



(12)发明专利

(10)授权公告号 CN 106933926 B

(45)授权公告日 2019.11.12

(21)申请号 201511032703.3

(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 106933926 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 何恺铎 黄健

(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 16/242(2019.01)

G06F 16/2455(2019.01)

(56)对比文件

CN 104408169 A,2015.03.11,

CN 105045877 A,2015.11.11,

CN 102334098 A,2012.01.25,

US 2014214796 A1,2014.07.31,

张崇.基于GIS的农产品指标数据库表结构
设计研究.《计算机工程与设计》.2011,第32卷
(第1期),

审查员 吕姣

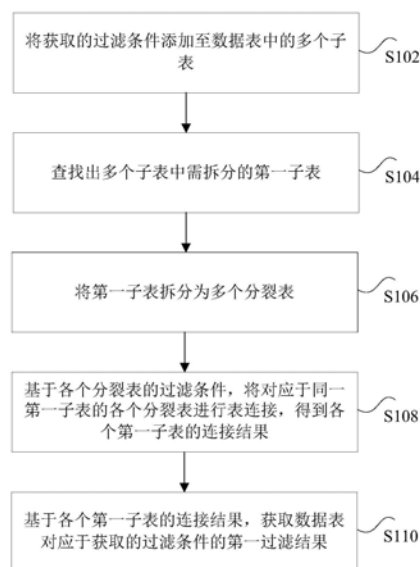
权利要求书2页 说明书9页 附图2页

(54)发明名称

数据表的过滤方法及装置

(57)摘要

本申请公开了一种数据表的过滤方法及装置。其中,该方法包括:将获取的过滤条件添加至数据表中的多个子表;查找出多个子表中需拆分的第一子表;将第一子表拆分为多个分裂表;基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果;基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果。本申请解决了对数据表进行同一维度的多次且过滤时的处理效率低的技术问题。



1. 一种数据表的过滤方法,其特征在于,包括:

将获取的过滤条件添加至数据表中的多个子表;

查找出所述多个子表中需拆分的第一子表,其中,所述第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一所述维度的多个指标之间具有逻辑且关系;

将所述第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件;

基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个所述第一子表的连接结果;

基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果;

其中,查找出所述多个子表中需拆分的第一子表包括:

遍历所述数据表的所述多个子表的过滤条件;

判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操作的过滤条件;

若所述当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将所述当前遍历到的子表确定为所述第一子表。

2. 根据权利要求1所述的方法,其特征在于,将对应于同一第一子表的各个分裂表进行表连接包括:

按照内连接的方式连接对应于同一第一子表的各个所述分裂表。

3. 根据权利要求1所述的方法,其特征在于,基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果包括:

获取所述数据表中无需拆分的第二子表;

利用所述第二子表的过滤条件,对所述第二子表进行过滤,得到第二过滤结果;

获取所述第二子表的第二过滤结果和所述第一子表的连接结果的汇总结果,将所述汇总结果作为所述第一过滤结果。

4. 根据权利要求1所述的方法,其特征在于,在将获取的过滤条件添加至数据表中的多个子表之前,所述方法还包括:

根据过滤需求从数据库中获取所述过滤条件;或者

获取预先设置的所述过滤条件。

5. 一种数据表的过滤装置,其特征在于,包括:

添加单元,用于将获取的过滤条件添加至数据表中的多个子表;

查找单元,用于查找出所述多个子表中需拆分的第一子表,其中,所述第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一所述维度的多个指标之间具有逻辑且关系;

拆分单元,用于将所述第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件;

表连接单元,用于基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个所述第一子表的连接结果;

获取单元,用于基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果;

其中,所述查找单元包括:

遍历模块,用于遍历所述数据表的所述多个子表的过滤条件;

判断模块,用于判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操作的过滤条件;

确定模块,用于若所述当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将所述当前遍历到的子表确定为所述第一子表。

6. 根据权利要求5所述的装置,其特征在于,所述表连接单元包括:

连接模块,用于按照内连接的方式连接对应于同一第一子表的各个所述分裂表。

7. 根据权利要求5所述的装置,其特征在于,所述获取单元包括:

第一获取模块,用于获取所述数据表中无需拆分的第二子表;

过滤模块,用于利用所述第二子表的过滤条件,对所述第二子表进行过滤,得到第二过滤结果;

处理模块,用于获取所述第二子表的第二过滤结果和所述第一子表的连接结果的汇总结果,将所述汇总结果作为所述第一过滤结果。

8. 根据权利要求5所述的装置,其特征在于,所述装置还包括:

第二获取模块,用于在将获取的过滤条件添加至数据表中的多个子表之前,根据过滤需求从数据库中获取所述过滤条件;或者,

第三获取模块,用于获取预先设置的所述过滤条件。

数据表的过滤方法及装置

技术领域

[0001] 本申请涉及数据处理领域,具体而言,涉及一种数据表的过滤方法及装置。

背景技术

[0002] 数据表中的数据一般为多对多的关系,其中,多对一的关系为较常见的一种,由于电商和会话是多对一的关系,在一个会话中可能买多个订单,每个订单包含多种产品,所以会出来同一产品维度多次进行的且过滤的维度链。在现有技术中,对该同一维度多次且过滤的情况,一般采用直接用该过滤条件对数据表进行过滤,这样做会产生数据表的过滤数据量较大,使得表过滤的开销很大,效率较低。

[0003] 针对上述的对数据表进行同一维度的多次且过滤时的处理效率低的问题,目前尚未提出有效的解决方案。

发明内容

[0004] 本申请实施例提供了一种数据表的过滤方法及装置,以至少解决对数据表进行同一维度的多次且过滤时的处理效率低的技术问题。

[0005] 根据本申请实施例的一个方面,提供了一种数据表的过滤方法,该方法包括:将获取的过滤条件添加至数据表中的多个子表;查找出所述多个子表中需拆分的第一子表,其中,所述第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一维度的多个指标之间具有逻辑且关系;将所述第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件;基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个所述第一子表的连接结果;基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果。

[0006] 进一步地,将对应于同一第一子表的各个分裂表进行表连接包括:按照内连接的方式连接对应于同一第一子表的各个所述分裂表。

[0007] 进一步地,基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果包括:获取所述数据表中无需拆分的第二子表;利用所述第二子表的过滤条件,对所述第二子表进行过滤,得到第二过滤结果;获取所述第二子表的第二过滤结果和所述第一子表的连接结果的汇总结果,将所述汇总结果作为所述第一过滤结果。

[0008] 进一步地,在将获取的过滤条件添加至数据表中的多个子表之前,所述方法还包括:根据过滤需求从数据库中获取所述过滤条件;或者,获取预先设置的所述过滤条件。

[0009] 进一步地,查找出所述多个子表中需拆分的第一子表包括:遍历所述数据表的所述多个子表的过滤条件;判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操作的过滤条件;若所述当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将所述当前遍历到的子表确定为所述第一子表。

[0010] 根据本申请实施例的另一方面,还提供了一种数据表的过滤装置,该装置包括:添加单元,用于将获取的过滤条件添加至数据表中的多个子表;查找单元,用于查找出所述多个子表中需拆分的第一子表,其中,所述第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一维度的多个指标之间具有逻辑且关系;拆分单元,用于将所述第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件;表连接单元,用于基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个所述第一子表的连接结果;获取单元,用于基于各个所述第一子表的连接结果,获取所述数据表对应于所述获取的过滤条件的第一过滤结果。

[0011] 进一步地,所述表连接单元包括:连接模块,用于按照内连接的方式连接对应于同一第一子表的各个所述分裂表。

[0012] 进一步地,所述获取单元包括:第一获取模块,用于获取所述数据表中无需拆分的第二子表;过滤模块,用于利用所述第二子表的过滤条件,对所述第二子表进行过滤,得到第二过滤结果;处理模块,用于获取所述第二子表的第二过滤结果和所述第一子表的连接结果的汇总结果,将所述汇总结果作为所述第一过滤结果。

[0013] 进一步地,所述装置还包括:第二获取模块,用于在将获取的过滤条件添加至数据表中的多个子表之前,根据过滤需求从数据库中获取所述过滤条件;或者,第三获取模块,用于获取预先设置的所述过滤条件。

[0014] 进一步地,所述查找单元包括:遍历模块,用于遍历所述数据表的所述多个子表的过滤条件;判断模块,用于判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操作的过滤条件;确定模块,用于若所述当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将所述当前遍历到的子表确定为所述第一子表。

[0015] 在本申请实施例中,采用将用于过滤所述数据表的所有过滤条件分配给所述数据表中的多个子表;查找出需拆分的所述数据表中的子表;将所述需拆分的子表拆分为多个分裂表,所述拆分后的每个所述分裂表包括一个维度且过滤条件;基于所述过滤条件将所述拆分后的所述分裂表进行表连接,以得到新的子表的方式,通过将包括多个同一维度的多个且过滤的子表拆分成多个仅含有一个维度且过滤的分裂表,再将该拆分后的分裂表进行表连接来表示且的关系,相比于现有技术中的直接进行同一维度的多个且过滤,达到了减小且过滤的数据处理量的目的,从而实现了提高数据表进行同一维度的多次且过滤时的处理效率的技术效果,进而解决了对数据表进行同一维度的多次且过滤时的处理效率低的技术问题。

附图说明

[0016] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,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一种数据表的过滤方法的流程图;以及

[0018] 图2是根据本申请实施例的一种数据表的过滤装置的示意图。

具体实施方式

[0019] 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。

[0020] 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能够以除了在这里图示或描述的那些以外的顺序实施。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。

[0021] 实施例1

[0022] 根据本申请实施例,提供了一种数据表的过滤方法的实施例,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。

[0023] 图1是根据本申请实施例的一种数据表的过滤方法的流程图,如图1所示,该方法包括如下步骤:

[0024] 步骤S102,将获取的过滤条件添加至数据表中的多个子表。

[0025] 步骤S104,查找出多个子表中需拆分的第一子表,其中,第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一维度的多个指标之间具有逻辑且关系。

[0026] 步骤S106,将第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件。

[0027] 步骤S108,基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果。

[0028] 步骤S110,基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果。

[0029] 采用本申请实施例,通过采用将获取的过滤条件添加至数据表中的多个子表;查找出多个子表中需拆分的第一子表;将第一子表拆分为多个分裂表;基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果,基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果的方式,相比于现有技术中的直接对数据表进行具有多个指标的维度进行逻辑且操作的过滤(相当于同一维度的多个且过滤),本申请通过将数据表中的第一子表拆分为多个具有一个指标的维度进行逻辑且操作的过滤条件(相当于一个维度的且过滤)的分裂表,然后将各个分裂表进行表连接,达到了减小且过滤的数据处理量的目的,从而实现了提高数据表进行同一维度的多次且过滤时的处理效率的技术效果,进而解决了对数据表进行同一维度的多

次且过滤时的处理效率低的技术问题。

[0030] 上述实施例中的对具有多个指标的维度进行逻辑且操作的过滤条件,即同一维度具有多个指标。

[0031] 上述实施例中的方法可以应用在上述多对多关系的查询中,例如同一维度多次且过滤Sql Builder出现的场景中,在该场景下,可以通过如下形式的维度链和剖析维度来实现对数据的同一维度多次且过滤操作。

[0032] 其中,维度链表示用于实现且过滤的同一维度的筛选条件,例如,该维度链可以为:product like '%a%' AND product like '%b%',该维度链表示既买了产品名称包含'a'又买了产品名称包含'b'的会话,该剖析维度可以为:操作系统简介OS Brief,上述的维度链即维度链过滤的条件,该维度链中包含有并且,也即“AND”的关系,可看作同一产品维度的两个且过滤条件。

[0033] 同一维度且过滤可以理解为:两个过滤条件属于同一维度时,该两个过滤条件又存在“并且”的关联关系(即上述的逻辑且操作),用该两过滤条件对数据进行过滤即可以称之为同一维度且过滤,该过滤条件可以为上述的维度链,维度即可以为上述的剖析维度。

[0034] 上述实施例还可以用在电商领域,由于电商和会话是多对一的关系,即在一个会话中可能买多个订单,每个订单包含多种产品,所以会出来同一产品维度多次进行的且过滤的维度链。

[0035] 例如,选出既买了名称为'c'的产品又买了名称为'f'的产品的会话,则可以将各个会话数据表中的会话Session1和会话Session2中的各个订单进行拆分(即子表的拆分),拆分成一个维度的且过滤的形式(即拆分为分裂表),以找出符合既买了名称为'c'的产品又买了名称为'f'的产品的条件的会话,若会话Session1中包括订单A和订单B,会话Session2中包括订单C、订单D以及订单E,可以将会话Session1和会话Session2中各个订单进行拆分,得到一个维度的且过滤条件的数据。

[0036] 可选地,将对应于同一子表的各个分裂表进行表连接包括:按照内连接的方式连接对应于同一子表的各个分裂表。

[0037] 通过上述实施例,可以通过内连接的方式实现子表的表连接操作,使得经过表连接后的结果集可以表示且的关系,从而实现获取准确连接结果的效果。

[0038] 在一个可选的实施例中,可以通过处理器运行下面的程序来运行实现上述的方法。

[0039] with s as(select server session id,osbrief,pv count,session duration
from session),

[0040] product_a as(select distinct server session id from product where
product name like '%a%'),

[0041] product_b as(select distinct server session id from product where
product name like '%b%'),

[0042] product_final as(select product_a.server session id from product_a
inner join product_b on product_a.server session id=product_b.server session
id),

[0043] final as(select s.server session id,os brief,pv count,session

```
duration from s inner join product_final on s.server session id=product_
final.server session id),
```

```
[0044] select osbrief,sum(pv count)as session pageviews,count(*)as visits,
avg(session duration)as avg session duration
```

```
[0045] from final
```

```
[0046] group by os brief
```

```
[0047] order by count(*)desc
```

```
[0048] limit 10
```

[0049] 上述程序主要的过程为:查找会话中的产品名称为‘a’的产品,生成子表,查找会话中的产品名称为‘b’的产品,生成子表,然后将上述的两个子表进行内连接,来表达会话中既有产品名称为‘a’的产品,并且有产品名称为‘b’的产品。

[0050] 下面以会话Session1为例对上述程序进行同理介绍。

[0051] 会话Session1中包括订单A和订单B,其中,订单A中包括名称为‘a’、‘c’、‘d’的产品,可以将该订单数据表Session1中的订单从产品这一节点进行拆分,将其拆分成仅含有一个过滤条件的数据,其拆分后的形式可以为:

[0052] 订单会话Session1--订单A--名称为‘a’的产品;

[0053] 订单会话Session1--订单A--名称为‘b’的产品;

[0054] 订单会话Session1--订单A--名称为‘c’的产品。

[0055] 同样,会话Session1中还包括订单B,其中,订单B中包括名称为‘b’、‘d’、‘f’的产品,则可以将该订单数据表Session1中的订单拆分成仅含有一个且过滤条件的分裂表,其拆分后的形式可以为:

[0056] 会话Session1--订单B--名称为‘b’的产品;

[0057] 会话Session1--订单B--名称为‘d’的产品;

[0058] 会话Session1--订单B--名称为‘f’的产品。

[0059] 然后,可以根据预设条件(既买了名称为‘c’的产品又买了名称为‘f’的产品),将拆分后的各个分裂表进行表连接,如将“订单会话Session1--订单A--名称为‘c’的产品”和“订单会话Session1--订单B--名称为‘f’的产品”进行表连接,从而得到符合预设条件的会话Session1,从而得到同一产品维度的两次且过滤的结果。

[0060] 可选地,基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果的操作包括:获取数据表中无需拆分的第二子表;利用第二子表的过滤条件,对第二子表进行过滤,得到第二过滤结果;获取第二子表的第二过滤结果和第一子表的连接结果的汇总结果,将汇总结果作为第一过滤结果。

[0061] 在上述实施例中,通过对子表中的无需拆分的第二子表进行过滤,然后将过滤的结果和第一子表的过滤结果汇总,从而实现对子表的高效准确的过滤的效果。

[0062] 可选地,在将获取的过滤条件添加至数据表中的多个子表之前,该方法还包括:根据过滤需求从数据库中获取过滤条件;或者,获取预先设置的过滤条件。

[0063] 在上述实施例中,可以通过多种方式来获取所有过滤条件,从而实现为后续的准确过滤提供保障的效果。

[0064] 可选地,查找出多个子表中需拆分的第一子表包括:遍历数据表的多个子表的过

滤条件;判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操作的过滤条件;若当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将当前遍历到的子表确定为第一子表。

[0065] 在上述实施例中,每个子表可以包括多个维度的过滤条件,若其中存在有一个维度的过滤条件中包括两个同一维度的且过滤条件,则该子表被选定为需要拆分的子表,从而实现了准确高效选择需要拆分的子表的效果。

[0066] 可选地,将需拆分的子表拆分为多个分裂表包括:遍历各个子表;将每个子表的多个维度进行排序,从每个节点的第一个维度开始执行以下步骤,直至每个子表均拆分为多个分裂表为止,其中,将第一个维度初始化为当前维度;判断当前维度的过滤条件中是否包含至少两个同一当前维度的且过滤条件;若判断出当前维度的过滤条件中包含至少两个同一当前维度的且过滤条件,则对当前维度的过滤条件中包含至少两个同一当前维度的且过滤条件进行拆分,拆分为多个包括一个当前维度的且过滤条件的分裂表;基于排序获取当前维度对应的下一维度,将子表的下一维度作为新的当前维度。

[0067] 在将包括同一维度的多个且过滤条件的子表拆分为包括一个且过滤条件的分裂表时,可以将需要拆分的子表中的过滤条件的维度进行排序,依次判断各个维度的过滤条件是否包括至少两个同一维度的且过滤条件,若判断结果为是,则将该子表进行拆分,拆分成多个包括一个“且”过条件的分裂表,然后再将拆分后的分裂表进行下个维度的判断,若还需要拆分,则进一步进行拆分,直至拆分后的分裂表中每个维度的且过滤条件都为一个,则完成拆分程序。通过上述实施例,可以通过依次对各个维度进行判断和节点基的拆分,实现提高拆分效率的效果。

[0068] 在一个可选的实施例中,以叶子表这一数据表为例,可以通过如下步骤实施数据的同一维度多次且过滤,其中,叶子表相当于上述实施例中的子表。

[0069] 步骤S201,将所有过滤条件应用到叶子表。

[0070] 步骤S202,判断每个叶子表结点的所有过滤条件,是否包含同一维度的且过滤。

[0071] 步骤S203,对于包含同一维度的且过滤条件,进行解析树的分裂。

[0072] 具体地,其思想为:分裂意味着原先仅有一个叶子节点(即子表,包含两个相同的维度且过滤条件),分裂为两个叶子节点(即分裂表,每个分裂表包含一个维度且过滤条件),若多次出现,则进行多次分裂,最终每个子表对于多次出现的维度,仅出现一次,再将分裂后的节点采用内连接Inner Join的方式连接,表达且的关系,即既满足分裂子表1的条件,又满足分裂子表2的条件。其具体实现步骤可以包括:先将该事实表上其它所有维度过滤条件应用上,写成Sql里的一个子表表达式;再将包含多次同一维度且过滤的子表进行分裂,分裂所基于的表为上面应用所有过滤后的子表(即数据表),将分裂后的子表,进行表连接,组合成新的节点,连接方式为内连接Inner Join(意即在两表里均出现,也即且关系)。

[0073] 通过上述实施例,可以通过解析树节点分裂,再内连接Inner Join的方式表示同一维度的多次过滤;可以通过先定义所有过滤条件的子表,再对其进行分裂,避免所有分裂表都进行一次全表I/O的操作,只需要对过滤后的子表(小表)进行相关过滤即可,提升性能。

[0074] 实施例2

[0075] 根据本申请实施例的另一方面,还提供了一种数据表的过滤装置,图2是根据本申请实施例的一种数据表的过滤装置的示意图,该装置包括:添加单元10、查找单元20、拆分

单元30、表连接单元40以及获取单元50。

[0076] 其中,该装置包括:添加单元10,用于将获取的过滤条件添加至数据表中的多个子表。

[0077] 查找单元20,用于查找出多个子表中需拆分的第一子表,其中,第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一维度的多个指标之间具有逻辑且关系。

[0078] 拆分单元30,用于将第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件。

[0079] 表连接单元40,用于基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果。

[0080] 获取单元50,用于基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果。

[0081] 采用本申请实施例,通过采用将获取的过滤条件添加至数据表中的多个子表;查找出多个子表中需拆分的第一子表;将第一子表拆分为多个分裂表;基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果,基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果的方式,相比于现有技术中的直接对数据表进行具有多个指标的维度进行逻辑且操作的过滤(相当于同一维度的多个且过滤),本申请通过将数据表中的第一子表拆分为多个具有一个指标的维度进行逻辑且操作的过滤条件(相当于一个维度的且过滤)的分裂表,然后将各个分裂表进行表连接,达到了减小且过滤的数据处理量的目的,从而实现了提高数据表进行同一维度的多次且过滤时的处理效率的技术效果,进而解决了对数据表进行同一维度的多次且过滤时的处理效率低的技术问题。

[0082] 可选地,表连接单元包括:连接模块,用于按照内连接的方式连接对应于同一第一子表的各个分裂表。

[0083] 通过上述实施例,可以通过内连接的方式实现子表的表连接操作,使得经过表连接后的结果集可以表示且的关系,从而实现获取准确连接结果的效果。

[0084] 可选地,获取单元包括:第一获取模块,用于获取数据表中无需拆分的第二子表;过滤模块,用于利用第二子表的过滤条件,对第二子表进行过滤,得到第二过滤结果;处理模块,用于获取第二子表的第二过滤结果和第一子表的连接结果的汇总结果,将汇总结果作为第一过滤结果。

[0085] 在上述实施例中,通过对子表中的无需拆分的第二子表进行过滤,然后将过滤的结果和第一子表的过滤结果汇总,从而实现对子表的高效准确的过滤的效果。

[0086] 可选地,装置还包括:第二获取模块,用于在将获取的过滤条件添加至数据表中的多个子表之前,根据过滤需求从数据库中获取过滤条件;或者,第三获取模块,用于获取预先设置的过滤条件。

[0087] 在上述实施例中,可以通过多种方式来获取所有过滤条件,从而实现为后续的正确过滤提供保障的效果。

[0088] 可选地,查找单元包括:遍历模块,用于遍历数据表的多个子表的过滤条件;判断模块,用于判断当前遍历到的子表的过滤条件是否为对具有多个指标的维度进行逻辑且操

作的过滤条件;确定模块,用于若当前遍历到的子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,则将当前遍历到的子表确定为第一子表。

[0089] 在上述实施例中,每个子表可以包括多个维度的过滤条件,若其中存在有一个维度的过滤条件中包括两个同一维度的且过滤条件,则该子表被选定为需要拆分的子表,从而实现了准确高效选择需要拆分的子表的效果。

[0090] 可选地,该装置可以包括拆分单元,该拆分单元又可以包括:第二遍历模块,用于遍历各个子表;执行模块,用于将每个子表的多个维度进行排序,从每个节点的第一个维度开始执行以下步骤,直至每个子表均拆分为多个分裂表为止,其中,将第一个维度初始化为当前维度;第二判断模块,用于判断当前维度的过滤条件中是否包含至少两个同一当前维度的且过滤条件;拆分模块,用于若判断出当前维度的过滤条件中包含至少两个同一当前维度的且过滤条件,则对当前维度的过滤条件中包含至少两个同一当前维度的且过滤条件进行拆分,拆分为多个包括一个当前维度的且过滤条件的分裂表;第三获取模块,用于基于排序获取当前维度对应的下一维度,将子表的下一维度作为新的当前维度。

[0091] 在拆分包括同一维度的多个且过滤条件的子表为一个且过滤条件的分裂表时,可以将需要拆分的子表中的过滤条件的维度进行排序,依次判断各个维度的过滤条件是否包含至少两个同一维度的且过滤条件,若判断结果为是,则将该子表进行拆分,拆分成多个包括一个且过条件的分裂表,然后再将拆分后的分裂表进行下个维度的判断,若还需要拆分,则进一步进行拆分,直至拆分后的分裂表中每个维度的且过滤条件都为一个,则完成拆分程序。通过上述实施例,可以通过依次对各个维度进行判断和节点基的拆分,实现提高拆分效率的效果。

[0092] 上述本申请实施例序号仅仅为了描述,不代表实施例的优劣。

[0093] 在本申请的上述实施例中,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。

[0094] 在本申请所提供的几个实施例中,应该理解到,所揭露的技术内容,可通过其它的方式实现。其中,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,可以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。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,单元或模块的间接耦合或通信连接,可以是电性或其它的形式。

[0095] 所述数据表的过滤装置包括处理器和存储器,上述添加单元10、查找单元20、拆分单元30、表连接单元40以及获取单元50等均作为程序单元存储在存储器中,由处理器执行存储在存储器中的上述程序单元来实现相应的功能。

[0096] 处理器中包含内核,由内核去存储器中调取相应的程序单元。内核可以设置一个或以上,通过调整内核参数来减小且过滤的数据处理量,从而实现了提高数据表进行同一维度的多次且过滤时的处理效率的技术效果,进而解决了对数据表进行同一维度的多次且过滤时的处理效率低的技术问题。

[0097] 存储器可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM),存储器包括至少一个存储芯片。

[0098] 本申请还提供了一种计算机程序产品,当在数据处理设备上执行时,适于执行初始化有如下方法步骤的程序代码:将获取的过滤条件添加至数据表中的多个子表;查找出多个子表中需拆分的第一子表,其中,第一子表的过滤条件为对具有多个指标的维度进行逻辑且操作的过滤条件,同一维度的多个指标之间具有逻辑且关系;将第一子表拆分为多个分裂表,其中,每个分裂表的过滤条件为对多个具有一个指标的维度进行逻辑且操作的过滤条件;基于各个分裂表的过滤条件,将对应于同一第一子表的各个分裂表进行表连接,得到各个第一子表的连接结果;基于各个第一子表的连接结果,获取数据表对应于获取的过滤条件的第一过滤结果。

[0099] 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。

[0100] 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。

[0101] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可为个人计算机、服务器或者网络设备等)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0102] 以上所述仅是本申请的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本申请的保护范围。

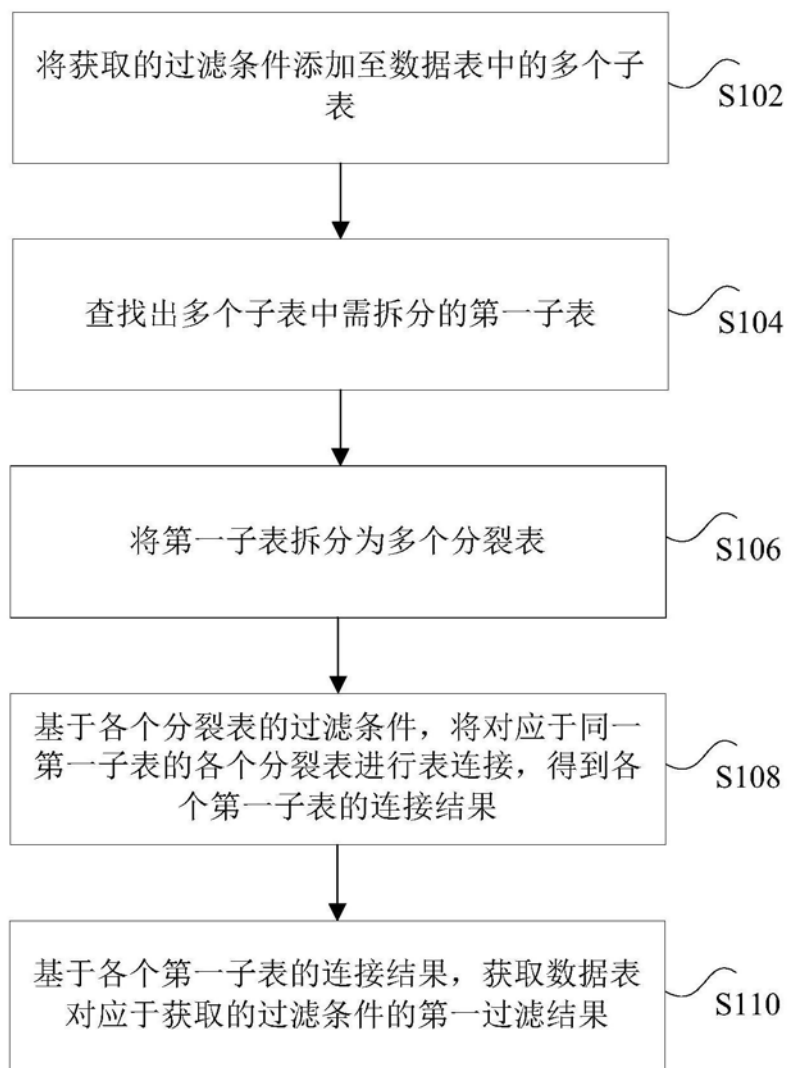


图1



图2