



(12) 发明专利申请

(10) 申请公布号 CN 102648466 A

(43) 申请公布日 2012. 08. 22

(21) 申请号 201080053432. 9

(74) 专利代理机构 北京市中咨律师事务所

(22) 申请日 2010. 11. 09

11247

代理人 于静 张亚非

(30) 优先权数据

09177866. 2 2009. 12. 03 EP

(51) Int. Cl.

G06F 17/30(2006. 01)

(85) PCT申请进入国家阶段日

2012. 05. 25

(86) PCT申请的申请数据

PCT/EP2010/067127 2010. 11. 09

(87) PCT申请的公布数据

W02011/067080 EN 2011. 06. 09

(71) 申请人 国际商业机器公司

地址 美国纽约

(72) 发明人 C·C·麦西 M·瓦兰德

J·P·伯兰德

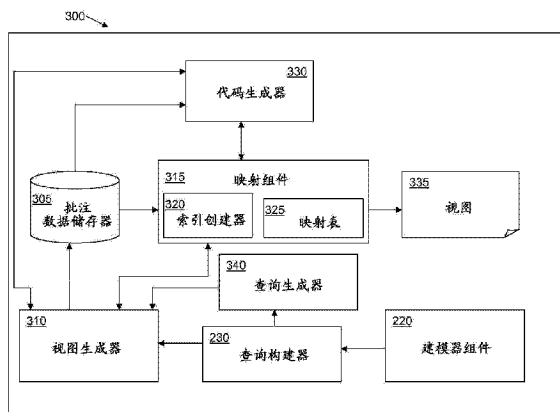
权利要求书 1 页 说明书 13 页 附图 7 页

(54) 发明名称

用于取回视图中的数据项批注的方法

(57) 摘要

一种取回与视图中的数据项关联的批注的方法,其中所述视图由信息管理系统查询数据源而生成,所述方法包括以下步骤:接收查询的输出;分析所述查询的输出以识别一个或多个数据项,其中数据项包括数据值和与该数据值关联的属性;对于每个所识别的数据值和属性,识别与所述数据值和所述属性中的每一个关联的唯一值,其中与所述数据值关联的所识别的唯一值和与所述属性关联的所识别的唯一值形成唯一值集合;从数据储存器识别与所述唯一值集合对应的先前记录的唯一值集合;响应于肯定的判定,判定所述先前记录的唯一值集合是否包括关联批注;以及响应于肯定的第二判定,从所述数据储存器取回所述批注。



1. 一种取回与视图中的数据项关联的批注的方法,其中所述视图由信息管理系统查询数据源而生成,所述方法包括以下步骤:

接收查询的输出;

分析所述查询的输出以识别一个或多个数据项,其中数据项包括数据值和与该数据值关联的属性;

对于每个所识别的数据值和属性,识别与所述数据值和所述属性中的每一个关联的唯一值,其中与所述数据值关联的所识别的唯一值和与所述属性关联的所识别的唯一值形成唯一值集合;

从数据储存器识别与所述唯一值集合对应的先前记录的唯一值集合;

响应于肯定的判定,判定所述先前记录的唯一值集合是否包括关联批注;以及

响应于肯定的第二判定,从所述数据储存器取回所述批注。

2. 如权利要求 1 的方法,还包括在视图中显示所识别的数据项中的每一个和所述关联批注。

3. 如权利要求 1 的方法,其中从所述数据储存器取回所述批注进一步包括:将所述唯一值集合内的所述唯一值中的每一个映射回如在所述查询的输出中识别的所述唯一值的关联数据项;以及在视图中显示所述数据项与所述关联批注。

4. 如权利要求 1 的方法,其中数据项包括数据和所述数据的已通过查询数据源而取回的关联属性。

5. 如权利要求 2 的方法,其中数据的属性包括如在视图中显示的列和 / 或行标题。

6. 如权利要求 1 的方法,其中批注包括字符串、整数值、URL、至信息源的其它指针或链接。

7. 如权利要求 1 的方法,其中批注可与视图中的任何位置处的数据项关联。

8. 如权利要求 1 的方法,其中批注与一个或多个视图中的多个数据项关联。

9. 如权利要求 1 的方法,其中所述数据源是在所述信息管理系统外部的数据源。

10. 一种取回与视图中的数据项关联的批注的装置,其中所述视图由信息管理系统查询数据源而生成,所述装置包括以下组件:

用于接收查询的输出的映射组件;

用于分析所述查询的输出以识别一个或多个数据项的映射组件,其中数据项包括数据值和与该数据值关联的属性;

用于识别与每个数据值和所述数据值中的每个数据值的属性关联的唯一值的映射组件,其中与所述数据值关联的所识别的唯一值和与所述属性关联的所识别的唯一值形成唯一值集合;

用于从数据储存器识别与所述唯一值集合对应的先前记录的唯一值集合的映射组件;

用于响应于肯定的判定而判定所述先前记录的唯一值集合是否包括关联批注的映射组件;以及

用于响应于肯定的第二判定而从所述数据储存器取回所述批注的映射组件。

11. 一种包括计算机程序代码的计算机程序,所述计算机程序代码在加载至计算机系统中且执行时执行如权利要求 1 至 9 中的任一权利要求的方法的所有步骤。

用于取回视图中的数据项批注的方法

技术领域

[0001] 本发明涉及信息管理的领域。具体地说,本发明涉及一种用于将批注与视图中的数据项关联的改进方法。

背景技术

[0002] 信息管理系统包括工具及应用,所述工具及应用存储数据、分析数据及对数据执行某一形式的计算,以便向用户提供对数据的某种有意义的理解。

[0003] 信息管理系统可包括收集、分析及报告信息的任何数目的应用,如数据库应用和电子表格应用。通过使用关系型数据库应用的实例,使用关系模型以便定义具有与其它数据元素同样的属性的数据元素之间的关系。例如,消费者与发票可具有“一对多”关系。这意味着一个消费者可具有“许多”发票,但一张发票不具有“许多”消费者,因为发票倾向于对特定消费者而言是唯一的。一旦已定义了关系,便可接着创建采用预定义关系的查询以便提供有意义的报告。例如,可定义一个查询以查询数据源,以便查明特定消费者具有多少张未付(outstanding) 发票。

[0004] 在一种在线分析处理系统中,通常在星型或雪花型模式(schema) 上对基础数据结构建模。所述系统包括称为测量且通过称为维度的事项分类的数值型事实(numeric fact)。从事实表中的记录导出测量,且从维度表导出维度。在数据仓库中,维度是将数据集中的每一项目分类为非重叠区域的数据元素。生成视图,所述视图显示测量及与测量关联的事实,即,数据及描述数据的性质的描述符。例如,数据可采取“等离子屏幕电视”的形式,且“等离子屏幕电视”的维度是“产品”。

[0005] 另一实例可见于电子表格中,其中数据被填入“组成”列及行的单元格中。通常,使用数据进行某种计算分析,且(例如)以表格形式等来显示结果。

[0006] 然而,所有以上类型的应用的共同点在于:分析数据,且向用户显示结果以供审阅和/或进一步分析。

[0007] 通常,当生成报告或视图时,报告的查看者可能希望对报告/视图中的一个或多个项目进行批注。尽管此情形通过向数据所在的单元格添加注释而是可能的,但当重新运行查询且随后刷新报告数据时,将出现问题。这是因为注释并不跟随显示于报告中的第一位置处的数据,且在刷新报告时,数据现在显示于报告中的第二位置处。问题在于,注释仍显示于第一位置处,但数据已移动至不同位置,且因此注释现在引用不正确的数据。

[0008] 用于解决此问题的一种现有技术解决方案可见于描述一种方法的美国专利申请案第 2006/0212469 号中,所述方法用于将项目元数据与电子表格中的项目相关联,使得当项目移动至电子表格中的不同单元格时,注释亦与项目一起移动。此情形通过在电子表格自身中创建索引来实现,所述索引创建至元数据表的索引。然而,此解决方案的问题在于,虽然项目元数据将与表中的项目一起移动至新单元格,但此解决方案仅在数据包含在电子表格自身内的情况下起作用。因此,现有技术仍依赖数据的单元格位置来确定将注释移动至的位置。此解决方案在以下情形下不起作用:数据由外部数据源提供和/或生成报告的

查询被重新运行。另外,现有技术解决方案需要将元数据索引“插入”项目的单元格中,因此需要通过修改基础电子表格的结构使其具有 ID 字段。因此,本发明试图减轻上述问题。

发明内容

[0009] 从第一方面考虑,本发明提供一种取回与视图中的数据项关联的批注的方法,其中所述视图由信息管理系统查询数据源而生成,所述方法包括以下步骤:接收查询的输出;分析所述查询的输出以识别一个或多个数据项,其中数据项包括数据值和与该数据值关联的属性;对于每个所识别的数据值和属性,识别与所述数据值和所述属性中的每一个关联的唯一值,其中与所述数据值关联的所识别的唯一值和与所述属性关联的所识别的唯一值形成唯一值集合;从数据储存器识别与所述唯一值集合对应的先前记录的唯一值集合;响应于肯定的判定,判定所述先前记录的唯一值集合是否包括关联批注;以及响应于肯定的第二判定,从所述数据储存器取回所述批注。

[0010] 有利地,本发明提供将批注(多个)与视图中的数据项关联并取回所述批注(多个)的能力。批注是传达待与数据项关联的信息的任何形式的注释或元数据。数据项包括数据及数据属性。数据被表征为已从诸如数据仓库的外部数据源查询的数据或已对其执行某一形式的计算分析的数据。数据属性包括列和行标题,即,给出数据的上下文含义的属性。

[0011] 当接收到查询时,将查询基础数据源,且返回查询的结果以供进一步分析。作为元组(tuple)集合而返回结果。每一维度(即,每一单元格位置)可通过元组(即,行、列、数据)来表示。如果批注与数据项关联,则元组可包括行、列、数据及批注。针对每一数据项(在元组中识别的行、列及数据)识别唯一生成的值。如果数据集包括额外标题或分组,则将此信息包括至所述额外标题或分组。所述唯一值中的每一个唯一值的组合形成唯一值集合,其被记录在批注数据储存器中的索引元组表中。将批注存储在批注表中,且针对批注表中存储的每一批注生成唯一关键字(key)。所述唯一关键字与批注的关联唯一值集合一起存储。唯一值可在所有视图上或基于每个视图为唯一的。

[0012] 当重新运行查询且将新维度或新属性引入查询的结果时,即,如果该查询正在查询一个数据源以判定“哪些国家已销售了多于 500 台电视机”,则第一次执行该查询,其可返回这些国家:英国、法国。然而,在后续查询时,可能存在新的可用数据,且该查询的结果可将额外国家(诸如,德国及美国)返回至所述结果中。

[0013] 因此,分析新查询的结果以识别形成所生成查询的结果的数据项。视图中的每一数据项(数据值和属性,或数据值及维度)与元组相关联,所述元组具有来自每一维度的属性。

[0014] 分析所述元组中的每一数据项以识别唯一值。记录每一所识别的唯一值以形成唯一值集合,其中所述唯一值集合表示查询结果中的元组。在表中执行查找以识别是否已记录先前唯一值集合,其中所述先前唯一值集合为与当前“正被分析”的唯一值集合相同的值。如果判定为肯定的,则取回与已记录的值集合关联的唯一关键字,且该唯一关键字允许执行查找以便定位关联批注。“正被分析”的唯一值集合然后被映射回至所述数据项中的每个数据项且呈现在视图中,从而显示所述查询的结果。

[0015] 优选地,本发明提供一种方法,还包括在视图中显示所识别的数据项中的每一个

和所述关联批注。

[0016] 优选地,本发明提供一种方法,其中从所述数据储存器取回所述批注进一步包括:将所述唯一值集合内的所述唯一值中的每一个映射回如在所述查询的输出中识别的所述唯一值的关联数据项;以及在视图中显示所述数据项与所述关联批注。

[0017] 优选地,本发明提供一种方法,其中数据项包括数据和所述数据的已通过查询数据源而取回的关联属性。

[0018] 优选地,本发明提供一种方法,其中数据的属性包括如在视图中显示的列和/或行标题。

[0019] 优选地,本发明提供一种方法,其中批注包括字符串、整数值、URL、至信息源的其他指针或链接。

[0020] 优选地,本发明提供一种方法,其中批注可与视图中的任何位置处的数据项关联。

[0021] 优选地,本发明提供一种方法,其中批注与一个或多个视图中的多个数据项关联。

[0022] 优选地,本发明提供一种方法,其中所述数据源是在所述信息管理系统外部的数据源。

[0023] 从另一方面考虑,本发明提供一种将批注与视图中的数据项关联的方法,其中所述视图从在视图中呈现查询的结果而生成,所述方法包括以下步骤:接收指令以将批注与视图中的数据项关联;从所述数据项确定与所述视图中的所述数据值关联的一个或多个属性;对于每一数据值及其关联属性,确定与所述数据值及其关联属性相关联的唯一值;记录所述唯一值中的每一个以创建唯一值集合,所述唯一值集合表示所述数据项及所述数据项的属性;以及生成用于将所请求的批注与所述唯一值集合关联的唯一值。

[0024] 从另一方面考虑,本发明提供一种取回与视图中的数据项关联的批注的装置,其中所述视图由信息管理系统查询数据源而生成,所述装置包括以下组件:用于接收查询的输出的映射组件;用于分析所述查询的输出以识别一个或多个数据项的映射组件,其中数据项包括数据值和与该数据值关联的属性;用于识别与每个数据值和所述数据值中的每个数据值的属性关联的唯一值的映射组件,其中与所述数据值关联的所识别的唯一值和与所述属性关联的所识别的唯一值形成唯一值集合;用于从数据储存器识别与所述唯一值集合对应的先前记录的唯一值集合的映射组件;用于响应于肯定的判定而判定所述先前记录的唯一值集合是否包括关联批注的映射组件;以及用于响应于肯定的第二判定而从所述数据储存器取回所述批注的映射组件。

[0025] 从另一方面考虑,本发明提供一种包括计算机程序代码的计算机程序,所述计算机程序代码在加载至计算机系统中且执行时执行如上文所描述的方法的所有步骤。

附图说明

[0026] 现在将仅通过实例参考附图描述本发明的各实施例,这些附图是:

[0027] 图 1a 及图 1b 是本发明的一个优选实施例可在其中操作的数据处理系统的示意性表示;

[0028] 图 2 是详细描述如本领域公知的信息管理系统的示意图;

[0029] 图 3 是根据本发明的一个优选实施例的详细描述信息管理系统上的组件的示意图;

[0030] 图 4 是根据本发明的一个优选实施例的展示信息管理系统生成的具有关联批注的报告的一个实例的示意图；

[0031] 图 5 是根据本发明的一个优选实施例的展示索引创建器组件的各组件的示意图；

[0032] 图 6a 及图 6b 是根据本发明的一个优选实施例的展示实例视图及表示视图中的维度的元组的示意图；及

[0033] 图 7 是根据本发明的一个优选实施例的详细描述数据库应用的处理步骤的流程图。

具体实施方式

[0034] 图 1a 及图 1b 详细描述与本发明的一个优选实施例一起使用的数据处理系统 100。数据处理系统 100 包括彼此协作以便提供本发明的益处的硬件 115 及软件 120。通常, 数据处理系统 100 包括: 某一形式的存储构件 120, 在存储构件 120 中本地地在数据处理系统上或经由数据处理系统 100 外部的存储构件 145 来存储数据; 存储及内存构件 115, 其用于存储并运行可操作以在所述数据处理系统上使用的信息管理应用; 输入构件 125, 其用于输入指令及与信息管理应用的操作关联的数据; 及显示构件 130, 其用于查看数据库应用的输出。

[0035] 信息管理应用可以服务器模式 135 或客户端模式 150 操作。当以服务器模式 135 操作时, 客户端设备 150 可操作以经由网络 140 连接至服务器。客户端设备 150 可经由任何形式的有线网络构件 140 或无线网络构件 140 连接至服务器 135。

[0036] 客户端设备 150 包括输入构件 155 及输出构件 160, 及用于与数据处理系统 100 交互的内存及存储构件 165。客户端设备 150 为包括输入/输出处理构件的任何设备, 诸如膝上型计算机、桌上型计算机、笔记本型计算机、移动电话或其它多功能手持型移动计算设备。以服务器模式 135 操作的数据处理系统 100 可操作以与多个客户端设备 150 通信且接收来自多个客户端设备 150 的指令。

[0037] 图 2 详细描述如本领域中公知的信息管理系统 200。然而, 本领域技术人员将了解, 本发明适用于任何活动数据库应用或电子表格应用, 藉此分析并报告数据。此定义被视为涵盖在线交易处理应用, 以及数据挖掘应用、关系型数据库应用、多维数据库以及电子表格应用等, 它们共享与数据库应用等关联的许多特性。如根据各优选实施例解释的本发明旨在适用于以格式化方式显示来自所生成的查询结果的任何应用, 且其中当所显示数据移动至格式化显示器中的另一位置时, 与所显示数据关联的批注需要与所显示数据持续地关联。

[0038] 数据存储器 225 存储与活动或实体有关的数据, 以便在由信息管理系统 200 进行分析时提供历史、当前及预测性分析以及商业运营 (business operation) 的视图。数据存储器 225 可采取如在本领域中公知的数据仓库或数据超市的形式。

[0039] 数据存储器 225 可与信息管理应用 200 分离, 即, 位于不同服务器上但在同一服务器机柜内, 或位于与信息管理系统 200 不同的地理位置。备选地, 数据存储器 225 可与信息管理系统 200 位于同一服务器上。数据可采取结构化数据的形式, 通常, 结构化数据是通过数据模型而建模的数据。备选地, 数据可为非结构化的数据, 即, 可见于需要语义分析以便分析并报告数据的含义的电子邮件、SMS、即时消息传送中的数据。

[0040] 在一个优选实施例中,建模器(modeller)组件220提供其中对基础数据建模以便描述表示和存取数据的方式的手段。数据模型通常定义数据元素及所定义数据元素之间的关系。数据模型可为关系实体模型或星型或雪花型模式等。

[0041] 查询构建器205使得用户能够通过查询生成器组件230构建查询以供执行。查询生成器组件230分析位于相关数据储存器225中的数据,且以视图或报告215的形式返回相关数据。查询可采取“1998年在多个地理区中销售了多少台电视机?”的形式。视图生成器210接收来自查询生成器组件230的所生成查询,且生成用于显示查询结果的视图215。

[0042] 视图215或报告215可为所述形式或通过用户定义的任何格式。视图215亦可描述为原始数据集的一个方面的视图215,所述原始数据集已基于用户定义的准则被查询。术语视图215用于描述任何形式的输出显示,藉此向用户显示查询结果。视图215或报告215通常在已对数据源执行查询之后生成。视图215或报告215显示查询结果。术语“视图”将贯穿本说明书的剩余部分使用,且理解为涵盖查询的所有显示输出。

[0043] 图4中示出查询的简化输出。视图215为“2001年全世界销售了多少台ABC公司的电视机”的查询的输出。藉此,第一列400列出相关产品(即,电视机),第二列405列出地理区,第三列410列出销售的数目或测量,且第四列415列出时间标度(即,所查询的时间段)。每一行435显示查询结果,且由列标题进行分类。列400与行435之间的相交点为单元格440,或亦称为维度425,并且维度包括数据值445。可将数据值视为可从查询生成的数据。

[0044] 图3详细描述本发明的一个优选实施例。存在与现有技术信息管理应用共享的多个核心组件,即,数据储存器225、建模器组件220及查询构建器205,因此,将不进一步解释这些组件。

[0045] 根据本发明的一个优选实施例,修改后的信息管理系统300包括批注数据储存器305、修改后的查询生成器组件340、用于生成视图335的修改后的视图生成器组件310、映射组件315、映射表325、用于创建索引元组表的索引创建器组件320,及代码生成器组件330。

[0046] 当用户提交查询时,查询构建器组件230将查询格式化为查询语言,且查询生成器组件340查询数据储存器225以搜寻所需信息。在本发明的一个优选实施例中,查询生成器组件340将查询作为元组集合返回至视图生成器组件310。元组包括值的集合,其中元组中的所述值中的每个值表示数据库表中的一个维度。例如,如果生成询问“2001年在德国销售了多少台电视机”的查询,则查询可返回以下数据:

[0047] “2001年在德国销售了500台电视机”

[0048] 本领域技术人员将认识到,此实例仅出于说明目的,且通常可以如在图4中展示的结构化表格格式返回结果。

[0049] 然而,出于高效存储及取回的目的,结果可如下作为元组存储于数据储存器305中:

[0050] <产品=电视机,国家=德国,时间=2001,测量=500>

[0051] 本领域技术人员将认识到,以上元组仅出于说明目的,且在实际中元组将具有更复杂的数据结构。

[0052] 重要的是指出:元组不仅存储查询结果(数据值),而且存储与查询结果有关的行

标题、列标题及子标题。因此，元组存储数据值及数据值的属性。将贯穿说明书使用术语数据项以描述数据值及其关联维度或属性。

[0053] 再次参看图 4，其展示针对 ABC 公司的等离子电视机的销售的查询的结果的简化视图。仅出于说明目的，展示第五列，其中用户可将注释 / 批注 440 添加至报告中的行中。在此实例中，用户已添加的对 2001 年在德国销售的电视机的数目的批注为“进步”440。此批注 440 与整个行 435 关联，但亦可与数据项 425、445、450 之一而非组成该行的全体数据项关联。因此，对于此实例，元组可以为如下：

[0054] < 产品 = 电视机，国家 = 德国，时间 = 2001，测量 = 500，批注 = 进步 >

[0055] 批注 440 可采取字符串、整数值或至外部数据源的链接或指针的形式。外部数据源 225 可为网页、文档或用于传送信息的任何其它形式。

[0056] 批注 440 可与显示在组成整个报告的单元格、列及行中的任何数目的数据项关联，或批注可与关联于特定单元格位置的单个数据项关联。

[0057] 在用户正查看报告时，批注 440 可与数据项相关联，或在刷新查询且更新维度并将其显示于不同视图中时，可显示批注。

[0058] 批注 440 可显示于一个或多个关联数据项正被显示在其中的单元格位置 425、450、455 处或显示于额外列中，如图 4 中所说明。备选地，批注 440 可通过使鼠标“暂留”在包括与批注关联的数据项(多个)的行上来显示。批注 440 可经由其它显示手段显示于对话框中，所述对话框经由鼠标或菜单功能操作来触发。本领域技术人员将认识到，在不偏离本发明的范围的情况下，存在显示与数据项关联的批注 440 的多种方式。

[0059] 批注 440 存储在批注数据存储器 305 内的批注表中。批注 440 可被修正或删除(或与数据项关联的其它批注)，且所有改变被更新并反映在批注数据存储器 305 内的批注表中。

[0060] 因此，批注 440 亦可与显示在列而非行中的汇总项目关联，这是因为批注 440 可与视图或报告中的任何单元格位置中的任何数据值 440 或数据值的属性 400、405、410、415、420 关联。

[0061] 为了将批注与显示在视图 335 中的一个或多个数据项相关联，将批注链接至作为查询结果的一部分而生成的元组。对于所生成视图的一个方面，将存在许多元组。

[0062] 例如，使用展示于图 4 中的视图，包括“进步”的批注的元组可如下：

[0063] < 产品 = ABC 公司的等离子电视，地理区 = DE，时间 = 2001，值 = 5000，批注 = 进步 >

[0064] 参看图 5，包含在以上元组中的数据项存储于批注存储器 305 内的元组索引表 505 中，且批注 440 存储在批注存储器 305 内的批注表 510 中。存储在元组索引表 505 中的元组经由唯一生成的关键字 550 链接至其在批注表 510 中的关联批注 440。然而，本领域技术人员将认识到，在不偏离本发明的范围的情况下，存在可能的其它存储配置。在以上实例中，字符串“进步”440 亦将存储在批注数据存储器 305 中的批注表 510 中。然而，如果批注 440 引用网页，则至网页的 URL 引用将存储在批注数据存储器 305 内的批注表 510 中。

[0065] 为了取回与元组关联的批注，索引创建器组件 320 创建所存储元组的索引 520。

[0066] 在此实例中，索引创建器表 500 包括多个行 525 至 555，每一行 525 至 555 表示元组或基础数据模式中的唯一识别的数据项。例如，如果基础数据模式是包括事实表(其具有

以下维度：包括国家、地址及邮政编码的地理区，包括电视机、收音机、音频系统、烤面包器的产品，包括 2001、2002、2003、2004 及 2005 的年份)的星型模式，则这些数据项亦可在索引创建器表 500 中列出。因此，表内的行的数目随着与星型模式关联的维度的数目而线性地增大。备选地，可在第一次提交对元组中的数据项的存储操作时（即，在检测到对批注及批注的关联元组的保存操作时）创建索引创建器表 500 中的行 525 至 555。

[0067] 首先，索引创建器组件 320 分析待提交至存储器的元组、检测元组中的第一数据项、在索引创建器表 500 中执行查找，且检测在元组中识别的第一数据项是否位于表的行 525 至 555 中。

[0068] 如果判定为否定的，即，第一数据项并不存在，则将第一数据项放置在表 500 的行 525 至 555 中，且向第一数据项给予唯一生成的识别符 515。所生成的识别符 515 存储在索引创建器表 500 中。针对元组中的每一数据项继续此处理，直至元组中的所有数据项均已被分析为止。对于每一提交操作，可存在具有许多数据项的许多元组。

[0069] 例如，采用如下元组：

[0070] < 国家 = 德国，产品 = 电视机，时间 = 2001，测量 = 5000，批注 = 进步 >

[0071] 索引创建器组件 320 通过查看位于元组中的第一数据项（即，“国家”）而开始，且识别在索引创建表 500 中不存在针对“国家”的表项，并将数据项“国家”放置在表 500 内的可用行 525 中，且向“国家”指派为（例如）值 1 的唯一识别符 515（针对表 500 中的每个唯一表项生成唯一值）。索引创建器组件 320 定位元组中的下一项目（即，“德国”），判定在索引创建器表中不存在针对“德国”的表项且将数据项“德国”添加至索引创建器表 500 中的下一可用行 530，并向数据项“德国”指派值“2”。接下来，索引创建器组件 320 定位元组中的下一项目（即，“产品”），判定在索引创建器表 500 中不存在针对“产品”的表项且将数据项“产品”添加至索引创建器表 500 中的下一可用行 540，并向数据项“产品”指派值“4”。接下来，索引创建器组件 320 单位元组中的下一项目（即，“电视机”），判定在索引创建器表中不存在针对“电视机”的表项且将数据项“电视机”添加至索引创建器表 500 中的下一可用行 545，并向数据项电视指派值“5”。接下来，索引创建器组件 320 定位元组中的下一项目（即，“时间”），且判定在索引创建器表 500 中不存在针对“时间”的表项且将数据项“时间”添加至索引创建器表 500 中的下一可用行 545，并向数据项“时间”指派值“10”。针对在元组中识别的每个数据项继续此处理，使得每个数据项已记录于索引创建表 500 中，且生成唯一值并使所述唯一值与数据项中的每一个相关联。

[0072] 如果索引创建器组件 320 识别元组中的批注属性，则索引创建器组件 320 将批注值（即，字符串、整数值、至其它信息的指针或链接）写入数据存储器 305 内的批注表 510 中，且创建唯一关键字 550 并将唯一关键字 550 与批注 440 相关联。对于任何给定元组，可存在许多批注。唯一关键字 550 亦与元组索引表 505 中的数据值的相应集合 520 相关联。其它信息（诸如，创建批注的人员及创建批注的日期及时间）亦可与批注一起存储。

[0073] 针对提交至存储器的每一元组及关联批注执行上述处理。当用户在查看视图的同时添加批注时，或每当刷新报告且将所有现存批注“拉入”报告中时，此情形可发生。

[0074] 在随后检测到至批注数据存储器 305 的提交操作时，索引创建器组件 320 将再次分析元组的每一数据项。当索引创建器组件 320 在索引创建表 500 中执行查找且检测到元组的数据项已记录在索引创建表 500 中时，索引创建器组件 320 接着移动至元组中的下一

项目且检测下一数据项是否已记录在索引创建表 500 中。如果已记录该数据项,则索引创建器组件 320 再次移动至元组中的下一数据项,直至已分析且对照索引创建表 500 中记录的表项检查所有数据项为止。仅在索引创建器组件 320 判定元组的数据项未记录在索引创建表 500 中时,索引创建器组件 320 才将数据项记录在索引创建表 500 中且生成唯一识别符 515 以与记录在索引创建表 500 中的项目关联。数据项亦包括与视图中的数据项相关联的批注。

[0075] 如果索引创建器组件 320 检测到数据项已记录在索引创建表 500 中,则索引创建器组件 320 识别与数据项关联的唯一识别符 515 且将该唯一识别符 515 写入元组索引表 505。

[0076] 因此,得到的值集合 520 是唯一地识别元组中的所有数据项(包括任何关联批注 440)的值集合。该值集合 520 存储在元组索引表 505 中,且批注存储在批注表 510 中。备选地,可以一起存储值集合 520 及批注 440。

[0077] 作为额外步骤,所述值集合中的每个值可使用已知散列技术进行散列以便提供批注的更快搜索及取回。

[0078] 因此,存储在批注数据存储器 305 中的是这样的值集合:其唯一地识别与批注 440 关联的维度(即,视图中的位置参考点(单元格位置))的集合。因此,当由于 a) 来自外部数据源 225 的数据已被刷新而刷新视图 335 或 b) 已重新运行查询而刷新视图 335 时,接着对于视图 335 中的每一维度集合,映射组件 315 查询元组索引表 505 以识别是否存在与显示于当前视图 335 中的维度匹配的维度(即,元组参考)的集合。如果定位相同的参考集合,则从批注数据存储器 305 中的批注表 510 查询关联批注,且将其取回以便与适当数据项一起显示在视图 335 中。

[0079] 由于基础数据源已被更新或查询自身已改变,所以可刷新查询。因此,当查询构建器组件 230 接收新查询以进行处理时,以上文所描述的方式处理该查询。然而,此次视图生成器组件 310 需要判定待生成的视图是否包括需要与关联数据项一起显示的任何批注 440。此处理通过映射组件 315 来处置。

[0080] 当查询生成器组件 340 将来自查询的元组集合返回至视图生成器 310 时,映射组件 315 拦截查询生成器组件 340 与视图生成器组件 310 之间的此通信,且通过分析元组中的数据项而开始。

[0081] 如前所述,识别元组中的第一数据项,且在索引创建表 500 中执行查找,以识别与第一数据项关联的唯一识别符。映射组件 315 将所识别的值写入至内存。接下来,映射组件 315 识别元组中的第二数据项并在索引创建表 500 中执行查找,且定位与第二数据项关联的唯一识别符并将所识别的值写入至内存。此处理继续,直至元组中的每一数据项均与在索引创建表 500 中定位的唯一值关联为止。

[0082] 因此,映射组件 315 创建唯一地识别元组中的数据项的组组合的值集合,该元组为查询的输出。然而,如果映射组件未能在索引表中定位数据项,则处理停止,且需要针对索引创建器组件中的数据项创建唯一值。

[0083] 接下来,映射组件 315 获取创建的值集合且在元组索引表 505 中执行查找,以判定是否存在记录于元组索引表 505 中的对应的唯一值集合。如果映射组件 315 识别到对应的值集合,则连同关联批注 440 一起将所识别的值集合取回至视图生成器组件 310 以便生成

视图 335, 视图 335 现在包括与特定元组关联的批注 440。

[0084] 例如, 如果查询询问以下问题:

[0085] “2001 年在英国销售了多少台电视机?”

[0086] 则查询生成器组件 340 将返回以下元组:

[0087] <产品=电视机, 国家=英国, 时间=2001, 测量=80,000>

[0088] 映射组件 315 获取以上元组, 且在索引创建表 500 中执行查找并生成以下值集合:

[0089] <4.5.1.3.10.11.7.8>

[0090] 此时并不知晓是否存在已针对元组<产品=电视机, 国家=英国, 时间=2001, 测量=80,000>而存储的任何批注。

[0091] 因此, 映射组件 315 在元组索引表 505 中执行查找以搜寻值集合<4.5.1.3.10.11.7.8>, 从而识别对应的值集合 520。如果找到, 则映射组件 315 将此值写入至映射表 325。映射组件 315 针对在元组中识别的每一项目继续此处理。例如, 如果元组为:

[0092] <产品=电视机, 国家=英国, 测量=250000>

[0093] 则映射组件 315 将使用来自索引创建表 500 的信息得到以下值集合。

[0094] <4.5.1.3.7.9>

[0095] 映射组件 315 在元组索引表 505 中执行查找以搜寻对应的值组合 520。然而, 数据值的次序无关紧要, 只要值集合 520 包括相同数据值。如果在单个值集合中找到对应的值组合, 则取回所识别的值集合。执行进一步的查找以判定是否存在与所识别的值集合关联的唯一生成的关键字 550, 且如果存在, 则使用唯一关键字 550 来取回关联批注。

[0096] 执行反向查找以从所述值集合“回到”数据项以便使视图生成器组件 310 能够生成视图 335, 视图 335 显示所请求的查询的结果及任何关联批注 440。

[0097] 图 6a 例示了视图 335 的一个实例, 视图 335 包括许多不同零售渠道的列标题 605。每个行 610 与不同时间段有关, 且单元格 615 中的数据 615 为特定时间段的销售量。

[0098] 标签 a、b 及 c 描绘与箭头指向的数据值关联的批注。

[0099] 图 6b 例示了与图 6a 中例示的视图相同的视图, 但是例示由索引创建器组件 325 针对图 6a 中例示的数据而生成的元组。因此, 对于图 6a 中给出的批注, 给出以下维度:

[0100] a=(2004, 户外用品商店, 数量)

[0101] b=(2006, 高尔夫用品商店, 数量)

[0102] c=(2004, 体育用品商店, 数量)

[0103] 其中通过如参考图 5 描述的处理针对批注 a、b 及 c 而生成的元组为:

[0104] a=(12::10::3)

[0105] b=(26::19::8)

[0106] c=(14::10::5)

[0107] 且采取与批注“a”关联的元组, 可通过代码生成器组件 330 生成以下代码。

[0108]

```

<CD useValue="Outdoors Shop" pun="8" lun="7" mun="12" hun="9" dun="10" rdi="11" qry="4"
ctxId="3"/>
<CD useValue="2004" pun="21" lun="20" mun="19" hun="22" dun="23" rdi="24" qry="4"
ctxId="10"/>
<CD useValue="5801731" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="12"/>

```

[0109] 其中可以看到,批注“a”的元组值反映在称为 ctxid 属性的事项中。ctxid 属性的组合等效于图 5 的元组索引表 505 中所创建的索引。ctxid 属性唯一地识别文档内的元素。

[0110] 以上代码中的 pun/mun/lun/hun/dun 属性指代所生成视图 335 的数据项区段中的行。此情形更详细地展示于以下所生成的代码中。

[0111]

```

<DA vtype="2" lun="[Sales].[Retailer].[Retailer].[Retailer type]" uid="7"/>
<DA vtype="1" pun="[Sales].[Retailer].[Retailer].[Retailer(All)]-&gt;[all]" uid="8"/>
<DA vtype="6" hun="[Sales].[Retailer].[Retailer]" uid="9"/>
<DA vtype="7" dun="[Sales].[Retailer]" uid="10"/>
<DA vtype="3" rdi="Retailer(All) (visible items with calculations set)" dtype="1" drill="0"
usage="3" q="4" h="9" level="1" uid="11"/>
<DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[6]" uid="12"/>

```

[0112] 以下为来自 HTML 生成器组件 330 的针对图 6a 及图 6b 中所示的表的实例代码输出。

[0113]

```

<CONTEXT-METADATA>
  <META-DATA>

```

```

    <DA vtype="0" mun="[Sales].[Sales fact].[Quantity]" uid="1"/>
    <DA vtype="6" hun="[Sales].[Sales fact]" uid="2"/>
    <DA vtype="7" dun="[Sales].[Sales fact]" uid="3"/>
    <DA vtype="5" qry="Query1" uid="4"/>
    <DA vtype="9" uid="5"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[4]" uid="6"/>
    <DA vtype="2" lun="[Sales].[Retailer].[Retailer].[Retailer type]" uid="7"/>
    <DA vtype="1" pun="[Sales].[Retailer].[Retailer].[Retailer(All)]-&gt;[all]" uid="8"/>
    <DA vtype="6" hun="[Sales].[Retailer].[Retailer]" uid="9"/>
    <DA vtype="7" dun="[Sales].[Retailer]" uid="10"/>
    <DA vtype="3" rdi="Retailer(All) (visible items with calculations set)" dtype="1"
drill="0" usage="3" q="4" h="9" level="1" uid="11"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[6]" uid="12"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[7]" uid="13"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[8]" uid="14"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[2]" uid="15"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[5]" uid="16"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[1]" uid="17"/>
    <DA vtype="0" mun="[Sales].[Retailer].[Retailer].[Retailer type]-&gt;[all].[3]" uid="18"/>

```

[0114]

```

<DA vtype="0" mun="[Sales].[Time dimension].[Time dimension].[Year]-&gt;[all].[2004]"
uid="19"/>
<DA vtype="2" lun="[Sales].[Time dimension].[Time dimension].[Year]" uid="20"/>
<DA vtype="1" pun="[Sales].[Time dimension].[Time dimension].[Time dimension(All)]-
&gt;[all]" uid="21"/>
<DA vtype="6" hun="[Sales].[Time dimension].[Time dimension]" uid="22"/>
<DA vtype="7" dun="[Sales].[Time dimension]" uid="23"/>
<DA vtype="3" rdi="Time dimension(All) (visible items with calculations set)" dtype="1"
drill="0" usage="3" q="4" h="22" level="1" uid="24"/>
<DA vtype="3" rdi="Default Measure" dtype="6" drill="0" usage="2" q="4" h="2"
level="0" uid="25"/>
<DA vtype="0" mun="[Sales].[Time dimension].[Time dimension].[Year]-&gt;[all].[2006]"
uid="26"/>
<DA vtype="0" mun="[Sales].[Time dimension].[Time dimension].[Year] -
&gt;[all].[2005]" uid="27"/>
<DA vtype="0" mun="[Sales].[Time dimension].[Time dimension].[Year]-&gt;[all].[2007]"
uid="28"/>
</META-DATA>
<CONTEXT-DATA>
<Block id="1">
<CD useValue="Quantity" mun="1" hun="2" dun="3" rdi="5" qry="4" ctxld="1"/>
<CD useValue="Warehouse Store" pun="8" lun="7" mun="6" hun="9" dun="10"
rdi="11" qry="4" ctxld="2"/>
<CD useValue="Outdoors Shop" pun="8" lun="7" mun="12" hun="9" dun="10"
rdi="11" qry="4" ctxld="3"/>
<CD useValue="Eyewear Store" pun="8" lun="7" mun="13" hun="9" dun="10" rdi="11"
qry="4" ctxld="4"/>
<CD useValue="Sports Store" pun="8" lun="7" mun="14" hun="9" dun="10" rdi="11"
qry="4" ctxld="5"/>
<CD useValue="Department Store" pun="8" lun="7" mun="15" hun="9" dun="10"
rdi="11" qry="4" ctxld="6"/>
<CD useValue="Equipment Rental Store" pun="8" lun="7" mun="16" hun="9"
dun="10" rdi="11" qry="4" ctxld="7"/>
<CD useValue="Golf Shop" pun="8" lun="7" mun="17" hun="9" dun="10" rdi="11"
qry="4" ctxld="8"/>
<CD useValue="Direct Marketing" pun="8" lun="7" mun="18" hun="9" dun="10"
rdi="11" qry="4" ctxld="9"/>
<CD useValue="2004" pun="21" lun="20" mun="19" hun="22" dun="23" rdi="24"
qry="4" ctxld="10"/>
<CD useValue="1723153" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="11"/>
<CD useValue="5801731" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="12"/>
<CD useValue="610446" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="13"/>
<CD useValue="5136252" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="14"/>
<CD useValue="5055618" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="15"/>
<CD useValue="172689" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="16"/>
<CD useValue="1010427" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="17"/>
<CD useValue="664414" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="18"/>
<CD useValue="2006" pun="21" lun="20" mun="26" hun="22" dun="23" rdi="24"
qry="4" ctxld="19"/>
<CD useValue="1320698" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="20"/>
<CD useValue="9611226" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="21"/>
<CD useValue="1014217" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="22"/>
<CD useValue="6221728" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="23"/>
<CD useValue="4791417" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="24"/>
<CD useValue="395535" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="25"/>
<CD useValue="1715644" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="26"/>
<CD useValue="871325" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="27"/>
<CD useValue="2005" pun="21" lun="20" mun="27" hun="22" dun="23" rdi="24"
qry="4" ctxld="28"/>
<CD useValue="1182485" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="29"/>
<CD useValue="8114460" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="30"/>
<CD useValue="767413" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="31"/>
<CD useValue="5940504" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="32"/>
<CD useValue="5156496" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="33"/>
<CD useValue="308789" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxld="34"/>

```

[0115]

```

<CD useValue="1448435" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="35"/>
<CD useValue="606113" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="36"/>
<CD useValue="2007" pun="21" lun="20" mun="28" hun="22" dun="23" rdi="24"
qry="4" ctxId="37"/>
<CD useValue="885424" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="38"/>
<CD useValue="8188015" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="39"/>
<CD useValue="919735" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="40"/>
<CD useValue="4614255" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="41"/>
<CD useValue="2953809" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="42"/>
<CD useValue="377288" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="43"/>
<CD useValue="1235010" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="44"/>
<CD useValue="422350" mun="1" hun="2" dun="3" rdi="25" qry="4" ctxId="45"/>

```

</Block>

</CONTEXT-DATA>

</CONTEXT-METADATA>

[0116] 代码生成器组件 330 将输出传送至视图生成器组件 310, 以便呈现输出至图 6a 中示出的表的视图 335。因此可以看到, 一旦批注与数据项关联, 基础数据源是否被更新、查询是否被刷新以及不同维度集合是否显示在视图中便为无关紧要的, 因此改变视图的原始组织结构为无关紧要的, 因为批注将始终与其关联数据项一起显示。

[0117] 还可以在报告级别将批注分组在一起, 且因此可显示仅与特定报告有关的批注。这通过引入报告 id 来完成, 且可将批注链接至报告 id 作为过滤机制。可引入其它分组和过滤机制以便向视图的不同级别和方面提供“向下切入 (drill down) 及钻研 (drill through)”能力。

[0118] 图 7 例示用于在用户刷新并查看查询时取回批注的处理流程。

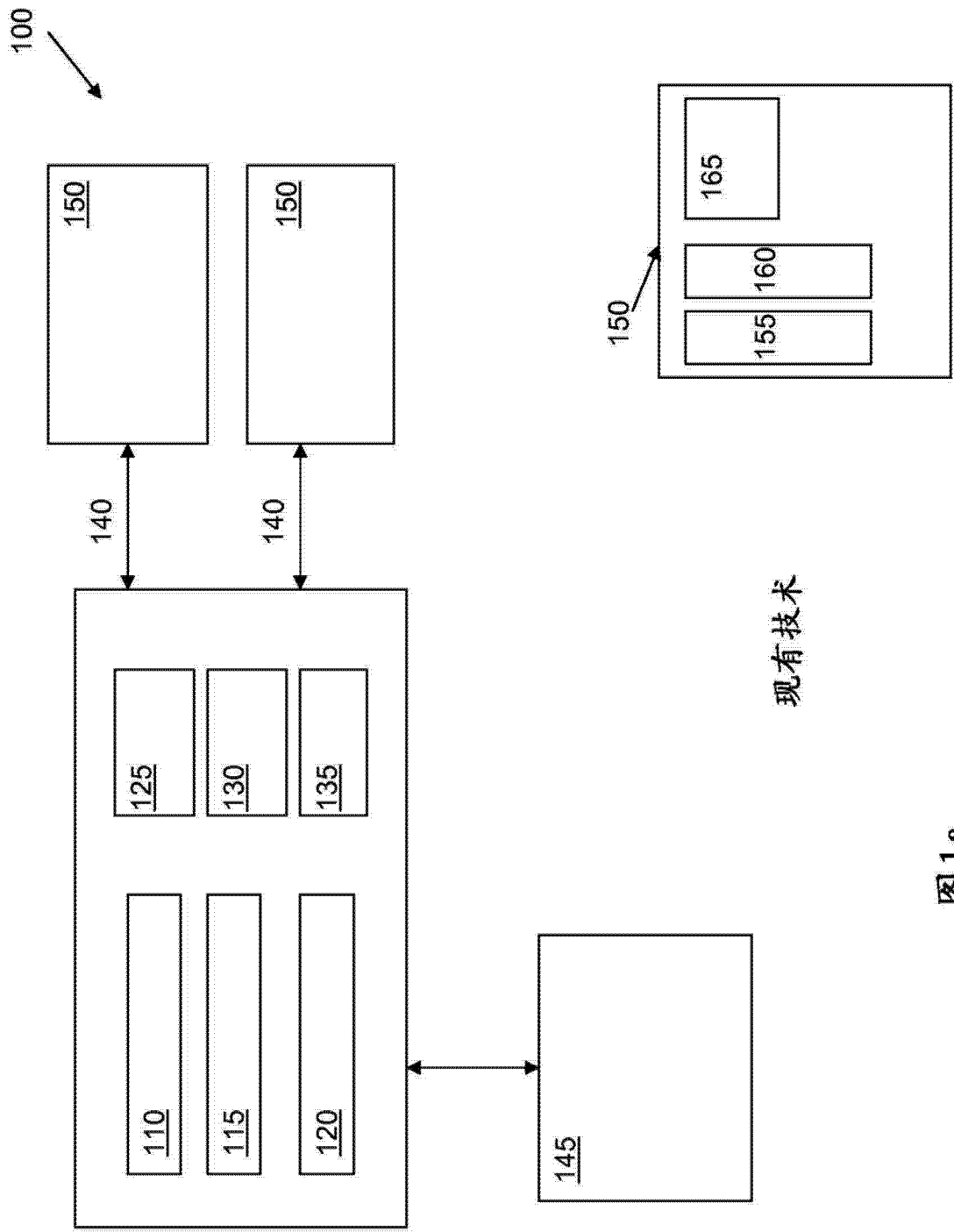
[0119] 在步骤 700, 查询生成器组件 340 接收查询, 且对基础数据源 225 进行查询。返回结果的集合。在步骤 705, 通过映射组件 315 分析查询的输出以识别输出的每个维度中的一个或多个数据项, 以便识别与每个数据项关联的唯一值 515。在步骤 710, 记录 505 每个所识别的唯一值 515, 并且其中每个所识别的唯一值形成表示查询输出中的每个所识别数据项的唯一值集合。在步骤 715, 映射组件 315 从数据存储器 305 识别先前记录的唯一值集合 520, 其对应于当前正在被分析的唯一值集合 520。在步骤 720, 响应于肯定的判定而判定先前记录的唯一值集合 520 是否包括关联批注, 且在步骤 725, 响应于肯定的第二判定而从数据存储器取回批注。

[0120] 本发明可采用完全硬件实施例、完全软件实施例或含有硬件及软件元素两者的实施例的形式。在一个优选实施例中, 本发明以软件实施, 所述软件包括 (但不限于) 固件、驻留软件、微代码等。

[0121] 此外, 本发明可以采取可从计算机可用或计算机可读介质访问的计算机程序产品的形式, 所述计算机可用或计算机可读介质提供了可以被计算机或任何指令执行系统使用或与计算机或任何指令执行系统结合的程序代码。出于此描述的目的, 计算机可用或计算机可读介质可以是任何能够包含、存储、传送、传播或传输由指令执行系统、装置或设备使用或与所述指令执行系统、装置或设备结合的程序装置。

[0122] 所述介质可以是电、磁、光、电磁、红外线或半导体系统 (或装置或设备) 或传播介质。计算机可读介质的实例包括半导体或固态存储器、磁带、可移动计算机盘、随机存取存储器 (RAM)、只读存储器 (ROM)、硬磁盘和光盘。光盘的当前实例包括压缩盘 - 只读存储器 (CD-ROM)、压缩盘 - 读 / 写 (CR-R/W) 和 DVD。

[0123] 可在不偏离本发明的范围的情况下对前述内容做出改进及修改。



现有技术

图 1a

图 1b

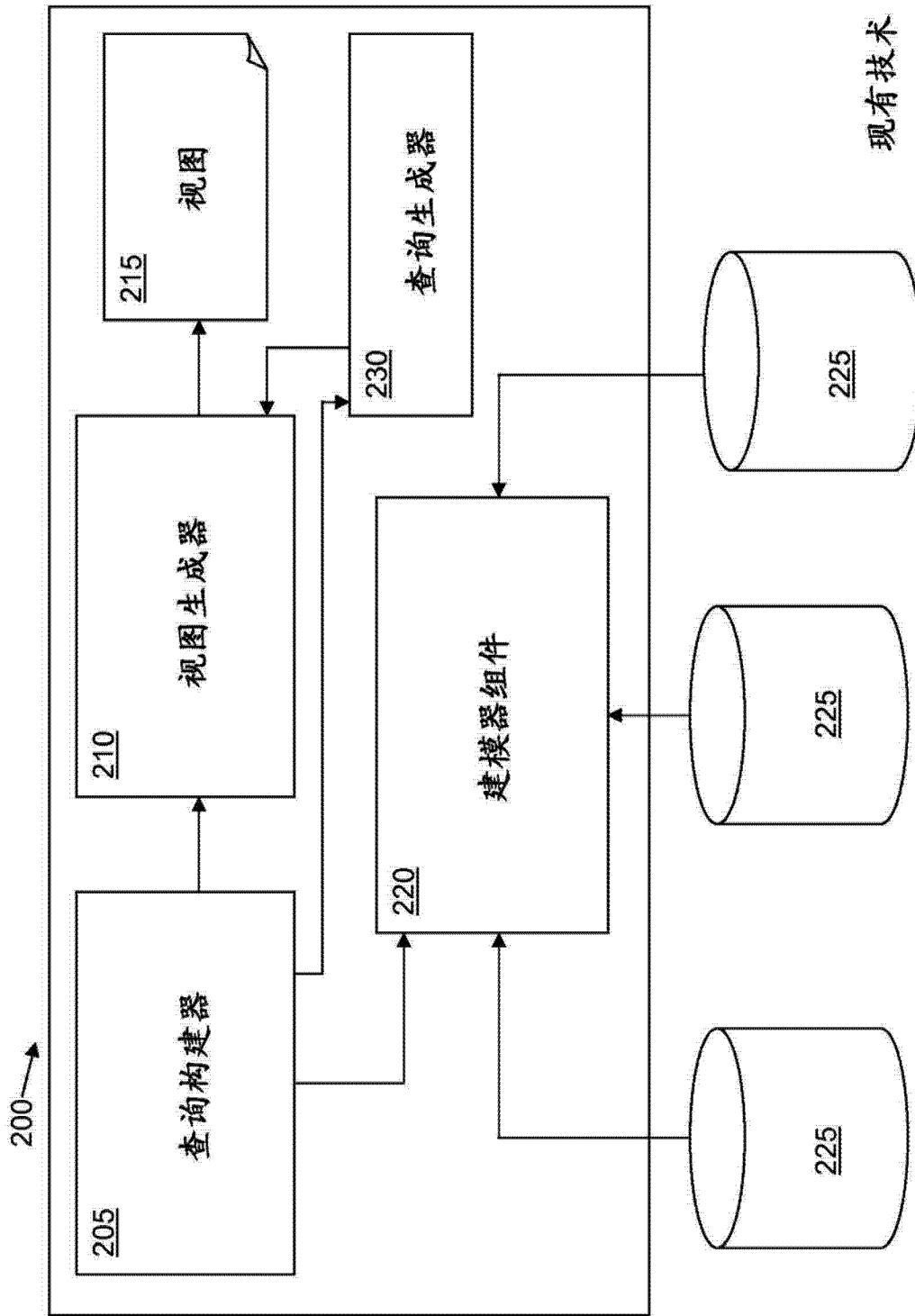


图 2

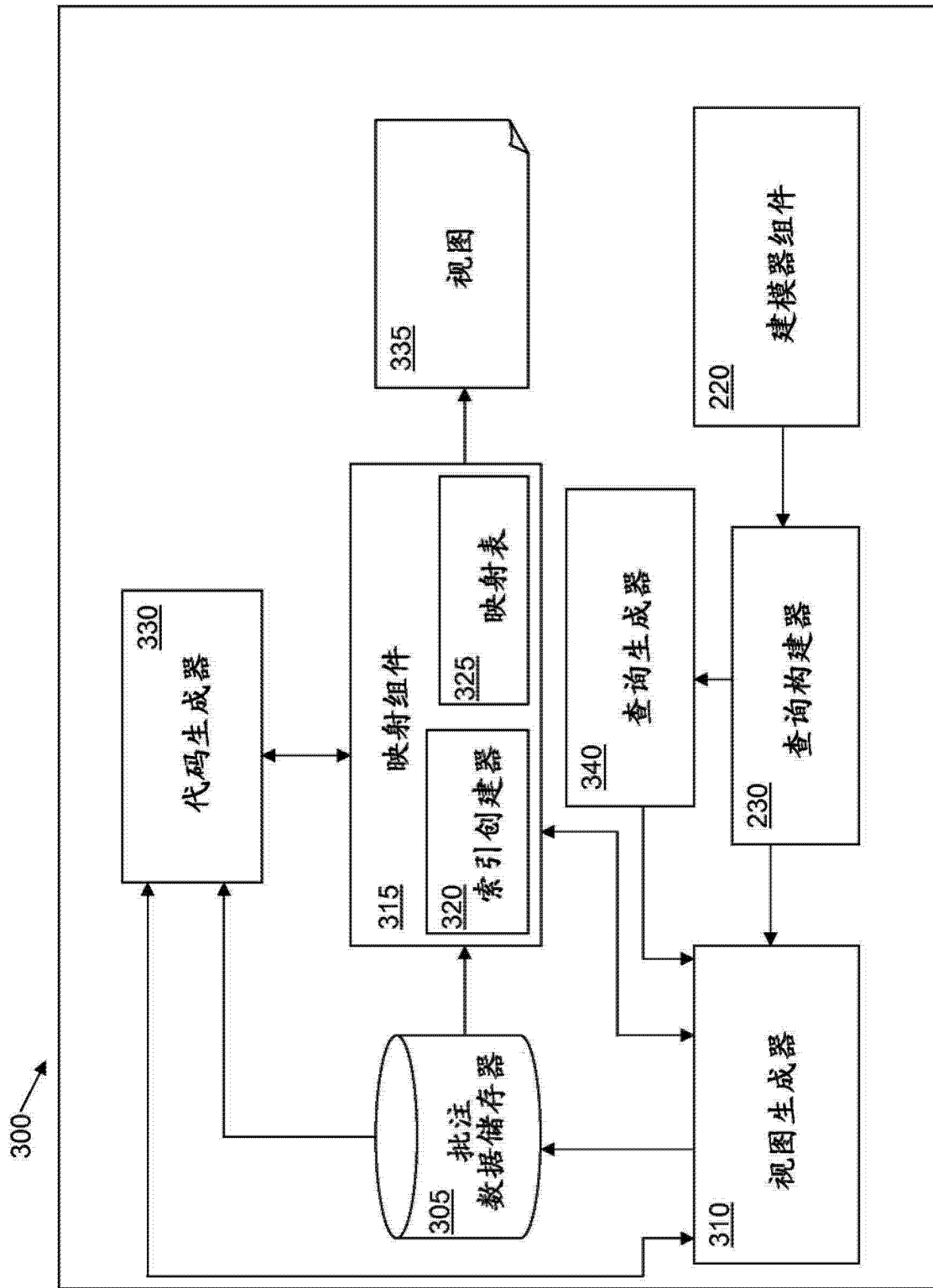


图 3

产品	地理区	销售量	时间	批注
ABC公司的电视 <u>435</u>	德国	5000	2001	进步 <u>440</u>
ABC公司的电视	英国	4500	2001	
ABC公司的电视	美国	10000	2001	预期
ABC公司的电视	<u>425</u> 爱尔兰	300 <u>445</u>	2001	
ABC公司的电视	瑞士	1500	2001	
ABC公司的电视	澳大利亚	<u>450</u> 500	<u>460</u> 2001	失望
ABC公司的电视	比利时	750	2001	
ABC公司的电视	法国	<u>455</u> 2500	2001	达到目标

图 4

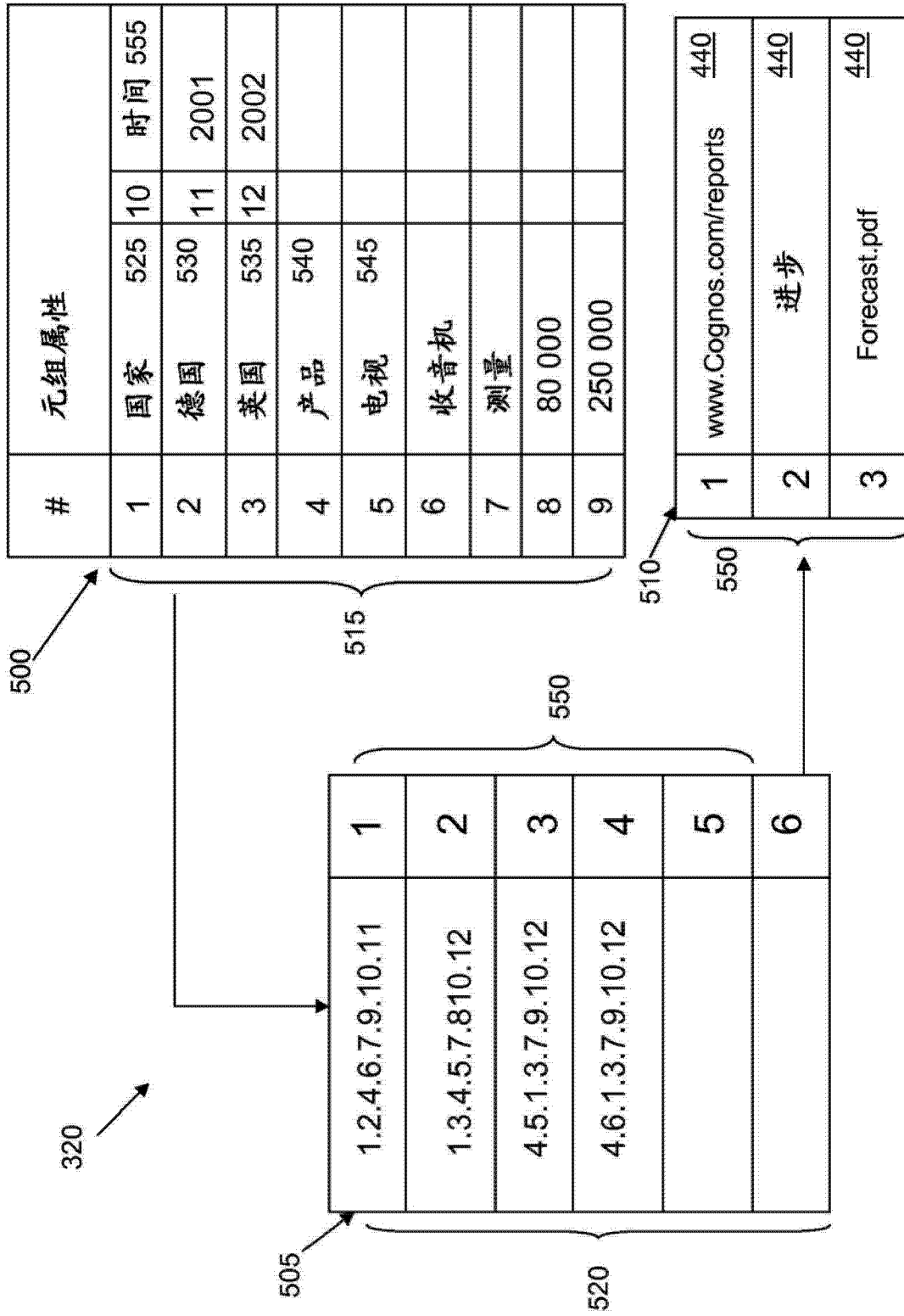


图 5

数量	量贩店	户外用品商店	眼镜商店	体育用品商店	百货商店	设备租赁商店	高尔夫用品商店	直接销售
2004	1,723,153	5,801,731	610,446	5,136,252	5,055,618	172,689	1,010,427	664,414
2006	1,320,698	9,611,226	1,014,217	6,221,728	4,791,417	395,535	1,715,644	871,325
2005	1,182,485	8,114,460	767,413	5,940,504	5,156,486	308,789	1,448,435	606,113
2007	885,424	8,188,015	919,735	4,614,255	2,953,809	377,288	1,235,010	422,350

图 6a

1	2	3	4	5	6	7	8	9
10	11::10::2	12::10::3 (a)	13::10::4	14::10::5 (c)	15::10::6	16::10::7	17::10::8	18::10::9
19	20::19::2	21::19::3	22::19::4	23::19::5	24::19::6	25::19::7	26::19::8 (b)	27::19::9
28	29::28::2	30::28::3	31::28::4	32::28::5	33::28::6	34::28::7	35::28::8	36::28::9
37	38::37::2	39::37::3	40::37::4	41::37::5	42::37::6	43::37::7	44::37::8	45::37::9

图 6b

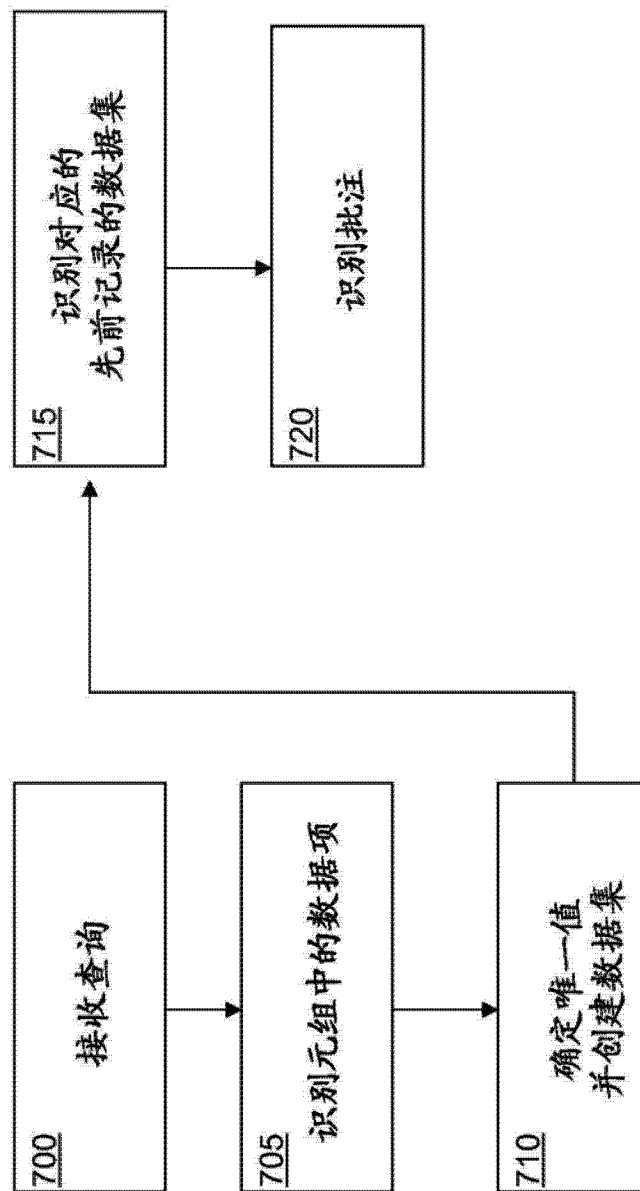


图 7