括维度剖析链和待查询指标的查询条件。

[0033] 可选地,在申请实施例的数据多维度的查询方法中,获取查询条件还可以通过以下步骤实现:接收外部输入的查询指令;根据查询指令确定维度剖析链和待查询指标;以及将维度剖析链和待查询指标作为查询条件。

[0034] 在获取查询条件之前,需要配置维度和指标的元数据信息,确定维度和指标的相关属性,例如,维度配置的一个示例:

[0035] "Dimensions":[

[0036] {

[0037] "name":"SourceType",//名称,类型

[0038] "column":"sourcetype",//列,类型

[0039] "Tables":["session","pageview","ecommerce","sitesearch","mouseclick","event"]},//表:会话表,页面游览表,订单表,网站搜索表,点击表,自定义事件表{

[0040] "name":"ChannelName",//名称,来源渠道名称

[0041] "column":"channelname",//列

[0042] "Tables":["session","pageview","ecommerce","sitesearch","mouseclick","event"]}//表:会话表,页面游览表,订单表,网站搜索表,点击表,自定义事件表]

[0043] 指标配置的一个示例:

[0044] "Measures":[

[0045] {

[0046] "Name":"Visits",//名称,访问量

[0047] "Table":"session",//表,会话表

[0048] "Column":null,

[0049] "AggregateType":"COUNT"

[0050] },

[0051] {

[0052] "Name":"BouncesVisits",//名称,访问量

[0053] "Table":"session",//表,会话表

[0054] "Column": "isbouce", //列, isbouce列

[0055] "AggregateType": "SUM"

[0056] }

[0057]]

[0058] 步骤S102, 获取与查询条件相关联的事实数据表, 得到关联事实数据表集合。

[0059] 维度(dimension)是多维数据集的结构特性。它们是事实数据表中用来描述数据的分类的有组织层次结构(级别)。这些分类和级别描述了一些相似的成员集合, 用户将基于这些成员集合进行分析。维度表可以看作是用户来分析数据的窗口, 维度表中包含事实数据表中事实记录的特性, 有些特性提供描述性信息, 有些特性指定如何汇总事实数据表数据, 以便为分析者提供有用的信息, 维度表包含帮助汇总数据的特性的层次结构。例如, 事实数据表为销量表, 维度表就是地区表。剖析某地区商品的销量, 是从地区这个角度观察商品销量, 剖析北京、上海和广州等地区商品的销量, 是从多个角度观察商品销量。

[0060] 步骤S102的查询条件中的至少一个维度为本申请实施例的数据多维度自由剖析的查询方法中需要进行剖析的维度, 在分布式数据库中, 获取与需要进行剖析的维度组成的维度剖析链相关联的事实数据表, 得到多个事实数据表, 多个事实数据表组成关联事实数据表集合。

[0061] 通过上述步骤, 获取到与查询条件相关联的多个事实数据表, 即获取到后续查询数据的元数据信息。

[0062] 可选地, 为了提升获取关联事实数据表集合的准确性, 在本申请实施例的数据多维度自由剖析的查询方法中, 获取与查询条件相关联的事实数据表, 得到关联事实数据表集合还可以通过以下步骤实现: 获取与查询条件相关联的事实数据表, 得到多个关联事实数据表; 获取关联键, 其中, 关联键用于关联多个关联事实数据表; 以及通过关联键关联多个关联事实数据表, 得到关联事实数据表集合。

[0063] 例如, 关联键为会话标识(SessionID), 通过会话标识(SessionID) 关联多个关联事实数据表, 得到关联事实数据表集合。

[0064] 需要说明的是, 本申请提供的数据多维度自由剖析的查询方法为在分析领域中的数据多维度自由剖析的查询方法。

[0065] 步骤S103, 根据关联事实数据表集合中事实数据表的连接关系生成解析树, 其中, 解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构。

[0066] 在本申请提供的数据多维度自由剖析的查询方法中, 解析树是一种数据结构, 该结构记录了生成结构化查询语言所需要的所有条件和样式, 根据查询需求的维度链, 剖析维度和指标(即上述的查询条件), 计算出所需要的表结点以及过滤信息, 并将所有表结点通过相关的关联关系连接起来。其中, 解析树中的每一个节点, 相当于定义了一个临时表数据结构。整个解析树反映了如何从原始的数据表中通过相应的过滤、选取、变化, 逐步得到最终结果的过程, 解析树中记录了所有的数据表结点。

[0067] 可选地, 在本申请实施例的数据多维度自由剖析的查询方法中, 在获取与查询条件相关联的事实数据表, 得到关联事实数据表集合之后, 在根据关联事实数据表集合中事实数据表的连接关系生成解析树之前, 该方法还包括: 获取关联事实数据表集合中各事实数据表对应的连接键值; 以及根据连接键值将关联事实数据表集合中的各个事实数据表进

行连接。

[0068] 例如,关联事实数据表集合中各事实数据表对应的连接键值为相应的ID等共有维度字段,根据ID等共有维度字段将关联事实数据表集合中的各个事实数据表进行连接。

[0069] 可选地,在本申请实施例的数据多维度自由剖析的查询方法中,根据关联事实数据表集合中事实数据表的连接关系生成解析树还可以通过以下步骤实现:确定关联事实数据表集合中各个事实数据表之间连接的顺序关系;按照预设条件调整各个事实数据表之间连接的顺序关系;以及根据调整后的各个事实数据表之间连接的顺序关系生成解析树。

[0070] 各个事实数据表中包括大表和小表,在本申请中的将数据表中数据量超过阈值的表称为大表,数据表中数据量未超过阈值的表称为小表。由于各表之间的链接顺序会极大的影响性能开销,在本申请中按照预设条件调节各个事实数据表之间连接的顺序关系,可以将小表与小表先连接,再与大表连接,优化表的连接顺序,从而减小性能开销,极大的提高性能。

[0071] 步骤S104,根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理。

[0072] 本申请实施例的数据多维度自由剖析的查询方法中的数据查询语言为结构化查询语言(Structured Query Language,简称SQL),SQL为数据库查询和程序设计语言,用于存取数据以及查询、更新和管理关系数据库系统。

[0073] 例如,维度剖析链为:IsBounce=0AND PageDuration>10AND SessionPageViews>10;剖析维度为:OSBrief;待查询指标为:Visits,AvgSessionDuration,SessionpageViews。

[0074] 获取到与维度剖析链、剖析维度和待查询指标相关联的事实数据表,根据关联事实数据表集合中事实数据表的连接关系生成解析树,根据解析树生成相应的数据查询语言。

[0075] 例如,实现代码可以如下所示:

[0076] with s as(select serversessionid,osbrief,pvcount,sessionduration from session where isbounce=0and PVCount>10)

[0077] pv as(select serversessionid from pageview group by serversessionid having max(pageduration)>10)

[0078] select osbrief,sum(pvcount)as sessionpageviews,count(*)as visits,avg(sessionduration)as avgsessionduration

[0079] from s inner join pv on s.serversessionid=pv.serversessionid

[0080] group by osbrief

[0081] order by count(*)desc

[0082] limit 10

[0083] 通过上述生成的数据查询语言执行查询处理。

[0084] 又例如,维度剖析链为:IsBounce=0AND PageDuration>10AND SessionPageView>10;剖析维度:OSBrief;待查询指标为:Visits,AvgSessionDuratuon,SessionpageViews。

[0085] 获取到与维度剖析链、剖析维度和待查询指标相关联的事实数据表,根据关联事

实数据表集合中事实数据表的连接关系生成解析树,根据解析树生成相应的数据查询语言。

[0086] 实现代码可以如下所示:

[0087] `PV_Jion_Session(SelectColumns:osbrief,Measures GroupBy:OSBrief
OrderBy:Visits desc paging:1,10)`

[0088] 通过上述生成的数据查询语言执行查询处理。

[0089] 根据上述生成的数据查询语言执行数据查询,解决了现有技术中由于相关技术中缺乏统一将用户查询需求转化为查询的技术方案,导致分析领域中的数据多维度查询效率较低的问题。

[0090] 对于数据仓库,各事实数据表都可能具有相应的关系,并且通过相应的ID等共有维度字段会进行关联,本申请实施例提供的数据多维度的查询方法处理的是展平后的事实数据表,例如,会话等字段会派生到其子实体,如页面浏览,电商订单等事实数据表上,这样在仅剖析会话类的维度,如浏览器等,就不需要进行会话表的关联,以达到性能提升的目的。

[0091] 本申请实施例提供的数据多维度的查询方法,通过获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;获取与查询条件相关联的事实数据表,得到关联事实数据表集合;根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理,解决了数据多维度查询效率较低的问题,进而达到了提升数据多维度查询效率的效果。

[0092] 需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行,并且,虽然在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤。

[0093] 本申请实施例还提供了一种数据多维度的查询装置,需要说明的是,本申请实施例的数据多维度的查询装置可以用于执行本申请实施例所提供的用于数据多维度的查询方法。以下对本申请实施例提供的数据多维度的查询装置进行介绍。

[0094] 图2是根据本申请实施例的数据多维度的查询装置的示意图。如图2所示,该装置包括:第一获取单元10、第二获取单元20、生成单元30和处理单元40。

[0095] 第一获取单元10,用于获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤。

[0096] 第二获取单元20,用于获取与查询条件相关联的事实数据表,得到关联事实数据表集合。

[0097] 生成单元30,用于根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构。

[0098] 处理单元40,用于根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理。

[0099] 本申请实施例提供的数据多维度的查询装置,通过第一获取单元10获取查询条件,其中,查询条件包括维度剖析链和待查询指标,维度剖析链由至少一个维度组成,用于对多个事实数据表进行过滤;第二获取单元20获取与查询条件相关联的事实数据表,得到

关联事实数据表集合;生成单元30根据关联事实数据表集合中事实数据表的连接关系生成解析树,其中,解析树是由查询条件和连接关系生成并用于反应数据流过程的数据结构;处理单元40根据解析树生成相应的数据查询语言,并通过数据查询语言执行查询处理,解决了数据多维度查询效率较低的问题,进而达到了提升数据多维度查询效率的效果。

[0100] 可选地,在本申请实施例提供的数据多维度的查询装置中,该装置还包括:第三获取单元,用于获取关联事实数据表集合中各事实数据表对应的连接键值;以及连接单元,用于根据连接键值将关联事实数据表集合中的各个事实数据表进行连接。

[0101] 可选地,在本申请实施例提供的数据多维度的查询装置中,生成单元30包括:第一确定模块,用于确定关联事实数据表集合中各个事实数据表之间连接的顺序关系;调整模块,用于按照预设条件调整各个事实数据表之间连接的顺序关系;以及生成模块,用于根据调整后的各个事实数据表之间连接的顺序关系生成解析树。

[0102] 可选地,在本申请实施例提供的数据多维度的查询装置中,第二获取单元20包括:第一获取模块,用于获取与查询条件相关联的事实数据表,得到多个关联事实数据表;第二获取模块,用于获取关联键,其中,关联键用于关联多个关联事实数据表;以及关联模块,用于通过关联键关联多个关联事实数据表,得到关联事实数据表集合。

[0103] 可选地,在本申请实施例提供的数据多维度的查询装置中,第一获取单元10包括:接收模块,用于接收外部输入的查询指令;第二确定模块,用于根据查询指令确定维度剖析链和待查询指标;以及第三确定模块,用于将维度剖析链和待查询指标作为查询条件。

[0104] 需要说明的是,对于前述的各方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本申请并不受所描述的动作顺序的限制,因为依据本申请,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定是本申请所必须的。

[0105] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述。

[0106] 在本申请所提供的几个实施例中,应该理解到,所揭露的装置,可通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。

[0107] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0108] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0109] 显然,本领域的技术人员应该明白,上述的本申请的各模块或各步骤可以用通用的计算装置来实现,它们可以集中在单个的计算装置上,或者分布在多个计算装置所组成的网络上,可选地,它们可以用计算装置可执行的程序代码来实现,从而,可以将它们存储

在存储装置中由计算装置来执行,或者将它们分别制作成各个集成电路模块,或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样,本申请不限制于任何特定的硬件和软件结合。

[0110] 以上所述仅为本申请的优选实施例,并不用于限制本申请,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

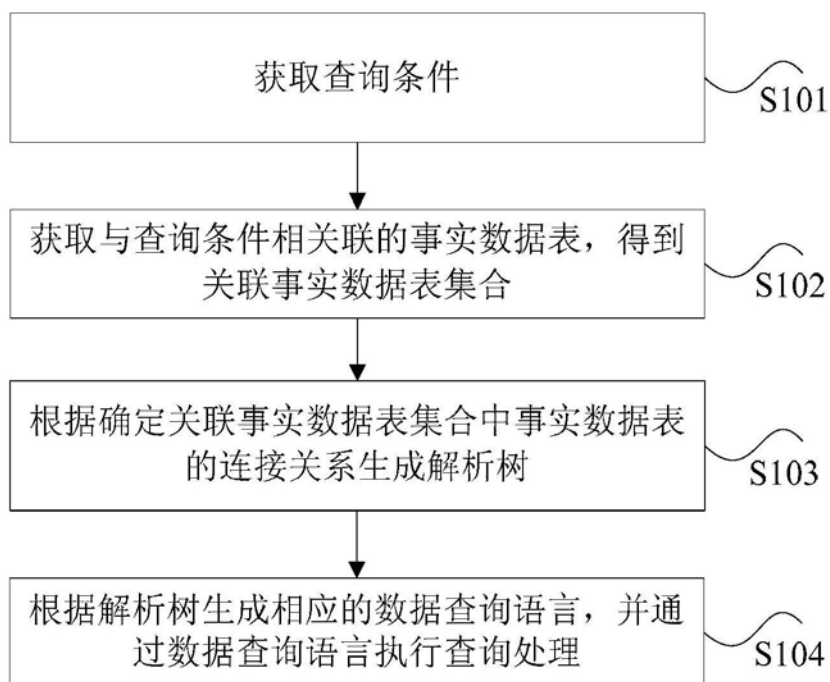


图1



图2