



(12) 发明专利申请

(10) 申请公布号 CN 104657370 A

(43) 申请公布日 2015. 05. 27

(21) 申请号 201310587065. 6

(22) 申请日 2013. 11. 19

(71) 申请人 中国移动通信集团天津有限公司
地址 300300 天津市东丽区空港经济加工区
中央大道 99 号

(72) 发明人 刘玉娜

(74) 专利代理机构 北京派特恩知识产权代理有
限公司 11270
代理人 张颖玲 张振伟

(51) Int. Cl.
G06F 17/30(2006. 01)

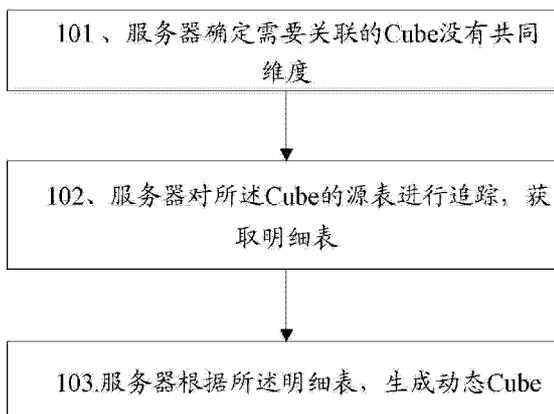
权利要求书1页 说明书7页 附图2页

(54) 发明名称

一种实现多维立方体关联的方法和装置

(57) 摘要

本发明公开了一种实现多维立方体(Cube)关联的方法,确定需要关联的 Cube 没有共同维度;对 Cube 的源表进行追踪,获取明细表;根据获得的明细表,生成动态 Cube。本发明同时还公开了一种实现 Cube 关联的装置。通过本发明的方案,能够实现无公共维度的 Cube 关联,为用户提供更加便捷的数据分析。



1. 一种实现多维立方体 Cube 关联的方法,其特征在于,所述方法包括:
服务器确定需要关联的 Cube 没有共同维度,对所述 Cube 的源表进行追踪,获取明细表,根据所述明细表生成动态 Cube。
2. 根据权利要求 1 所述的方法,其特征在于,所述服务器对所述 Cube 的源表进行追踪,获取明细表,包括:
所述服务器根据需要关联的 Cube 的元数据,查询需要关联的 Cube 的生成过程,获取生成需要关联的 Cube 的源表,根据所述源表得到明细表。
3. 根据权利要求 2 所述的方法,其特征在于,所述根据所述明细表,生成动态 Cube,包括:
服务器对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态 Cube ;其中,所述参数信息包括主键、维度和度量信息。
4. 根据权利要求 3 所述的方法,其特征在于,所述对生成需要关联的 Cube 的明细表进行关联包括:
服务器通过相同的主键信息,对生成需要关联的 Cube 的明细表进行关联。
5. 根据权利要求 3 或 4 所述的方法,其特征在于,所述根据所述虚拟明细表生成动态 Cube 包括 :通过任务 Task 将虚拟明细表生成事实表,再使用结构化查询语言 SQL 语句将所述事实表的维度信息和度量信息结合生成动态 Cube 。
6. 一种实现多维立方体 Cube 关联的装置,其特征在于,所述装置包括 :判断模块、源表分析模块和生成模块 ;其中,
判断模块,用于确定需要关联的 Cube 没有共同维度 ;
源表分析模块,用于对所述 Cube 的源表进行追踪,获取明细表 ;
生成模块,用于根据所述明细表,生成动态 Cube 。
7. 根据权利要求 6 所述的装置,其特征在于,所述判断模块,具体用于比较需要关联的 Cube 的维度信息,在没有相同的维度信息时,确定所述需要关联的 Cube 没有共同维度,并通知所述源表分析模块。
8. 根据权利要求 6 所述的装置,其特征在于,所述源表分析模块,具体用于根据需要关联的 Cube 的元数据,查询所述 Cube 的生成过程,获取生成所述 Cube 的源表,根据所述源表得到明细表。
9. 根据权利要求 6 所述的装置,其特征在于,所述生成模块,具体用于对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态 Cube ;其中,所述参数信息包括主键、维度和度量信息。
10. 根据权利要求 9 所述的装置,其特征在于,所述生成模块包括 :虚拟明细表生成模块以及 Cube 生成模块 ;其中,
虚拟明细表生成模块,用于将明细表进行关联,生成虚拟明细表 ;
Cube 生成模块,用于根据所述虚拟明细表,生成动态 Cube 。

一种实现多维立方体关联的方法和装置

技术领域

[0001] 本发明涉及分布式统计技术,尤其是涉及一种实现多维立方体(Cube)关联的方法和装置。

背景技术

[0002] 联机分析处理(OLAP)是由关系数据库之父 E. F. Codd 于 1993 年提出的, Codd 认为联机事务处理(OLTP)已不能满足终端用户对数据库查询分析的要求,结构化查询语言(SQL)对大数据的简单查询也不能满足用户分析的需求。用户的决策分析需要对关系数据库进行大量计算才能得到结果,而查询的结果并不能满足决策者提出的需求。因此,Codd 提出了多维数据库和多维分析的概念,即 OLAP。OLAP 委员会对联机分析处理的定义为:从原始数据中转化出来的、能够真正为用户所理解的、并真实反映企业多维特性的数据称为信息数据,使分析人员、管理人员或执行人员能够从多种角度对信息数据进行快速、一致、交互地存取,从而获得对数据的更深入了解的一类软件技术。

[0003] 现有的多维分析方案主要有以下两种:

[0004] ROLAP:基于关系数据库的 OLAP 实现(Relational OLAP);以关系数据库为核心,以关系型结构进行多维数据的表示和存储。ROLAP 将多维数据库的多维结构划分为两类表:一类是事实表,用来存储数据和维关键字;另一类是维表,即对每个维至少使用一个表来存放维的层次、成员类别等维的描述信息。维表和事实表通过主关键字和外关键字联系在一起,形成了“星型模式”。

[0005] MOLAP:基于多维数据组织的 OLAP 实现(Multidimensional OLAP);以多维数据组织方式为核心,也就是说,MOLAP 使用多维数组存储数据。多维数据在存储中将形成“Cube”的结构,在 MOLAP 中对“Cube”的“旋转”、“切块”、“切片”是产生多维数据报表的主要技术。

[0006] 随着数据量的增大,上述方案中生成的 Cube 已不能满足用户的需求,这就需要两个或者多个 Cube 进行关联,以得到用户所需要的信息。

[0007] 图 1 示出了 Cube 的生成过程:先由接口表通过任务(Task)生成汇总表,汇总表通过 Task 生成明细表,明细表再通过 Task 生成事实表,最后事实表通过 Task 生成 Cube;当需要对两个或多个具有共同维度的 Cube 进行关联时,会形成大数据运算,例如将两个 Cube 的数据汇总到一个 Cube 中,会通过两个 Cube 的共同维度进行关联,然后通过关联后的数据,计算各个 Cube 的明细表。

[0008] 因此,现有的两个方案中,在对两个 Cube 实现关联分析时都存在有共同的缺点:实现两个 Cube 的关联分析,必须要求这两个 Cube 具有公共维度,无公共维度的 Cube 无法实现关联分析。

发明内容

[0009] 本发明实施例主要提供一种实现 Cube 关联的方法和装置,能够实现无公共维度的 Cube 关联,提高多 Cube 关联分析的效率。

- [0010] 本发明的技术方案是这样实现的：
- [0011] 本发明提供了一种实现 Cube 关联的方法，所述方法包括：
- [0012] 服务器确定需要关联的 Cube 没有共同维度，对所述 Cube 的源表进行追踪，获取明细表，根据所述明细表生成动态 Cube。
- [0013] 上述方案中，所述服务器对所述 Cube 的源表进行追踪，获取明细表，包括：
- [0014] 所述服务器根据需要关联的 Cube 的元数据，查询需要关联的 Cube 的生成过程，获取生成需要关联的 Cube 的源表，根据所述源表得到明细表。
- [0015] 上述方案中，所述根据所述明细表，生成动态 Cube，包括：
- [0016] 服务器对所述明细表进行解析，记录组成所述明细表的参数信息，通过所述参数信息，对生成需要关联的 Cube 的明细表进行关联，得到虚拟明细表，根据所述虚拟明细表生成动态 Cube；其中，所述参数信息包括主键、维度和度量信息。
- [0017] 上述方案中，所述对生成需要关联的 Cube 的明细表进行关联包括：
- [0018] 服务器通过相同的主键信息，对生成需要关联的 Cube 的明细表进行关联。
- [0019] 上述方案中，所述根据所述虚拟明细表生成动态 Cube 包括：通过任务 Task 将虚拟明细表生成事实表，再使用结构化查询语言 SQL 语句将所述事实表的维度信息和度量信息结合生成动态 Cube。
- [0020] 本发明提供了一种实现 Cube 关联的装置，所述装置包括：判断模块、源表分析模块和生成模块；其中，
- [0021] 判断模块，用于确定需要关联的 Cube 没有共同维度；
- [0022] 源表分析模块，用于对所述 Cube 的源表进行追踪，获取明细表；
- [0023] 生成模块，用于根据所述明细表，生成动态 Cube。
- [0024] 上述方案中，所述判断模块，具体用于比较需要关联的 Cube 的维度信息，在没有相同的维度信息时，确定所述需要关联的 Cube 没有共同维度，并通知所述源表分析模块。
- [0025] 上述方案中，所述源表分析模块，具体用于根据需要关联的 Cube 的元数据，查询所述 Cube 的生成过程，获取生成所述 Cube 的源表，根据所述源表得到明细表。
- [0026] 上述方案中，所述生成模块，具体用于对所述明细表进行解析，记录组成所述明细表的参数信息，通过所述参数信息，对生成需要关联的 Cube 的明细表进行关联，得到虚拟明细表，根据所述虚拟明细表生成动态 Cube；其中，所述参数信息包括主键、维度和度量信息。
- [0027] 上述方案中，所述生成模块包括：虚拟明细表生成模块以及 Cube 生成模块；其中，
- [0028] 虚拟明细表生成模块，用于将明细表进行关联，生成虚拟明细表；
- [0029] Cube 生成模块，用于根据所述虚拟明细表，生成动态 Cube。
- [0030] 本发明实施例提供了一种实现 Cube 关联的方法和装置，当需要关联的 Cube 没有共同维度时，服务器通过对生成所述 Cube 的源表进行追踪，获取明细表；然后根据获取的明细表生成动态 Cube；如此，能够实现无公共维度的 Cube 关联，提高多 Cube 关联分析的效率，为用户分析数据提供便利；并且通过追踪源表数据的方法，从而减小 Cube 关联时的系统性能消耗和执行时间。

附图说明

- [0031] 图 1 为现有技术中生成 Cube 的方法的流程示意图；
 [0032] 图 2 为本发明实施例一提供的实现 Cube 关联的方法流程示意图；
 [0033] 图 3 为本发明实施例二提供的实现 Cube 关联的装置的结构示意图。

具体实施方式

[0034] 本发明实施例中,当需要关联的 Cube 没有共同维度时,服务器对所述需要关联的 Cube 的源表进行追踪,获取明细表,然后对明细表进行解析,获得相同的参数信息,并通过相同的参数信息,生成动态 Cube,从而实现无公共维度的 Cube 关联,解决了没有共同维度的 Cube 不能关联的问题。

[0035] 其中,所述需要关联的 Cube,在生成过程中,具有共同的参数信息;所述参数信息包括主键信息。

[0036] 下面通过附图及具体实施例对本发明做进一步的详细说明。

[0037] 实施例一

[0038] 本发明实施例一实现 Cube 关联的方法,如图 2 所示,该方法包括以下几个步骤:

[0039] 步骤 101:服务器确定需要关联的 Cube 没有共同维度。

[0040] 具体的,服务器比较需要关联的 Cube 的维度信息,在没有相同的维度信息时,确定所述需要关联的 Cube 没有共同维度;所述服务器可以是用于数据管理的服务器,如移动公司后台的数据管理服务器。

[0041] 在具有相同的维度信息时,所述服务器直接通过所述维度信息对所述需要关联的 Cube 进行关联。

[0042] 例如,对于表 1 中的 Cube1 和表 2 中的 Cube2,当用户需要对各个地区的各个品牌的用户数量和收入进行统计查看时,就需要对 Cube1 和 Cube2 进行关联,生成一个具有品牌信息、收入信息、在网客户数信息和收入信息的新的 Cube;通过查询表 3 所示的 Cube1 的事实表 1 和表 4 所示的 Cube2 的事实表 2,比较 Cube1 和 Cube2 是否具有共同的维度信息,由于 Cube1 的事实表 1 的维度信息为品牌信息,Cube2 的维度信息为地区信息,可知 Cube1 和 Cube2 没有共同维度,所以通过 Cube1 的维度信息或 Cube2 的维度信息无法实现 Cube1 和 Cube2 的直接关联。

[0043]

Cube1	
品牌	总收入
全球通	27490.672
动感地带	13878.082
神州行	23091.921

[0044] 表 1

[0045]

Cube2	
地区	在网客户数
和平河西	327490
南开红桥	113878
河北河东	423091
。 。 。 。	。 。 。 。

[0046] 表 2

[0047]

事实表 1		
属性	类型	说明
品牌	varchar2 (10)	维度
总收入	number	度量

[0048] 表 3

[0049]

事实表 2		
属性	类型	说明
地区	varchar2 (10)	维度
在网客户数	number	度量

[0050] 表 4

[0051] 步骤 102 :服务器对所述 Cube 的源表进行追踪,获取明细表。

[0052] 具体的,所述服务器根据所述 Cube 的元数据,查询所述 Cube 的生成过程,获取生成所述 Cube 的源表,根据所述源表得到明细表;其中,所述源表包括明细表和事实表。

[0053] 所述元数据为记录 Cube 的生成过程以及底层数据逻辑关系的数据,通过所述元数据,可以追踪到生成 Cube 的源表,所述源表包括明细表和事实表;图 1 中示出了所述元数据记录的 Cube 的生成过程:所述 Cube 是由明细表通过 Task 生成事实表,再由所述事实表通过 Task 生成的;因此,通过所述元数据记录的 Cube 的生成过程,可以追踪到生成所述 Cube 的源表,从而获得明细表。

[0054] 对于步骤 101 中的 Cube1 和 Cube2,本步骤可以通过元数据,由 Cube1 的事实表 1 追踪到表 5 所示的 Cube1 的明细表 1,由 Cube2 的事实表 2 追踪到表 6 所示的 Cube2 的明细表 2。

明细表 1		
属性	类型	说明
[0055] 用户标识	char (20)	主键
品牌	varchar2 (10)	维度
总收入	number	度量

[0056] 表 5

明细表 2		
属性	类型	说明
[0057] 用户标识	char (20)	主键
地区	varchar2 (10)	维度
在网客户数	number	度量

[0058] 表 6

[0059] 步骤 103 :服务器根据所述明细表,生成动态 Cube。

[0060] 具体的,服务器对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态 Cube ;其中,所述参数信息包括主键、维度和度量信息。

[0061] 通过所述参数信息,对生成需要关联的 Cube 的明细表进行关联,一般可以是服务器通过相同的主键信息,对生成需要关联的 Cube 的明细表进行关联 ;

[0062] 服务器通过 Task 将虚拟明细表生成事实表,使用 SQL 语句将所述事实表的维度信息和度量信息结合生成动态 Cube。

[0063] 例如,服务器可以对步骤 102 中的 Cube1 的明细表 1 和 Cube2 的明细表 2 进行解析,记录明细表 1 和明细表 2 的参数信息,对于 Cube1 :记录主键信息为用户标识信息,记录维度信息为品牌信息,记录度量信息为总收入信息 ;对于 Cube2 :记录主键信息为用户标识信息,记录维度信息为地区信息,记录度量信息为在网客户数信息 ;通过比较明细表 1 和明细表 2 得知,明细表 1 和明细表 2 具有相同的主键信息(即用户标识),通过相同的主键信息,对明细表 1 和明细表 2 进行关联,生成表 7 所示的虚拟明细表 1,将虚拟明细表 1 结合记录的维度信息(品牌和地区)和度量信息(总收入和在网客户数),生成表 8 所示的动态 Cube3,表 8 中包含了各个地区的各个品牌的用户数量和收入,从而满足了用户的需求。

[0064]

虚拟明细表 1		
属性	类型	说明
用户标识	char (20)	主键
品牌	varchar2(10)	维度
地区	varchar2(10)	维度
总收入	number	度量
在网客户数	number	度量

[0065] 表 7

[0066]

动态 Cube3			
品牌	地区	总收入	在网客户数
全球通	和平河西	49.632	227
	南开红桥	21.56	878
	河北河东	48.091	391
	。。。。	17.378	878
动感地带	和平河西	82.091	291
	南开红桥	71.878	178
	河北河东	30.91	482
	。。。。	23.67	765
神州行	和平河西	23.9	123
	南开红桥	20.1	156
	河北河东	35.21	231
	。。。。	83.25	67

[0067]

[0068] 表 8

[0069] 实施例二

[0070] 本发明实施例二提供了一种实现 Cube 关联的装置,如图 3 所示,所述装置包括:判断模块 301、源表分析模块 302 以及生成模块 303;其中,

[0071] 判断模块 301,用于确定需要关联的 Cube 没有共同维度;

[0072] 源表分析模块 302,用于对所述 Cube 的源表进行追踪,获取明细表;

[0073] 生成模块 303,用于根据所述明细表,生成动态 Cube;

[0074] 其中,所述需要关联的 Cube 为:在生成过程中,具有相同的参数信息;所述参数信息包括主键、维度和度量信息;

[0075] 具体的,所述判断模块 301 比较需要关联的 Cube 的维度信息,在没有相同的维度信息时,确定所述需要关联的 Cube 没有共同维度,并通知所述源表分析模块 302;

[0076] 其中,在具有相同的维度信息时,直接通过所述维度信息对所述需要关联的 Cube 进行关联;

[0077] 所述源表分析模块 302 根据需要关联的 Cube 的元数据,查询所述 Cube 的生成过程,获取生成所述 Cube 的源表,根据所述源表得到明细表;其中,所述源表包括明细表和事

实表；

[0078] 其中,所述元数据为记录 Cube 的生成过程以及底层数据逻辑关系的数据,通过所述元数据,可以追踪到生成 Cube 的明细表。

[0079] 所述生成模块 303,具体用于对所述明细表进行解析,记录组成所述明细表的参数信息,通过所述参数信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表,根据所述虚拟明细表生成动态 Cube ;其中,所述参数信息包括主键、维度和度量信息；

[0080] 具体的,所述生成模块 303 通过相同的主键信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表,通过 Task 将虚拟明细表生成事实表,使用 SQL 语句将所述事实表的维度信息和度量信息结合生成动态 Cube。

[0081] 其中,所述判断模块 301 可以由服务器的 CPU 进行实现,或者由对比电路进行实现；

[0082] 所述源表分析模块 302 可以由服务器的 CPU 和内存或硬盘进行实现；

[0083] 所述生成模块 303 可以由服务器的 CPU 和显示器进行实现。

[0084] 所述生成模块 303 包括 :虚拟明细表生成模块 401 以及 Cube 生成模块 402 ;其中,

[0085] 虚拟明细表生成模块 401,用于将明细表进行关联,生成虚拟明细表；

[0086] Cube 生成模块 402,用于根据所述虚拟明细表,生成动态 Cube ；

[0087] 具体的,所述虚拟明细表生成模块 401 通过相同的主键信息,对生成需要关联的 Cube 的明细表进行关联,得到虚拟明细表 ;所述 Cube 生成模块 402 通过 Task 将虚拟明细表生成事实表,使用 SQL 语句将所述事实表的维度信息和度量信息结合生成动态 Cube。

[0088] 通过本发明所提供的方法和装置,能够对无公共维度的 Cube 进行关联 ;通过追踪到生成所述 Cube 的明细表,查询相同的主键信息,通过所述主键进行关联合并生成虚拟明细表,再将所述虚拟明细表生成动态 Cube,实现对无公共维度的 Cube 的关联,提高系统的运行效率,为用户提供所需要关联的相关信息。

[0089] 以上所述,仅为本发明的较佳实施例而已,并非用于限定本发明的保护范围,凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本发明的保护范围之内。

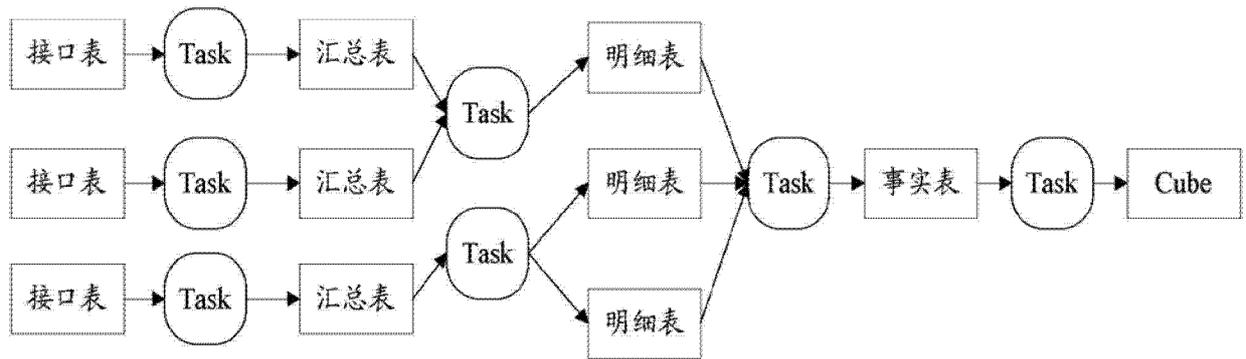


图 1

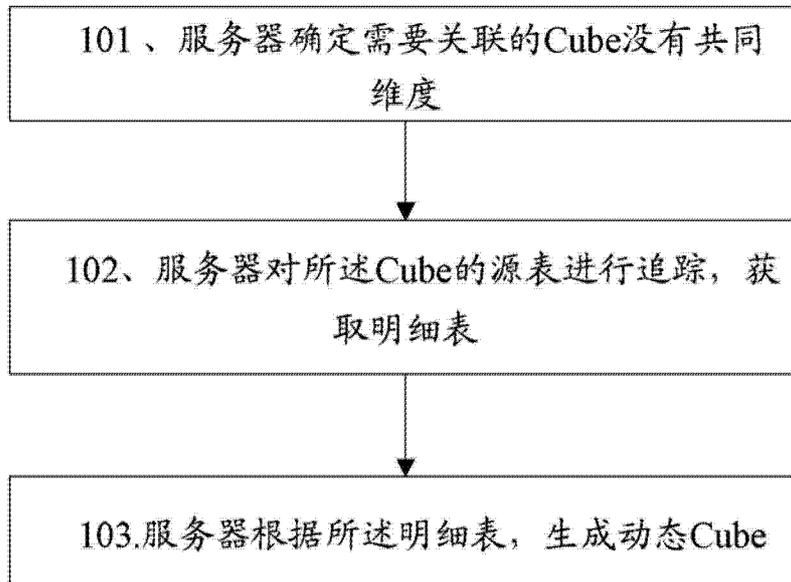


图 2

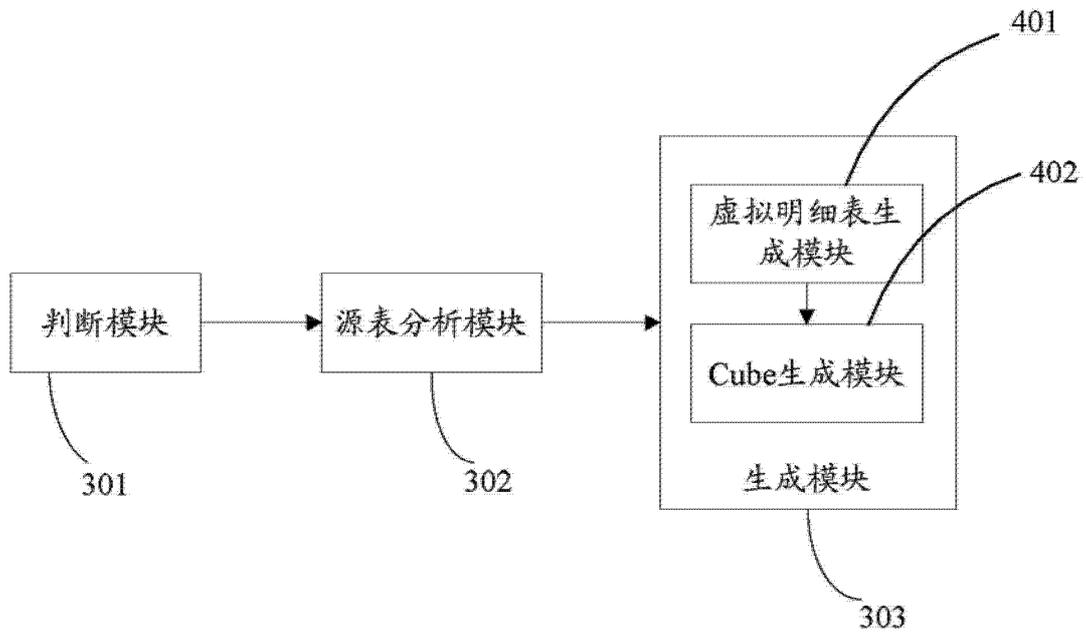


图 3