



## (12) 发明专利

(10) 授权公告号 CN 102542012 B

(45) 授权公告日 2015. 10. 28

(21) 申请号 201110420064. 3

CN 101339564 A , 2009. 01. 07,

(22) 申请日 2011. 12. 15

CN 101681352 A , 2010. 03. 24, 说明书第  
13 页第 2 段 - 第 14 页第 2 段, 附图 6.

(30) 优先权数据

12/969140 2010. 12. 15 US

US 2010306249 A1 , 2010. 12. 02, 说明书  
第 0008-0314 段, 附图 1-12.

(73) 专利权人 微软技术许可有限责任公司

WO 2010022655 A1 , 2010. 03. 04,

地址 美国华盛顿州

审查员 王艳臣

(72) 发明人 G. 库马 S. 阿哈里 F. 候赛尼

N. 罕 A. 阿杜尔卡德 A. 古普塔

G. 库马兰 V. 奈尔

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 谢建云 刘鹏

(51) Int. Cl.

G06F 17/30(2006. 01)

(56) 对比文件

CN 101169797 A , 2008. 04. 30,

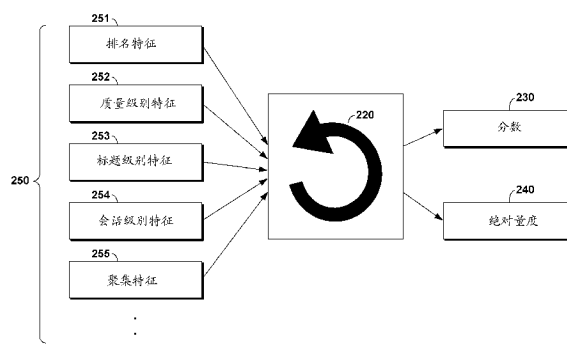
权利要求书2页 说明书12页 附图6页

(54) 发明名称

分类搜索查询的结果

(57) 摘要

提供了用于将搜索结果分类为高质量或低质量的计算机可读介质、计算机系统和计算方法。首先,选择部分搜索结果(如最高排名的文档)用以评价。使用分类处理确定每个选择的搜索结果的质量水平,分类处理包括以下步骤:确定由待评价的选定搜索结果展示的目标特征;评价所选择的特征来为每个选择的搜索结果产生质量水平分数;将该分数与预先定义的阈值进行比较;以及基于该比较为每个选择的搜索结果分配绝对量度。当分数少于阈值时,该绝对量度表明低质量。在识别出选择的搜索结果为低质量时,自动执行重新编制所发布的搜索查询的校正动作。



1. 一种采用搜索结果集的质量水平来激活一个或多个校正动作的方法,该方法包括:  
访问(502)该搜索结果集;

通过向其应用多个分类标准而单独地确定(504)该搜索结果集的质量水平,所述多个分类标准包括至少一个排名特征,它单独度量该搜索查询内的关键用语和该搜索结果集的内容之间的关系,以及至少一个标题级别特征,其列举该搜索查询内的关键用语和搜索引擎预备的、搜索结果集的概要之间的匹配数;

基于所确定的质量水平,为该搜索结果集中至少一个结果分配(506)在绝对标度中的分数;

确定该分数低于在绝对标度中的阈值(508);

计算搜索引擎撰写的搜索查询将提供改进的搜索结果的置信度级别;以及  
激活一个或多个与该计算的置信度级别相关联的校正动作。

2. 如权利要求1所述的方法,其中该方法进一步包括:

接收用户发布的搜索查询;以及

返回基于该搜索查询内的至少一个用语而产生的搜索结果集。

3. 如权利要求2所述的方法,其中该方法进一步包括:

部分地基于与该搜索查询中隐含的用户意图的相关性,将该搜索结果集中的搜索结果相对彼此排名;以及

以在排名中高排名的搜索结果为目标,来启动确定该搜索结果集的质量水平。

4. 如权利要求1所述的方法,其中该方法进一步包括,在比较揭示分数高于阈值时,将至少一个搜索结果分类为高质量。

5. 如权利要求1所述的方法,其中该方法进一步包括,在比较揭示该分数低于阈值时,将至少一个搜索结果分类为低质量。

6. 如权利要求2所述的方法,其中多个分类标准包括一个或多个质量级别特征,其单独量化该搜索结果集的整体质量而独立于搜索查询,其中整体质量是该搜索结果集的内容或集成到该搜索结果集的域名中的至少一个的函数。

7. 如权利要求2所述的方法,其中多个分类标准进一步包括一个或多个会话级别特征,其基于在在线会话过程期间用户发起的、对搜索结果的改进,将该搜索结果集中的搜索结果标记为相关或不相关。

8. 如权利要求2所述的方法,其中多个分类标准进一步包括一个或多个聚集特征,其考虑该搜索结果集的内容、在搜索结果页面中发布的广告或包括由搜索引擎与该搜索结果集一起提供的相关搜索的用语中的至少一个之间的变化。

9. 如权利要求1所述的方法,其中作为比较的函数来激活一个或多个校正动作包括:

当根据比较分数少于阈值而置信度级别高时,自动重写搜索查询;以及

发布重写的搜索查询以产生新的搜索结果集。

10. 如权利要求1所述的方法,其中作为比较的函数来激活一个或多个校正动作包括:

当根据比较分数少于阈值且置信度级别低时,呈现重新编制用户界面UI;以及

在该重新编制用户界面UI处接收替代或移除搜索查询中的一个或多个关键用语的、用户发起的选择。

11. 一种用于将搜索结果分类为高质量或低质量的方法,该方法包括:

提供(602) 响应于用户撰写和发布的搜索查询而返回的搜索结果集；

作为与发布的搜索查询的相关性的函数确定该搜索结果集相对彼此的排名(604)；

基于排名,从返回的被排名为与发布的搜索查询高度相关的搜索结果集中选择(606)一个或多个搜索结果来评价；

通过采用分类处理,独立于搜索查询用语,为从返回的搜索结果集中选择的一个或多个搜索结果中的每一个确定(608) 质量水平,其中该分类处理包括：

(a) 确定(610) 由一个或多个要评价的选定搜索结果展示的目标特征；

(b) 评价(612) 目标特征来为一个或多个选定搜索结果中的每一个产生质量水平分数；

(c) 将该分数相对于预先定义的阈值进行比较；以及

(d) 基于该比较,为一个或多个选定搜索结果中的每一个分配绝对量度,其中当分数高于阈值时,该绝对量度表明高质量,其中当分数少于阈值时,该绝对量度表明低质量；

计算搜索引擎撰写的搜索查询将提供改进的搜索结果的置信度级别；以及

响应于该计算的置信度级别,在识别出一个或多个选定搜索结果被分配了表明低质量的绝对量度时,用自动生成并自动重新编制的搜索查询自动进行搜索。

## 分类搜索查询的结果

### 背景技术

[0001] 搜索引擎一般设计为具有通过帮助用户以快速有效地方式来找到与他们的搜索查询相匹配的相关搜索结果(如文档)来简化用户生活的意图。例如,搜索引擎的用户可撰写并提交搜索查询,以针对服务、产品、特定信息或该用户在该搜索查询背后的意图所目标确定的其他任何数据来进行在线搜索。由传统搜索引擎产生的搜索结果集常常质量低下,这意味着该搜索结果集中至少一部分没有恰当满足该用户在该搜索查询背后的意图。特别地,在一项研究中,研究人员发现十个搜索结果集中大致有一个列在顶端的结果与该用户相应的搜索查询的意图不相关。

[0002] 产生低质量的搜索结果有各种原因。一些原因涉及该搜索引擎所询问的网络索引中缺少高质量的文档,或核心排名算法未能定位存在在该网络索引中的高质量文档。

### 发明内容

[0003] 提供本发明内容来以简化形式介绍将在下面具体实施方式中进一步描述的概念选集。本发明内容不意图标识所要求保护主题的关键特征或基本特征,也不意图孤立地作为确定所要求保护主题的范围的辅助。

[0004] 本发明的实施例将产生低质量搜索结果的各种原因归结为包括以下:搜索查询的贫乏用户构词具有与低质量文档或不满足用户在该搜索查询背后的意图的其他文档相匹配的格式。这些缺陷性撰写的搜索查询可要么包括极其常见的、可以大量不同方式解释的词条(产生离题的搜索结果),要么包括不匹配任何返回的有意义的文档的罕见词(产生模糊的搜索结果)。

[0005] 然而,搜索引擎将这些离题的或模糊的搜索结果返回给该用户而不管其是否实际上与该用户的意图相关。这给了该用户假象:因为搜索结果与其搜索查询相匹配,因此返回的搜索结果一定是相关的。然而经检查,用户常常发现这些搜索结果不相关。进一步地,如果该用户想要更好的、高质量的文档(在此可与短语“搜索结果”交换使用),该用户被迫重新撰写他们的搜索查询而没有得到任何帮助。搜索引擎的这个缺陷导致了低效率和混乱的用户搜索体验。

[0006] 因此,因为搜索引擎未能识别出作为该用户缺陷性撰写的搜索查询的结果而提供给该用户的低质量搜索结果,因此一个自动确定返回的搜索结果何时是低质量的、以及当确定低质量搜索结果时、采取校正动作来改进该搜索查询(即在在线搜索的环境下重新明确表达该用户的意图)的新颖搜索引擎会提供增强的整体用户体验。

[0007] 本发明的实施例解决许多用户不具有熟练的搜索查询表述技能的问题。例如,用户常常输入过于广泛的或过少包含的搜索查询,且在将该搜索查询发送至搜索引擎后,用户接收与该用户的真实意图不相关的搜索结果。其他实施例解决这样的有疑问的问题:一些搜索查询本来就模糊(如使用常见词)或难以满足(如使用罕见词)而不管怎么表述它们。

[0008] 相应地,本发明的实施例引入了检测不相关的搜索结果并辅助用户改进搜索结果相关性的各种方法。在一个示例中,对不相关搜索结果的检测由搜索引擎执行,它试图建立

分类处理,以便在运行时针对给定的用户发起的搜索查询,自动确定在搜索引擎结果页面(SERP)中的搜索结果的质量水平。在示例实施例中,该搜索引擎实现此分类处理来部分基于所确定的质量水平确定搜索结果就绝对意义(如,要么高质量要么低质量)而言是否可接受。进一步地,当这些搜索结果确定为低质量时,它们被如此标记以使得如果适当可执行一个或多个校正动作。

[0009] 因为对用户搜索查询的修改(如对其用语的小调整)常常改进了返回的搜索结果的相关性,所以校正动作之一可涉及代表该用户来隐含地重写该搜索查询。在其中该改变会改进搜索结果有较高置信度的情况下,该重写的搜索查询由该搜索引擎以透明方式自动地准备并提交。在其中该改变会改进搜索结果的置信度较低的情况下,修改关键用语(如移除或替换)的选项将明确地呈现给该用户,以征求他或她关于重新编制该搜索查询的反馈。除了帮助该用户重新编制他或她的搜索查询的之外,另一个校正动作还可包括呈现标识为高质量的搜索结果,而抑制低质量结果的那些。因此节省了该用户在低质量搜索结果中详细查看所涉及的时间。

## 附图说明

[0010] 下文中通过参考附图详细说明本发明的实施例,其中:

[0011] 图 1 是适合实现本发明的实施例的示例计算设备的框图;

[0012] 图 2 是示出了根据本发明的实施例的示例分类处理的各种输入和输出的框图;

[0013] 图 3 是适合在实现本发明的实施例中使用的示例网络环境的框图;

[0014] 图 4 是根据本发明的实施例的、在协作地实施在线搜索的实体之间的示例处理流的框图;

[0015] 图 5 是示出了根据本发明的实施例的、使用搜索结果集的质量水平来激活一个或多个校正动作的整体方法的流程图;以及

[0016] 图 6 是示出了根据本发明的实施例的、将搜索结果分类为高质量或低质量的整体方法的流程图。

## 具体实施方式

[0017] 在此公开的本发明的实施例的主题通过特别描述来满足法定要求。然而,描述本身不意图限制本专利的范围。相反地,发明者已经预期要求保护的主题也可以其他方式体现,来包括与本文档描述的步骤相似的不同步骤或步骤组合,并结合其他当前或未来的技术。

[0018] 根据下文公开部分可推断,本发明的各种实施例涉及确定搜索结果的质量水平,并可能利用基于任务的引擎(如决定引擎、任务引擎、单独应用或操作、小程序系统、操作系统以及基于任务的移动系统)或试图将搜索结果与在发布的搜索结果内固有的用户意图相匹配的一般系统,来采取校正动作以改进所确定的质量水平。为了简化论述,这些引擎和/或系统在此后将被称为“搜索引擎”。

[0019] 在此描述的本发明的实施例包括具有其上包含计算机可执行指令的计算机可读介质。在(如用处理器)执行时,该计算机可执行指令执行一种使用搜索结果集的质量水平来激活一个或多个校正动作的方法。在实施例中,该方法包括步骤:访问该搜索结果集,并

通过向那里应用多个分类标准来分别确定该搜索结果集的质量水平。在示例中,应用分类标准包括评价由搜索结果分别评价的排名特征、质量级别特征、标题级别特征、会话级别特征和 / 或聚集特征。下文参考图 2 更全面地讨论这些各种类型特征中的每一个。

[0020] 该方法可通过为该搜索结果集中至少一个搜索结果分配在绝对标度中的分数来继续,其中该分数基于所确定的质量水平。该分数可与在绝对标度的阈值进行比较。因此,可基于该比较的结果,选择并执行一个或多个校正动作。

[0021] 在第二个示例实施例中,本发明引入了一种用于进行至少一个搜索结果的质量的绝对分类的计算机系统。该计算机系统典型地包括计算机软件部件,这些部件包括下列:查询接收部件、分类部件、分配部件和应用部件。在实施例中,查询接收部件被配置为接收用户发布的搜索查询。典型地,该查询接收部件(即前端部件)返回基于该搜索查询而(如用后端部件)产生的搜索结果集。分类部件一般负责实现分类处理,其自动确定在该搜索结果集中的一个或多个搜索结果的质量分类。如下文参考图 3 更全面的讨论所述,该分类处理至少涉及以下步骤(没有特定顺序):检查该搜索结果呈现的各种特征以确定质量的整体水平;以及分别根据为每个搜索结果确定的整体质量水平来导出绝对量度。在示例实施例中,该绝对量度将搜索结果标识为高质量或低质量。

[0022] 在实施例中,分配部件基于分类处理的结果将每个搜索结果标记为高质量或低质量。在随后的时间点处,或在访问搜索结果时实时地,应用部件可激活与该搜索结果被如何标记相对应的适当校正动作。一般地,部分基于自动重写的搜索查询会提供新的、与该用户发布搜索查询的意图更相关(比较当前搜索结果)的搜索结果的置信度级别,来选择适当的校正动作。

[0023] 在第三个示例实施例中,提供由(在处理器上运行的)搜索引擎执行的计算机化的方法,以将搜索结果分类为高质量或低质量。首先,该方法包括提供响应于用户撰写并发布的搜索查询而返回的搜索结果集。该搜索结果集可能互相排名。在一个示例中,根据每个搜索结果与发布的搜索查询的相关性而计算该排名。该搜索引擎可然后选择排名为与发布的搜索查询高度相关的一个或多个搜索结果来进行评价。

[0024] 每个选择的搜索结果的质量水平可通过使用分类处理来确定。在示例实施例中,该分类处理包括一个或多个以下步骤:确定由要评价的选定搜索结果展示的目标特征;评价该目标特征来为每个选定搜索结果分别产生质量水平的分数;将该分数与预先定义的阈值进行比较;以及,基于该该比较,为每个选定搜索结果分配绝对量度。在示例中,当该分数比阈值高时,该绝对量度指示高质量。否则,该绝对量度指示该选定搜索结果具有低质量。在识别到选定搜索结果被分配了指示低质量的绝对量度时,自动执行在识别时重新表述所发布的搜索查询的校正动作。

[0025] 已经描述了本发明的实施例的概述,下文描述了本发明的实施例可在其中实现的示例操作环境,以便为本发明的各个方面提供一般环境。

[0026] 首先特别地参考图 1,示出了用于实现本发明实施例的示例操作环境,并一般性地将其指定为计算设备 100。计算设备 100 只是合适的计算环境的一个示例,不意图暗示关于本发明使用或功能范围的限制。计算设备 100 也不应解释为具有任何涉及示出的部件的任何一个或其组合的依赖性要求。

[0027] 本发明可在计算机代码或机器可用指令的一般环境中进行描述,机器可用指令包

括由计算机或其他机器(如个人数据助手或其他手持设备)执行的计算机可执行指令(如程序模块)。程序模块一般包括例程、程序、对象、部件、数据结构等,是指执行特定任务或实现特定抽象数据类型的代码。本发明可在各种系统配置中使用,包括手持设备、消费者电子产品、通用计算机、更专用的计算设备等。本发明也可在其中任务由通过通信网络连接的远程处理设备执行的分布式计算环境中实践。

[0028] 参考图 1, 计算设备 100 包括总线 110, 总线直接或间接地耦接以下设备: 存储器 112、一个或多个处理器 114、一个或多个呈现部件 116、输入 / 输出(I/O) 端口 118、输入 / 输出(I/O) 部件 120 和示例电源 122。总线 110 代表可能是一个或多个总线(如地址总线、数据总线或它们的组合)。虽然为了清晰起见, 图 1 的各个框图显示为带线的, 但实际上, 划界各个部件并不如此清晰, 打个比方, 这些线条更准确地将是灰色和模糊的。例如, 人可将呈现部件(如显示设备)认为是输入 / 输出部件。处理器也具有存储器。发明者意识到这是本领域的本性, 并重申图 1 中的图表只是可结合一个或多个本发明的实施例使用的示例计算设备的说明。在如“工作站”、“服务器”、“膝上型计算机”、“手持设备”等这样的类别之间不作区别, 因为它们所有被考虑在图 1 的范围内并称为“计算设备”。

[0029] 计算设备 100 典型地包括各种计算机可读介质。计算机可读介质可以是任何可得到的、可被计算设备 100 访问的介质, 它包括易失的和非易失的介质、可移除和不可移除的介质。通过示例(但不是限制), 计算机可读介质可包括计算机存储介质和通信介质。计算机存储介质包括以任何方式或技术实现的、用于存储信息(如计算机可读指令、数据结构、程序模块或其他数据)的易失和非易失的、可移除和不可移除的介质。计算机存储介质包括(但不限于)随机访问存储器 RAM、只读存储器 ROM、带电可擦可编程只读存储器 EEPROM、闪存存储器或其他存储技术、CD-ROM、数字多用途光盘(DVD)或其他光学存储设备, 盒式磁带、卡式磁带、磁盘存储设备或其他磁性存储设备或任何可用于存储想要的信息并可被计算设备 100 访问的其他介质。在实施例中, 计算机存储介质代表有形的计算机可读介质, 如体现在处理器上的介质。

[0030] 通信介质典型地以调制的数据信号(如载波或其他传送机制)体现计算机可读指令、数据结构、程序模块或其他数据, 并包括任何信息递送介质。术语“调制的数据信号”意指具有一个或多个其特征以在信号中编码信息的方式设置和改变的信号。通过示例(但不是限制), 通信介质包括有线介质(如有线网络或直接有线连接)和无线介质(如声音、RF、红外线或其他无线介质)。以上任何的组合也应该包括在计算机可读介质的范围中。在实施例中, 通信介质代表非暂时的计算机可读介质。

[0031] 存储器 112 包括易失和 / 或非易失存储器形式的计算机存储介质。该存储器可以是可移除、不可移除的或它们的组合。示例硬件设备包括固态存储器、硬盘驱动器、光盘驱动器等。计算设备 100 包括从各种实体(如存储器 112 和 / 或 I/O 部件 120)中读取数据的一个或多个处理器。呈现部件(多个)116 将数据指示呈现给用户或其他设备。典型的呈现部件包括显示设备、扬声器、打印部件、振动部件等。

[0032] I/O 端口 118 允许计算设备 100 逻辑地耦接至其他设备(包括 I/O 部件 120, 可内置它们中的一些)。示例部件包括麦克风、控制杆、游戏垫、圆盘式卫星、扫描仪、打印机、无线设备等。

[0033] 现在转向图 2, 示出了根据本发明实施例的示例分类处理 220 的各种输入 250 和

输出 230 及 240 的框图。一般地,在访问从发布的搜索查询返回的搜索结果集时,该搜索引擎使用分类处理 220 确定整个搜索结果集中的一部分的质量水平。首先,分类处理 220 可涉及将各种分类标准分别应用于一个或多个搜索结果。在示例中,短语“分类标准”广泛地涉及用来评价由可被检查的搜索结果展示的任何特征的参数。如示例,一些可被评价的、由搜索结果展示的特征是排名特征 251、质量级别特征 252、标题级别特征 253、会话级别特征 254 和 / 或聚集特征 255。在操作中,可确定这些特征中的一个或多个目标用于评价,聚集这些特征(多个)的评价来为每个搜索结果产生质量水平分数。

[0034] 虽然已经描述了被用作到分类处理 220 的输入 250 的该搜索结果的特定特征,但应该理解和领会的是,可使用用于确定质量水平的其他类型的合适输入和与搜索结果相关的特征,且本发明的实施例不限于在此描述的这 5 个不同的特征 251-255。也就是说,可使用实际上或可能有助于分类处理 220 的输出 230 及 240 的准确性的其他输入 250。例如,如果存在并且可访问用户简档(如包括为特定用户追踪的会话交互)或查询日志(如,包括为多个用户记录的社会行为),则它可以用作输入 250,以影响该搜索结果的质量水平。

[0035] 首先,排名特征 251 代表用于量度搜索查询内的关键用语和搜索结果集的内容(如文档、网页、博客之类的)之间的关系的那些特征。在示例实施例中,排名特征 251 与搜索结果依赖查询的特征有关,如搜索查询用语和搜索结果内容之间匹配的数量。为了在搜索引擎结果页面(SERP)上建立搜索结果顺序,常常采用一个或多个排名特征 251 用于将搜索结果相对于彼此排名的任务。

[0036] 在一个示例中,排名特征 251 可包括搜索查询中的用语和该搜索结果的标题、指向该搜索结果的锚文本、该搜索结果的正文和 / 或该搜索结果的关键部分(如页面顶端)之间匹配的数量 / 频率。在实施例中,未考虑该搜索查询的所有用语。搜索引擎典型地关注该搜索结果是否与该搜索查询的语言头(如关键用语、罕见用语之类)在某种方式上相匹配。因此,在实施例中,可以从搜索查询中提取语言头并采用该语言头来评价搜索结果的排名特征 251,同时为了分类目的可以忽略搜索查询中非关键、常用和自然的语言。

[0037] 在另一个示例中,排名特征 251 涉及搜索结果中匹配的关键用语的接近性。典型地,搜索结果内匹配的词出现得越靠近,搜索结果就相对于该搜索查询越相关,因此保证了高的质量水平指定。正如全文所述,通常给指定为具有高质量水平的搜索结果分配在绝对标度中的高分,这常常转化为“高质量”的绝对量度 240 中。

[0038] 质量级别特征 252 一般代表单独量化搜索结果集的整体质量而独立于搜索查询的特征。在实施例中,通过检查该搜索结果的内容 / 标题 / 锚文本和 / 或集成到搜索结果的域名来确定整体质量。在一个示例中,检查该搜索结果的内容涉及列举文档正文中的若干拼写错误,其中的大量拼写错误典型地降低了与该搜索结果相关的质量水平。在另一个示例中,检查该搜索结果的域名涉及确定该搜索结果是否导航至可信任网站。正如所看到的,质量级别特征 252 关注该搜索结果的属性而独立于搜索查询中的用语。

[0039] 在一个示例中,标题级别特征 253 代表列举该搜索查询中的关键用语与搜索结果集中的每一个特定的、搜索引擎预备的概要(即,摘要或示例回答)之间的匹配数目的特征。在另一个示例中,标题级别特征与确定是否良好地形成这些摘要(表示高质量搜索结果)、或这些摘要是否包括不完备的句子、美国信息交换标准码(ASCII)垃圾资料、大量数字 / 符号 / 统一资源定位符(URL),和 / 或非文本数据(表示低质量搜索结果)有关。

[0040] 会话级别特征 254 一般代表根据在线会话过程期间用户发起的搜索结果改进将搜索结果集中的搜索结果标记为相关或不相关的特征。在一个示例中,该改进与该用户在在线会话期间修改其该搜索查询的次数有关。经常地,该搜索查询的越多改进表示该用户对返回的搜索结果不满意,这相当于分配给响应该搜索查询而给出的搜索结果的较低质量水平。在另一个示例中,该改进与指向基于该搜索查询而产生的搜索结果的用户交互(如选择)的量有关。经常地,当几乎没有指向该搜索结果的用户发起的选择(如鼠标点击动作)时,用户常常对该搜索结果没有兴趣,从而表明该搜索结果在某种方式上不相关或不受欢迎。被发现不受欢迎的搜索结果一般被指定为具有低质量水平。

[0041] 聚集特征 255 一般代表考虑搜索结果集的内容、在搜索结果页面中发布的广告和/或组成由搜索引擎提供的连同搜索结果集的相关搜索的用语之间的变化的特征。作为一般规则,在实施例中,高质量搜索结果是呈现足够量的彼此之间的相互差异而没有过多改变的搜索结果。在一个示例中,这种相互差异表述为交叉结果熵。正如在此使用的,“交叉结果”熵指的是由 SERP 的组成表示的随机性或不一致性。在一个示例中,在搜索结果相似(如共享一个或两个域名和/或在其内容中透露共同的主题)的地方,该交叉结果熵非常低。这个低交叉结果熵可在该用户输入过于详细的搜索查询时出现,这降低了搜索结果的质量水平,因为一些相互差异对该用户有益。在另一个示例中,在搜索结果完全不同(如包含广泛的各种域名和/或由于在其内容中缺少共同的主题而妨碍搜索结果的聚集)的地方,该交叉结果熵非常高。这个高交叉结果熵可在该用户输入不够详细的搜索查询时出现,这也降低了搜索结果的质量水平,因为太多的相互差异对于该用户是混乱的。

[0042] 虽然已经描述了与搜索结果相关联的内容和域名,应理解和领会的是,可以采用搜索结果的任何其他合适的属性或 SERP 中的合适信息来计算或影响交叉结果熵,且本发明的实施例不限于以上描述的属性。例如,可在调整搜索结果的质量水平时考虑广告和/或相关搜索数据之间的熵。

[0043] 在假设的示例中,现在将关于“显示蓝钻石筹码价值多少(SHOW HOW MUCH BLUE DIAMOND CHIPS ARE WORTH)”的搜索查询来讨论特征 251-255,其中该搜索查询中该用户的意图是确定在特定赌场赌博用的“蓝钻石”扑克牌筹码的价值。首先,该搜索引擎可从该搜索查询中提取语言头,以便关键用语“蓝(blue)”、“钻石(diamond)”、“筹码(chips)”和“价值(worth)”在分类处理 220 中使用,而为了分类目的可忽略搜索查询中非关键的、常见和自然的语言(如“show how much”)。在一个示例中,排名首位的搜索结果可以是标题为“大批的散钻石碎片(Wholesale Loose Diamond Chips)”的文档。这个示例的搜索结果将用于下文的讨论中,以说明分类处理 220 的功能性应用。

[0044] 首先,可应用该示例搜索结果的排名特征 251 来度量该搜索查询中的关键用语和该示例搜索结果的内容之间的关系。在示例实施例中,在搜索查询用语“diamond”、“chips”和“worth”之间可以有很高的匹配数量,而很少与“blue”的匹配。这些匹配可具体出现在包括“Diamond Chips”的标题中,并可频繁地出现在该搜索结果的内容中。(这可能是相对于 SERP 上的其他搜索结果,示例搜索结果是排名首位的搜索结果的原因。)。结果,分类处理 220 可根据对排名特征 251 的分析,将该示例搜索结果评价为与搜索查询的用户意图相关。因此,排名特征 251 可将示例搜索结果标识为“高质量”搜索结果,且排名特征 251 的评价结果可另外地提高该示例搜索结果在绝对标度中的分数值。

[0045] 在另一个示例中,可应用该示例搜索结果的质量级别特征 252 来单独量化示例搜索结果的整体质量而独立于搜索查询。在这个假设的示例中,该示例搜索结果的整体质量通过检查该示例搜索结果的内部特征而确定。在一个示例中,可检查该示例搜索结果的内容,以列举该文档正文中的拼写错误数量。因为该示例搜索结果由钻石碎片经销商维护,所以可在该文档的正文中检测出很多拼写错误(如比起一般网页有相对更高频率的拼写错误)。因此,根据质量级别特征 252 的分析,分类处理 220 可将该示例搜索结果评价为具有低质量。因此,质量级别特征 252 可将该示例搜索结果标识为“低质量”搜索结果,且质量水平特征 252 的评价结果可另外地降低该示例搜索结果在绝对标度中的分数值。

[0046] 在另一个示例中,关于质量级别特征 252,可检查该示例搜索结果的域名,以确定该示例搜索结果是否导航至可信任的网站。再次因为该示例搜索结果由钻石碎片经销商维护,与一个大的、知名且声誉好的公司的网站相反,提供该示例搜索结果的网站的地址将可能是不可信任的网站。作为结果,根据质量级别特征 252 的分析,分类处理 220 可再次将该示例搜索结果评价为具有低质量。因此,质量级别特征 252 可将该示例搜索结果标识为“低质量”搜索结果,且可再次另外地降低该示例搜索结果在绝对标度中的分数值。

[0047] 在另一个示例中,可以应用该示例搜索结果的标题级别特征 253 来列举该搜索查询中的关键用语和搜索引擎准备好的、上面写着“钻石购买是买方注意销售…你获得了一幅可能显示了该钻石在其颜色最好时的图片(DIAMOND BUYING IS BUYER BEWARE SELLING … YOU GET A PICTURE THAT MAY SHOW THE DIAMOND IN ITS BEST COLORS)”的示例搜索结果的摘要。在这个摘要中,关键用语“diamond”匹配了两次,关键用语“blue”与摘要用语“colors”有关,且关键用语“chips”和“worth”根本没有匹配。因而,分类处理 220 可根据标题级别特征 253 的分析,将该示例搜索结果评价为具有差的摘要。

[0048] 在另一个示例中,可以应用该示例搜索结果的会话级别特征 254 来根据在线会话过程期间用户发起的搜索结果改进、将该示例搜索结果标记为相关或不相关。在一个示例中,如果在相同的在线会话期间在 SERP 中出现了该示例搜索结果,且对查看该示例搜索结果作出响应,用户改进了该搜索查询,则会话级别特征 254 表明用户对该示例搜索结果不满意,且该示例搜索结果是低质量的。在另一个示例中,如果在相同的在线会话期间在 SERP 中出现了该示例搜索结果,但它几乎没有接收到用户发起的选择(如鼠标点击动作),则会话级别特征 254 表明用户发现该示例搜索结果与用户的意图不相关。

[0049] 在最后的示例中,可以应用该示例搜索结果的聚集特征 255 来考虑该搜索结果的内容相比较该示例搜索结果之间的差异。例如,如果所有的搜索结果(包括该示例搜索结果在内)都与钻石经销商相关(在其内容中透露共同的主题),那么交叉结果熵非常低,这降低了搜索结果质量水平,这是因为一些相互差异对该用户有益。

[0050] 根据评价各种特征(如图 2 中的特征 251-255)和/或其他分类标准,分类处理 220 可预测搜索结果质量水平。在示例实施例中,该预测部分地基于每个输入 250 经评价后指示哪个评级(好或差)。当大量的输入 250 指向共同方向时,则完成该预测。例如,参考以上假设的示例,输入 250 的评价导致以下结果:表明高质量的排名特征 251;表明低质量的质量级别特征 252;表明低质量的标题级别特征 253;表明低质量的会话级别特征 254;以及表明低质量的聚集特征 255。因为根据在该示例搜索结果上的应用,五类分析特征 251-255 中的四个表明该示例搜索结果是低质量的,因此,分类处理 220 产生的预测可将该示例搜

索结果分类为低质量水平,从而可能触发校正动作。

[0051] 在产生预测后,在实施例,分类处理 220 可采用机器学习步骤。即该预测可随后与(或映射至)相似搜索结果集的人类判断的评级结果进行比较。根据比较,分类处理 220 确定该预测的准确性,并通过相对于人类判断的评级进行训练来调适评价方法。调适可涉及在汇编评级以确定预测的搜索结果质量水平时,学习如何加权各种输入 250、特征类别以及特征 251-255。换句话说,将分类处理配置为学习标识最能确定质量水平的分类标准。

[0052] 在产生质量水平之后,可基于该质量水平为正被分类处理 220 评价的搜索结果分配分数 230。该分数 230 可随后与阈值进行比较以确定该搜索结果是高质量还是低质量。在示例实施例中,分配的分数 230 代表绝对标度中的值(0 到 100 之间的数值),阈值代表折合的累计增益(DCG)阈值,它是在绝对标度中量化搜索结果质量的度量标准。

[0053] 因此,该分数与阈值的比较将搜索结果质量水平从预先定义的范围中的参数变换为绝对量度 240。在实施例中,绝对量度 240 是排除了难处理的回答范围的“是或否”或“肯定或否定”的回答。换句话说,绝对量度 240 仅限于两个不同结果(即高质量或低质量)中的一个。在操作中,绝对量度 240 的绝对值指令搜索引擎是否执行一个或多个校正动作,而在由绝对量度 240 触发后,搜索引擎采用下文讨论的置信度级别来选择一个或多个校正动作来执行。

[0054] 现在将参考图 3 讨论用于实现进行至少一个搜索结果质量的绝对分类(如图 2 中的绝对量度 240)并基于绝对分类值采取一个或多个校正动作方法的系统体系结构。首先,图 3 是说明适于在实现本发明实施例中使用的分布式计算环境 300 的框图。示例计算环境 300 包括用户设备 310、数据存储设备 335、网络服务器 350、移动设备 370 和互联这些项中的每一个的网络 380。图 3 所示的用户设备 310、数据存储设备 335、网络服务器 350 和移动设备 370 中的每个都可采用为各种类型计算设备的形式,例如上文参考图 1 描述的计算设备 100。只是举例不是限制,用户设备 310、网络服务器 350 和 / 或移动设备 370 可以是个人计算机、台式计算机、膝上型计算机、消费者电子设备、手持设备(如个人数字助理)、各种服务器、处理设备。然而应注意的是,本发明不限于在这些计算设备上实现,而是在本发明实施例的范围内可在不同类型计算设备的任何一种上实现。

[0055] 用户设备 310、网络服务器 350 和移动设备 370 中的每个典型地都包括(或链接于)某种形式的计算单元(如中央处理单元、微处理器等),以支持其上运行的部件(多个)(如查询接收部件 351、分类部件 352、分配部件 353、应用部件 354 等)的操作。正如在此使用的,短语“处理单元”一般指的是带有处理能力和存储器的专用计算设备,其支持软件、应用、计算机程序在其上执行的基础的操作软件。在一个示例中,为了使每个设备能执行通信相关的处理和其他操作(如执行图 2 中的分类处理 220),计算单元配置有分别集成到或可操作地耦接至用户设备 310、网络服务器 350 和 / 或移动设备 370 的有形硬件元件或机器。在另一个示例中,该计算单元可包括耦接至由用户设备 310、网络服务器 350 和 / 或移动设备 370 中的每个容纳的计算机可读介质的处理器(多个,未显示)。

[0056] 通常,计算机可读介质包括至少暂时地存储多个可由该处理器执行的计算机软件部件的物理存储器。如在此使用的,术语“处理器”不意图是限定性的,并可包括以计算能力行动的计算机单元中的任何元件。在这个能力中,该处理器可被配置为处理指令的有形物体。在示例实施例中,处理可涉及读取、解码 / 解释、执行和回写指令。

[0057] 除了处理指令之外,该处理器还可往返于布置在用户设备 310、网络服务器 350 和移动设备 370 上或集成到其中的其他资源来传递信息。资源一般指使用户设备 310、网络服务器 350 和移动设备 370 能执行特定功能的软件部件或硬件机构。只是举例来说,由网络服务器 350 容纳的资源(多个)进行操作以辅助搜索引擎 345 从用户设备 310 或移动设备 370 的用户处接收输入,和 / 或响应于输入而提供合适的通信(如,呈现搜索结果 325)。

[0058] 用户设备 310 可包括输入设备(未显示)和呈现设备 315。一般,提供输入设备以接收特别地影响由搜索引擎 345 呈现并且在用户界面(UI)显示器 320 处表面显示的搜索结果 325 的输入(多个)。说明性的输入设备包括图 1 中的鼠标、控制杆、小键盘、麦克风、I/O 部件 120 或能够接收用户输入并将该输入的指示传输给用户设备 310 的其他任何部件。只是举例而言,输入设备便于搜索查询 375 的输入,其通过网络 380 传输至网络服务器 350 以便由搜索引擎 345 处理。

[0059] 在实施例中,呈现设备 315 被配置为呈现和 / 或在其上给出 UI 显示 320。可与用户设备 310 的输出操作耦接的呈现设备 315 可被配置为能将信息呈现给用户的任何呈现部件,如数字监视器、电子显示面板、触摸屏、模拟机顶盒、等离子显示屏、扬声器、盲文面板等。在一个示例实施例中,呈现设备 315 配置成呈现丰富内容,如搜索结果 325、相关广告和数字图像和视频。在另一个示例实施例中,呈现设备 315 能够呈现其他形式的介质(即音频信号)。

[0060] 见图 2,数据存储设备 335 通常被配置为存储与评价搜索结果质量水平、确定分数 230 和基于以上导出绝对量度 240 相关联的信息 395。在操作中,当用户正在搜索引擎 345 处实施搜索时,搜索引擎 345 可访问驻留在数据存储设备 335 处的信息 395,如该用户的简档(捕捉该用户的历史搜索动作)、一般社会行为的数据记录(捕捉多个用户的历史在线动作)和 / 或一个或多个到分类处理 220 的输入 250 (见图 2)。

[0061] 在其他实施例中,可将数据存储设备 335 配置为可被搜索,以便适当地访问由此维护的信息。例如,针对分类处理 220 的输入 250 和人类判断评级可搜索数据存储设备 335 来训练分类处理 220。本领域普通技术人员要理解和领会的是,数据存储设备 335 中存储的信息可以是可配置的,并可包括与搜索结果质量水平的确定和 / 或检索、以及对从中触发的校正动作的选择相关的任何信息。这样的用户交互的信息内容和量不意图在任何方式上限制本发明实施例的范围。进一步地,虽然作为单个和独立的部件而示出,数据存储设备 335 实际上可以是多个数据库,例如数据库集群,其部分可驻留在客户端设备 310、网络服务器 350、移动设备 370、其他外部计算设备(未显示)和 / 或它们任何的组合上。

[0062] 这个分布式计算环境 300 只是可实现来执行本发明的方面的合适环境的一个实例,而并不意图暗示关于本发明的使用或功能范围的任何限制。示出的分布式计算环境 300 也不应解释为具有涉及所示的设备 310、350 和 370、数据存储设备 335 和部件 351-354 中的任何一个或者组合的任何依赖或要求。在一些实施例中,部件 351-354 的一个或多个可作为独立设备而实现。在其他实施例中,部件 351-354 的一个或多个可直接集成至网络服务器 350 中,或集成到互相连接以形成网络服务器 350 的分布式节点上。要领会和理解的是,部件 351-354 (图 3 所示)本质上和数量上是示例性的而不应解释为限制。

[0063] 因此,可以采用任何数量的部件来实现在本发明的实施例范围中的想要的功能。虽然为了清晰起见而利用线示出了图 3 的各种部件,然后划分各个部件并不如此清晰,而

且象征性地,这些线条更准确地会是灰色的或模糊的。进一步地,虽然图 3 的一些部件描绘成单个块,这些描述在本质上和数量上是示例性的而不被解释为限制(如,虽然只显示了一个呈现设备 315,但是可以有更多的设备通信地耦接至客户端设备 310)。

[0064] 进一步地,示例系统结构中的设备可由相关领域已知的任何方法相互连接。例如,用户设备 310、网络服务器 350 和移动设备 370 可通过包括多个通过一个或多个网络(未显示)彼此耦接的计算设备的分布式计算环境而操作地耦接。在实施例中,该网络可包括(但不限于)一个或多个局域网(LAN)和/或广域网(WAN)。这种联网环境常见于办公室、企业范围的计算机网络、企业内部网和互联网。因此,在此不再进一步描述该网络。

[0065] 在操作中,部件 351-354 被设计成执行以下处理,该处理至少包括逐个地自动确定搜索结果的绝对量度(高质量或低质量),且采取合适的校正动作。应注意的是,将共同参照图 3 和 4,描述搜索引擎 345 的操作。

[0066] 首先,搜索引擎 345 包括查询接收部件 351,其被配置为接收由用户发布的搜索查询 375,并返回部分基于该搜索查询 375 而产生的搜索结果集 325。以这种方式,查询接收部件 351 充当图 4 中与用户 410 接口的前端体验 420,而顶级聚集器 440 提供后端体验。在实施例中,该后端体验涉及顶级聚集器 440 将已处理的查询 411 发布至网络索引 450 并从中检索搜索结果 325。

[0067] 继续参考图 4,将应答顶级聚集器 430 配置为通过将搜索查询 375 发送至组合的修改服务/拼写器 435 来处理搜索查询 375。在应答顶级聚集器 430 接收已处理的查询 411 时(如使拼写错误得到修正及附加了其他同义词/逻辑修改),已处理的查询 411 用于产生搜索结果 325。在接收搜索结果 325 时,顶级聚集器 440 执行图 2 中的分类处理 220,以产生回传至应答顶级聚集器 430 的输出 416(如图 2 中的绝对量度 240)。

[0068] 现在转向在图 3 的计算环境 300 内的系统体系结构,提供类似图 4 的顶级聚集器 440 的分类部件 352 或者由顶级聚集器 440 提供分类部件 352。也就是说,分类部件 352 被配置为实现自动地确定搜索结果集 325 中的一个或多个搜索结果的质量度量的分类处理。如上面参考图 2 详细描述的那样,采用分类处理来自动地确定就绝对意义来说搜索结果 325 是否是可接受的(即,满足搜索查询 375 的用户意图)。典型地,该分类处理包括评价由搜索结果(多个)325 呈现的各种特征来确定整体质量水平、并且根据为每个搜索结果(多个)325 确定的整体质量水平导出绝对度量的步骤。通常,绝对度量以互相排斥的方式将搜索结果(多个)325 分别标识为要么高质量要么低质量。

[0069] 在示例实施例中,分配部件 353 被配置为基于分类处理的结果将每个搜索结果(多个)325 标记为“高质量”或者“低质量”。将这些被标记的搜索结果传递到应用部件 354,其在当一个或多个搜索结果(例如,X 数目个最高排名的搜索结果或者在 SERP 上的大多数搜索结果)被标记为“低质量”时,激活适当的校正动作。

[0070] 在示例实施例中,作为置信度级别的函数来选择适当的校正动作。作为此处的使用,“置信度级别”不意指在定义上的限制,而是涵盖涉及下述可能性的任何度量,该可能性为由应用部件 354 提供的重写搜索查询内的任何替换用语会产生相对于当前“低质量”搜索结果 325 有改进的新的搜索结果。简明而言,在实施例中,置信度级别指示搜索引擎编写的搜索查询将产生提高的搜索结果的可能性。在应用中,置信度级别用于选择一个或多个校正动作,诸如呈现重新编制用户接口(UI)340、经由通知消息警告用户搜索结果(多个)

325 被标记为“低质量”、避免向用户呈现低质量搜索结果、或者自动地重写该搜索查询 375 以产生搜索结果的更新集。

[0071] 在示例实施例中,在计算到由应用部件 354 建议的替换用语的高置信度级别时,自动地重写和自动地发布搜索查询 375。相反,在计算到由应用部件 354 建议的替换用语的低置信度级别时,呈现重新编制 UI 340。在一种情况下,重新编制 UI 340 请求和接受用户发起的交互 385 (诸如鼠标点击选择),以便引导改变搜索查询内的关键用语。举例来说,改变关键用语可能涉及替代或者删除搜索查询 375 内的一个或多个关键用语。

[0072] 虽然已经描述了在确定一个或多个搜索结果被标记为“低质量”时增强搜索结果 325 的各种不同的方法,但是应当理解和领会的是,可以使用其它类型的适当方法和接口用于选择和实现校正动作,而且本发明的实施例不局限于上述校正动作。例如,附加校正动作可涉及给予第三方机构使用该第三方机构的经验和资源来重新编制搜索查询 375 的机会。

[0073] 现在转到图 5,示出了描述根据本发明的实施例、采用搜索结果集的质量水平来激活一个或多个校正动作的示例方法 500 的流程图。首先,要领会和理解的是,除非且除了个别步骤顺序被明确描述时之外,否则虽然在此可能使用词条“步骤”和 / 或“框”以暗示采用的不同方法单元,但这些术语不应解释为暗示在此公开的各个步骤之中或之间的任何特定顺序。

[0074] 在示例实施例中,方法 500 涉及访问搜索结果集(见框 502)并通过向其应用多个分类标准而单独地确定该搜索结果集的质量水平(见框 504)。如上文更充分的描述,该分类标准包括评价由该搜索结果单独表达的排名特征、质量级别特征、标题级别特征、会话级别特征或聚集特征中的至少一个。如框 506 所示,该方法可通过为该搜索结果集中的至少一个结果分配在绝对标度中的分数而继续。该分数典型地基于确定的质量水平。如框 508 所示,该分数可与在绝对标度中的阈值进行比较。因此,可基于该比较的结果选择并执行一个或多个校正动作。这如框 510 所示

[0075] 参考图 6,示出了描述根据本发明的实施例的将搜索结果分类为高质量或低质量的示例方法 600 的流程图。首先,提供方法 600,它可代表由搜索引擎(在处理器上运行的)执行的用于将搜索结果分类为高质量或低质量的计算机化的方法。在实施例中,如框 602 所示,方法 600 可涉及提供响应于由用户撰写并发布的搜索查询而返回的搜索结果集的步骤。如框 604 所示,该搜索结果集可相对于彼此排名。在一个示例中,作为每个搜索结果与发布的搜索查询的相关性的函数来计算该排名。如框 606 所示,搜索引擎随后可选择一个或多个被排名为与发布的搜索查询高度相关的搜索结果用于评价。

[0076] 如框 608 所示,可通过采用分类处理来确定每个选择的搜索结果的质量水平。在示例实施例中,该分类处理包括一个或多个以下步骤:确定由待评价的、选定搜索结果展示的目标特征(见框 610);评价该目标特征来为每个选择的搜索结果单独产生质量水平的分数(见框 612);将该分数与预先定义的阈值进行(见框 614);以及,基于该比较为每个选择的搜索结果分配绝对量度(见框 616)。在示例中,当该分数比阈值高时,绝对量度指示高质量。否则,该绝对量度指示所选择的搜索结果低质量。如框 618 所示,在识别出所选择的搜索结果被分配了指示低质量的绝对量度后,在识别后自动执行重新编制所发布的搜索查询的校正动作。

[0077] 如上所述,该校正动作可涉及(如当自动撰写的搜索查询中的置信度高时)搜索引

擎隐含地重写搜索查询。或者该校正动作可涉及在重新编制 UI 中明确地为用户提供预置的建议,其征求用户修改搜索查询的关键用语的反馈(如当自动撰写的搜索查询中的置信度低时)。要领会和理解的是,可实现其他校正动作以辅助用户有效地到达与他或她的用户意图相关的搜索结果。

[0078] 已经说明性而不是限制性地描述了本发明的各种实施例。时不时地,替换实施例会变得明显而不脱离本发明实施例。要理解的是,某些特征和子组合很实用,并且可以被采用而不用参考其他特征和子组合。这是本权利要求所预期的且在本权利要求的范围内。

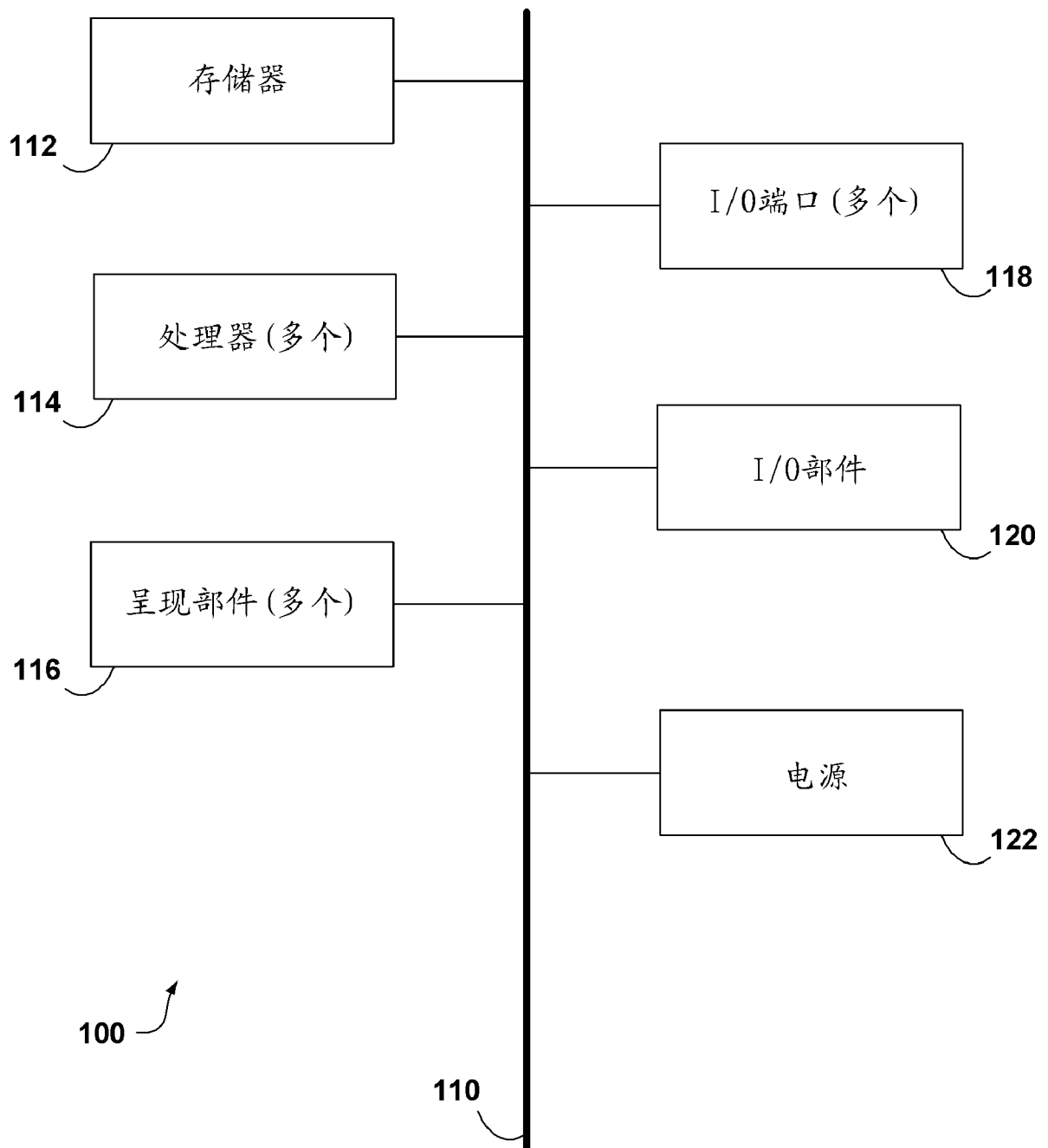


图 1

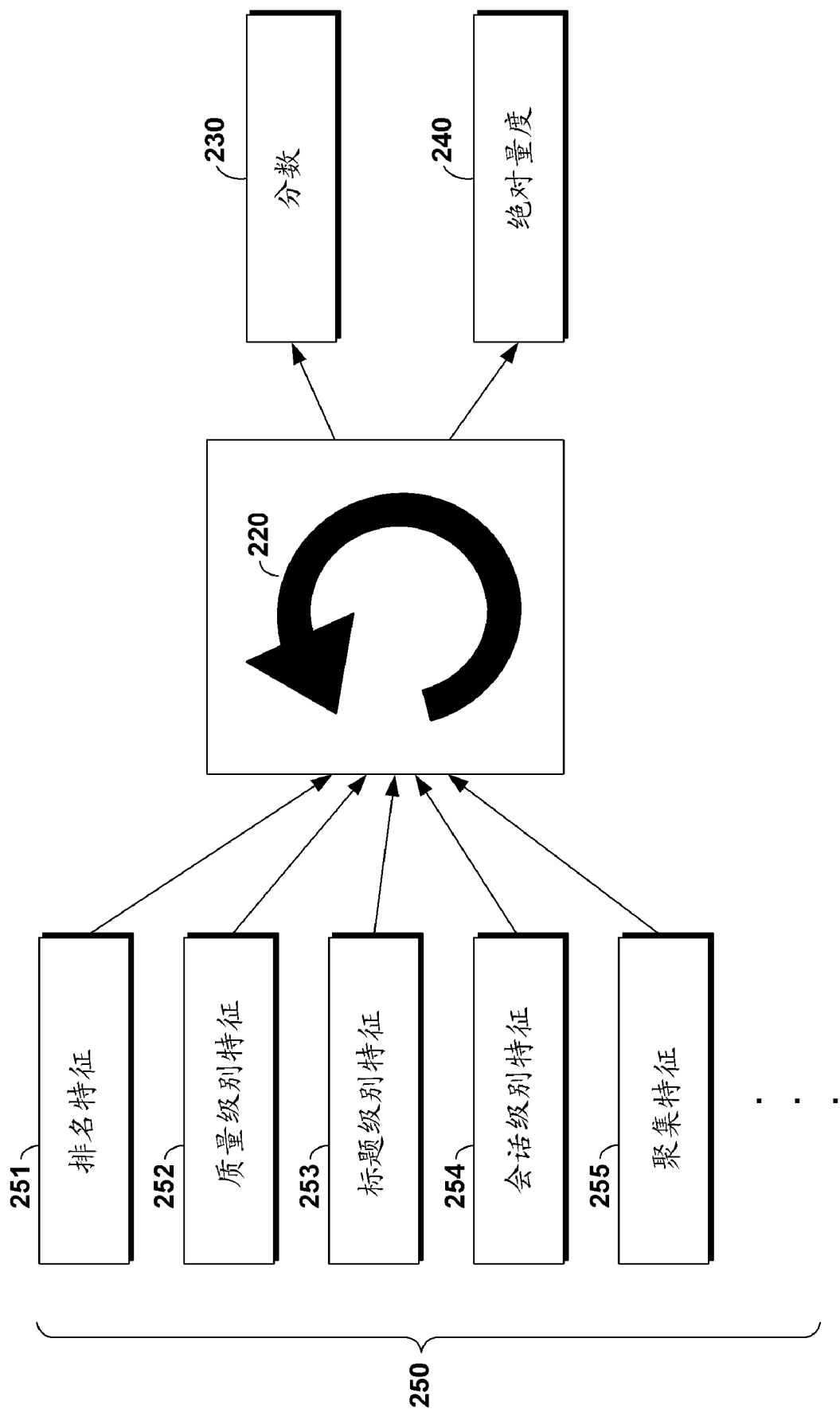


图 2

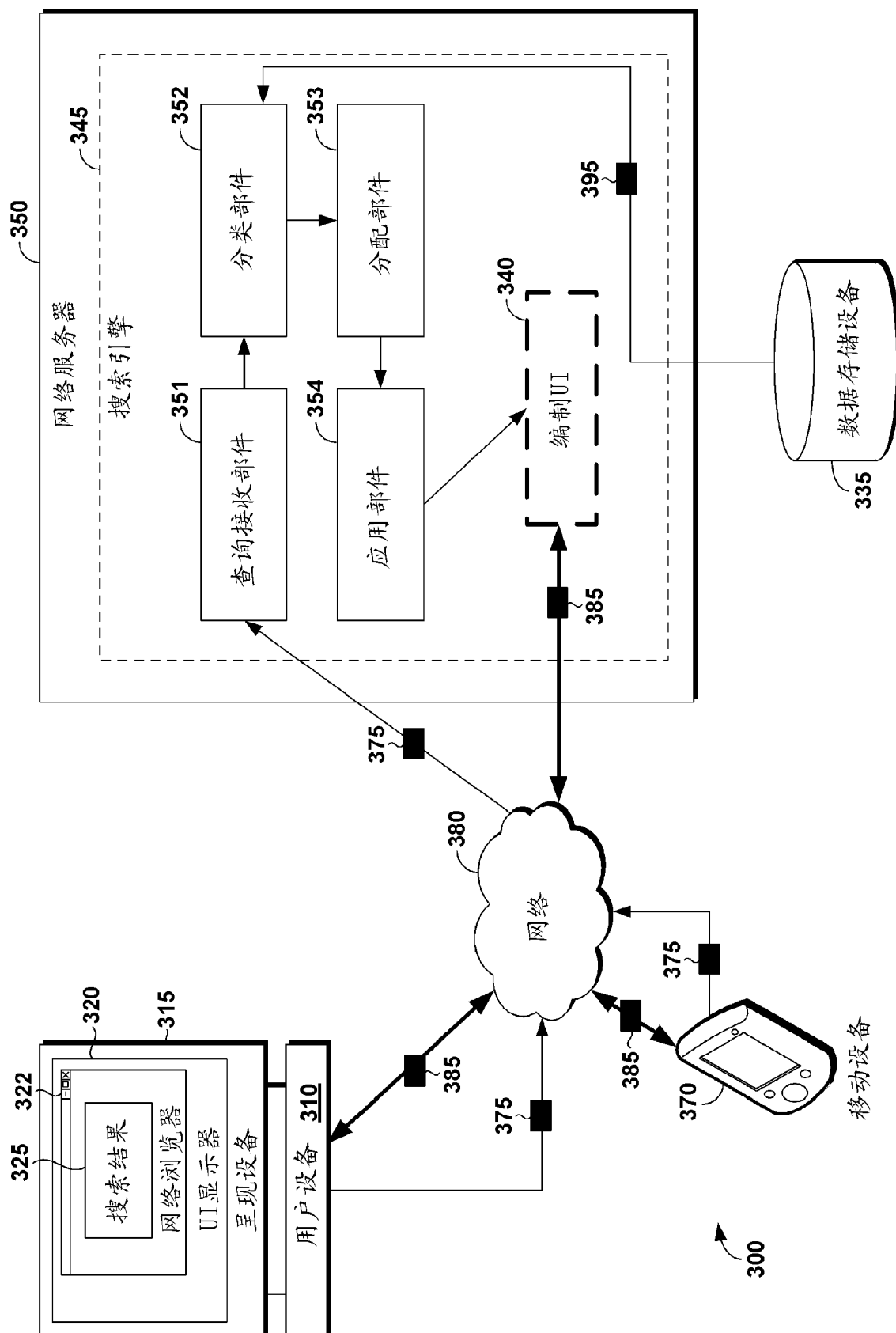


图 3

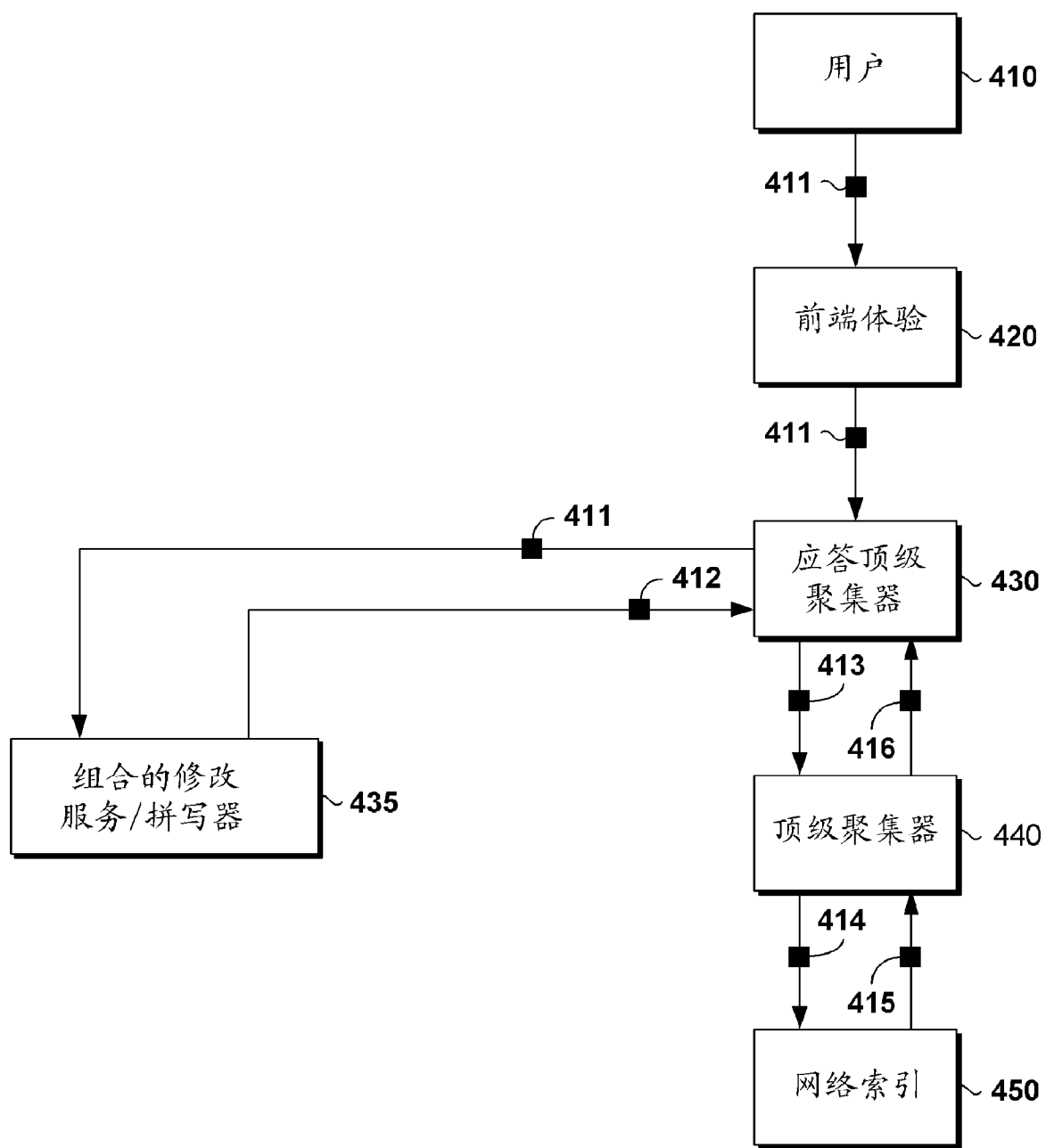


图 4

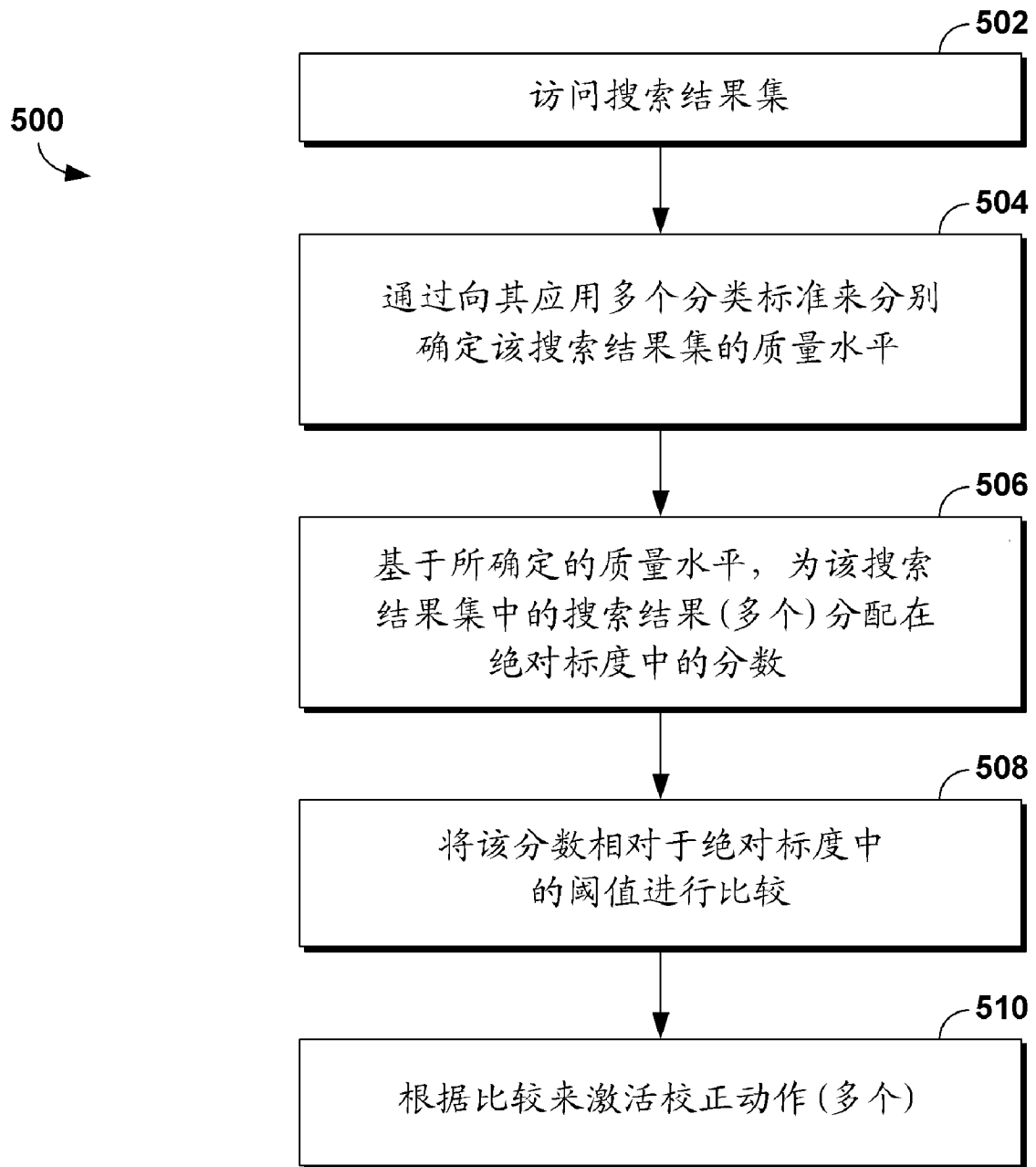


图 5

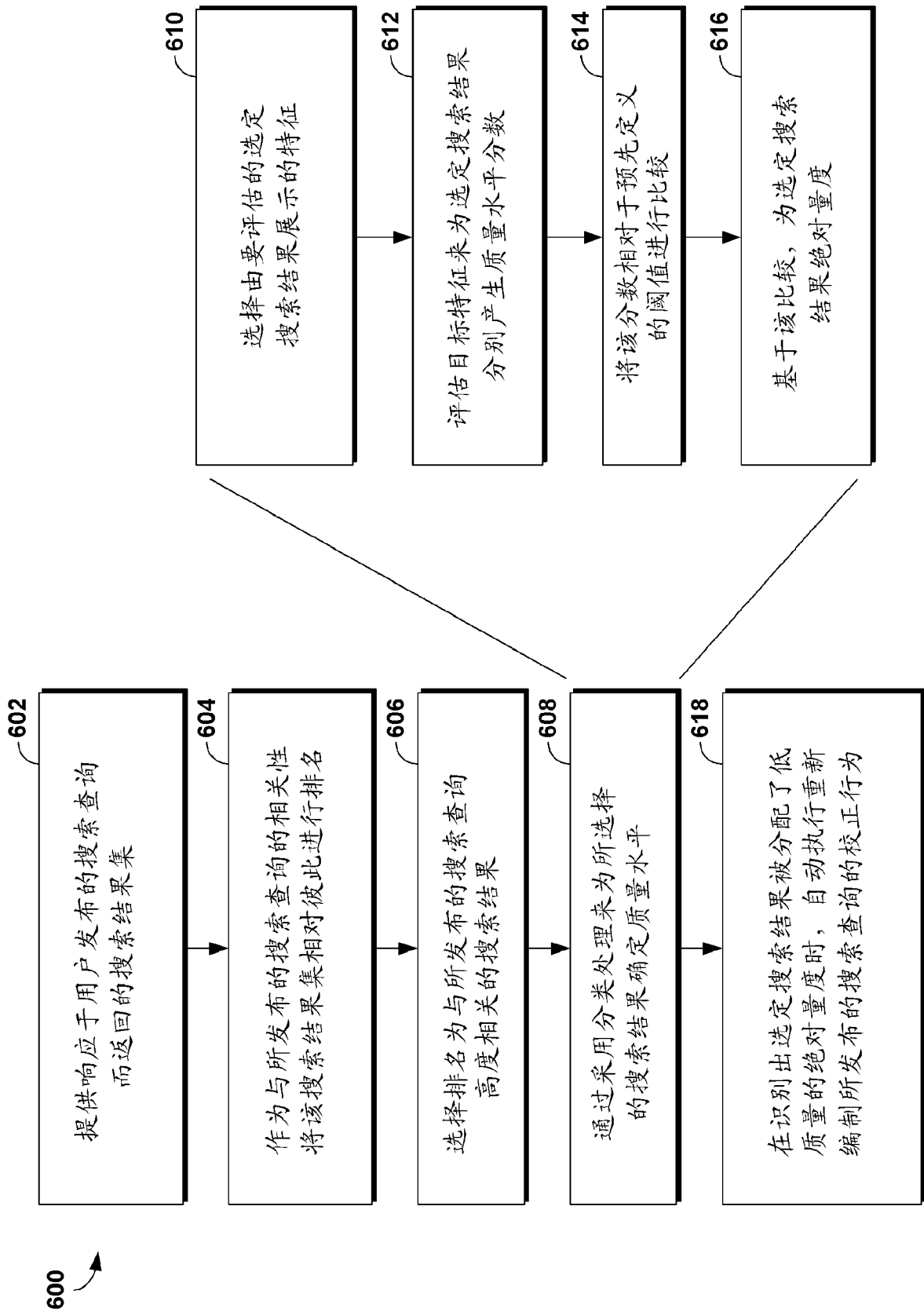


图 6