



(12)发明专利

(10)授权公告号 CN 102105879 B

(45)授权公告日 2017.03.29

(21)申请号 200980128749.1

(72)发明人 基思·A·马洛

(22)申请日 2009.07.23

(74)专利代理机构 北京东方亿思知识产权代理  
有限责任公司 11258

(65)同一申请的已公布的文献号

代理人 李晓冬

申请公布号 CN 102105879 A

(51)Int.Cl.

G06F 17/30(2006.01)

(43)申请公布日 2011.06.22

(56)对比文件

(30)优先权数据

US 6944612 B2, 2005.09.13, 第1栏第22–  
24、51–59行, 第5栏第1–32行, 第4栏46–67行, 2栏  
第20–43行, 第9栏40–56行, 图1.

12/181,268 2008.07.28 US

US 2005/0149496 A1, 2005.07.07, 全文.

(85)PCT国际申请进入国家阶段日

CN 1761962 A, 2006.04.19, 全文.

2011.01.21

审查员 李玉坤

(86)PCT国际申请的申请数据

PCT/US2009/004308 2009.07.23

(87)PCT国际申请的公布数据

W02010/014185 EN 2010.02.04

(73)专利权人 埃克斯凯利博IP有限责任公司

权利要求书4页 说明书10页 附图11页

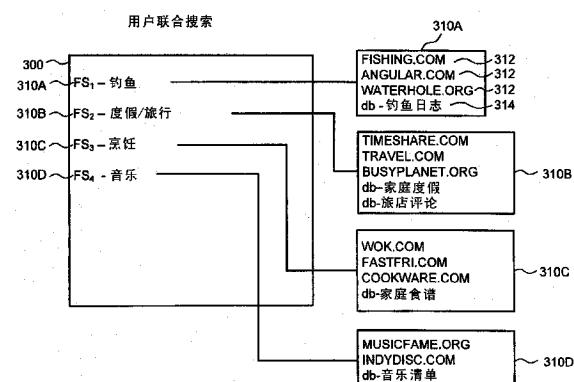
地址 美国加利福尼亚州

(54)发明名称

联合团体搜索

(57)摘要

提供一种用于联合搜索的方法、物品和设备。联合搜索可以由其创建者限定来在预先限定的一组数据源中搜索数据。相比于传统的基于爬虫的搜索，提高了搜索结果的准确度和关联度。



1.一种联合搜索方法,包括:

接收识别的由一个或多个第一用户计算平台提交或上载的多个联合搜索源,所述多个联合搜索源包括至少一个第一联合搜索源和至少一个第二联合搜索源,所述多个联合搜索源中各个联合搜索源包括到多个不同指定数据源的链接;

索引所述多个联合搜索源以确定索引标准,所述索引包括确定所述至少一个第一联合搜索源以及所述至少一个第二联合搜索源中与一个或多个第一关键词和一个或多个第二关键词中的各个关键词有关的至少一个联合搜索源;

在执行所述索引之后,接收来自第二用户计算平台的查询,并且将所接收到的查询与所述索引标准进行匹配以确定所述多个联合搜索源中与所述查询相关的一个或多个联合搜索源并对这一个或多个联合搜索源进行排名;以及

经由通信网络向所述第二用户计算平台发送包括结果的电信号,所述结果包括到和所述查询相关的所述一个或多个被排名的联合搜索源的各个链接以及针对一个或多个被排名的联合搜索源中的各个联合搜索源的有用性度量的至少一个可显示图形指示符。

2.根据权利要求1所述的方法,还包括:

确定与被索引的联合搜索源相关联的签名;

将查询和所述签名进行比较;以及

确定所述查询与和被索引的联合搜索相关联的签名之间的最接近匹配。

3.根据权利要求2所述的方法,还包括:

至少部分地基于被索引的联合搜索源来提供搜索结果,其中所述查询与和该被索引的联合搜索源相关联的签名之间为最接近匹配。

4.根据权利要求3所述的方法,还包括:

基于基于爬虫的搜索来提供搜索结果。

5.根据权利要求2所述的方法,其中,所述索引至少部分地基于由所述联合搜索的创建者所提供的关键字;并且其中,所述比较至少部分地基于由被索引的联合搜索的创建者所提供的关键字。

6.根据权利要求2所述的方法,其中,索引至少部分地基于将联合搜索中的域与由该查询定义的域相匹配;并且

其中,所述比较至少部分地基于将联合搜索的域与通过该查询限定的域相匹配。

7.根据权利要求2所述的方法,其中,所述索引至少部分地基于将被索引的联合搜索的域与基于分类路径返回结果的搜索引擎的域相匹配,其中被索引的联合搜索结果的分类路径被与和基于爬虫的搜索相关联的分类路径相匹配。

8.根据权利要求1所述的方法,其中,索引至少一组搜索结果包括:

从与联合搜索相关联的站点获取数据;

从所述数据中提取至少包括关键字和域的加权因子;以及

基于所述加权因子来创建签名的表格;并且

其中将查询与所述至少一个被索引的联合搜索进行比较包括将该查询与和所述联合搜索相关联的签名相比较;并且

其中,确定该查询与被索引的联合搜索源之间的最接近匹配包括:

确定该查询与签名之间的最接近匹配,并且还包括:

至少部分地基于被索引的联合搜索来提供搜索结果,其中该查询和与该被索引的联合搜索相关联的签名之间为最接近匹配。

9.一种用于联合搜索的设备,包括:

用于接收从一个或多个第一用户计算平台接收的多个联合搜索源的装置,所述多个联合搜索源包括至少一个第一联合搜索源和至少一个第二联合搜索源,所述多个联合搜索源中各个联合搜索源包括到多个不同指定数据源的链接;

用于索引所述多个联合搜索源以确定索引标准的装置,所述索引包括确定所述至少一个第一联合搜索源以及所述至少一个第二联合搜索源中与一个或多个第一关键词和一个或多个第二关键词中的各个关键词有关的至少一个联合搜索源;

用于在执行所述索引之后,接收来自第二用户计算平台的查询,并且将所接收到的查询与所述索引标准进行匹配以确定所述多个联合搜索源中与所述查询相关的一个或多个联合搜索源并对这一个或多个联合搜索源进行排名的装置;以及

用于经由通信网络向所述第二用户计算平台发送包括结果的电信号的装置,所述结果包括到和所述查询相关的所述一个或多个被排名的联合搜索源的各个链接以及针对一个或多个被排名的联合搜索源中的各个联合搜索源的有用性度量的至少一个可显示图形指示符。

10.根据权利要求9所述的设备,还包括:

用于确定与被索引的联合搜索源相关联的签名的装置;

用于将查询和所述签名进行比较的装置;以及

用于确定所述查询与和被索引的联合搜索源相关联的签名之间的最接近匹配的装置。

11.根据权利要求10所述的设备,还包括:

用于至少部分地基于被索引的联合搜索源来提供搜索结果的装置,其中所述查询与和该被索引的联合搜索源相关联的签名之间为最接近匹配。

12.根据权利要求11所述的设备,还包括:

用于基于基于爬虫的搜索来提供搜索结果的装置。

13.根据权利要求10所述的设备,其中,所述索引至少部分地基于由所述联合搜索源的创建者所提供的关键字;并且其中,所述比较至少部分地基于由被索引的联合搜索源的创建者所提供的关键字。

14.根据权利要求10所述的设备,其中,索引至少部分地基于将联合搜索源中的域与由该查询定义的域相匹配;并且

其中,所述比较至少部分地基于将联合搜索源的域与通过该查询限定的域相匹配。

15.根据权利要求10所述的设备,其中,所述索引至少部分地基于将被索引的联合搜索源的域与基于分类路径返回结果的搜索引擎的域相匹配,其中被索引的联合搜索结果的分类路径被与和基于爬虫的搜索相关联的分类路径相匹配。

16.根据权利要求9所述的设备,其中,用于索引至少一组搜索结果的装置包括:

用于从与联合搜索源相关联的站点获取数据的装置;

用于从所述数据中提取至少包括关键字和域的加权因子的装置;以及

用于基于所述加权因子来创建签名的表格的装置;并且

其中将查询与被索引的联合搜索源的各个联合搜索源进行比较包括将该查询与和所

述联合搜索源的各个联合搜索源相关联的签名相比较;并且

其中,确定该查询与被索引的联合搜索源之间的最接近匹配包括:

确定该查询与签名之间的最接近匹配,并且还包括:

至少部分地基于被索引的联合搜索源来提供搜索结果,其中该查询和与该被索引的联合搜索源相关联的签名之间为最接近匹配。

17.一种用于联合搜索的设备,包括:

计算平台,所述计算平台适于:

处理从一个或多个第一用户计算平台接收的多个接收到的联合搜索源,所述多个联合搜索源包括至少一个第一联合搜索源和至少一个第二联合搜索源,所述多个联合搜索源中的各个联合搜索源包括到多个不同指定数据源的链接;

索引所述多个联合搜索源以确定索引标准,所述索引包括确定所述至少一个第一联合搜索源以及所述至少一个第二联合搜索源中与一个或多个第一关键词以及一个或多个第二关键词中的各个关键词有关的至少一个联合搜索源;

在执行所述索引之后,处理来自第二用户计算平台的查询,并且将所接收到的查询与所述索引标准进行匹配以确定所述多个联合搜索源中与所述查询相关的一个或多个联合搜索源并对这一个或多个联合搜索源进行排名;以及

经由通信网络向所述第二用户计算平台发送包括结果的电信号,所述结果包括到和所述查询相关的所述一个或多个被排名的联合搜索源的各个链接以及针对一个或多个被排名的联合搜索源中的各个联合搜索的有用性度量的至少一个可显示图形指示符。

18.根据权利要求17所述的设备,所述计算平台还:

确定与具体被索引的联合搜索源相关联的签名;

将所述查询与所述签名相比较;以及

确定该查询与和被索引的联合搜索源相关联的所述签名之间的最接近匹配。

19.根据权利要求18所述的设备,所述计算平台还:

至少部分地基于被索引的联合搜索源来呈现搜索结果,其中该查询与和该被索引的联合搜索源相关联的签名之间为最接近匹配。

20.根据权利要求19所述的设备,所述计算平台还:

基于基于爬虫的搜索来呈现所述搜索结果。

21.根据权利要求18所述的设备,其中,所述计算平台还至少部分地基于由所述联合搜索源的创建者所提供的关键字来索引;并且

其中,所述比较是至少部分地基于由被索引的联合搜索源的创建者所提供的关键字被确定的。

22.一种用于联合搜索的设备,包括:

搜索引擎(1006),所述搜索引擎(1006)通过前端服务器(1004)与客户端通信,所述搜索引擎(1006)包括联合搜索引擎(1010),所述联合搜索引擎(1010)包括一个或多个联合搜索源的数据库,所述一个或多个联合搜索源基于至少一个或多个关键词被索引来索引所述联合搜索源,所述联合搜索引擎(1010)通过以下步骤,至少部分地基于在所述客户端处发起的查询来确定多个联合搜索源:

处理一个或多个接收到的联合搜索源,所述一个或多个接收到的联合搜索源包括所述

多个联合搜索源中的至少一个第一联合搜索源和至少一个第二联合搜索源,所述多个联合搜索源中的各个联合搜索源包括到多个不同指定数据源的链接;

在所述索引执行之后,处理从用户计算平台接收到的查询,并且对所接收到的查询与索引标准进行匹配以确定与所述查询相关的一个或多个联合搜索并对这一个或多个联合搜索源进行排名;

经由通信网络向所述用户计算平台发送包括结果的电信号,所述结果包括到和所述查询相关的所述一个或多个被排名的联合搜索源的各个链接以及针对一个或多个被排名的联合搜索源中的各个联合搜索源的有用性度量的至少一个可显示图形指示符。

23.根据权利要求22所述的设备,其中,所述被索引的搜索与签名相关联,其中,将所述查询与所述签名中的一个或多个签名相比较,并且其中,所述查询与被索引的联合搜索源之间的所述匹配包括确定所述查询与和所述被索引的联合搜索源相关联的签名之间的最接近匹配。

## 联合团体搜索

[0001] 相关申请

[0002] 本申请要求2008年7月28日提交的美国正式专利申请第12/181,268号的优先权，其全部内容通过引用被结合于此。

### 背景技术

[0003] 1. 技术领域：

[0004] 在此公开的主题涉及网络搜索协议。

[0005] 2. 信息

[0006] 在诸如Google之类的传统互联网搜索引擎中，仅仅是通过搜索引擎的爬虫(crawler)技术已搜索和索引(index)的数据源可被用户搜索、获取和访问。当用户查询传统的搜索引擎时，用户通常不是搜索整个互联网，而实际上是搜索搜索引擎的索引来找到查询的匹配并随后访问数据源。数据源的示例包括网页和数据库，但是不限于此。从这样的所谓基于爬虫的搜索得到的搜索结果有很多限制。例如，依赖于搜索引擎的爬虫技术用来更新具体网站的索引的时间表和时间，搜索结果可能明显过期。此外，大量数据被包含在由于爬虫技术的限制和访问约束而不能被传统互联网搜索引擎访问的数据库和网页中。

[0007] 联合搜索(federated search)与传统的基于爬虫的搜索不同。联合搜索可以由其创建者限定以在预定一组数据源中搜索数据。联合搜索允许用户利用来自单个用户界面的单个查询来搜索多个数据源。用户可以在界面的搜索框中输入搜索查询并且该查询可以被发送给联合搜索中所限定的每个数据源。由于联合搜索通常包括作为目标的一组数据源，所以对于给定查询，可以提高各个搜索的准确度和关联度。此外，相比由基于爬虫的搜索引擎提供的那些搜索，可以减少搜索目标数据源所需要的时间量。所产生的搜索结果的关联度通常依赖于具体的联合搜索中所包括的数据源的关联度。另外，联合搜索通常固有地作为各个数据源的当前数据，这是因为它们是被实时地搜索的。

### 附图说明

[0008] 将参考以下附图来描述非限制性和非穷尽性实施例，其中，若非另外指明，在各个附图中，类似的标号表示类似的部分。

[0009] 图1是能够通过网络在客户端和一个或多个服务器之间建立会话的系统的示意图；

[0010] 图2图示出根据一个实施例如显示装置上所显示的网络浏览器；

[0011] 图3是图示出根据一个实施例的联合搜索库的示例的示图；

[0012] 图4是用于处理联合搜索的方法的实施例的流程图；

[0013] 图5是提供用于处理联合搜索的方法的实施例的可能的图形表示的实施例的示图；

[0014] 图6是用于处理联合搜索的方法的实施例的流程图；

[0015] 图7是提供用于处理联合搜索的方法的实施例的可能的图形表示的实施例的示

图；

- [0016] 图8是用于处理联合搜索的方法的实施例的流程图；
- [0017] 图9是提供用于处理联合搜索的方法的实施例的图形表示的实施例的示图；
- [0018] 图10是提供用于处理联合搜索的系统的实施例的图形表示的示图；以及
- [0019] 图11是图示出可以实现实施例的计算机系统的框图。

### 具体实施方式

[0020] 贯穿本说明书，对“一个实施例”或“实施例”的提及是指与该实施例关联地描述的具体特征、结构或特点被包括在要求保护的主题的至少一个实施例中。因此，贯穿本说明书的各个地方出现的短语“在一个实施例中”或“实施例”不一定都是指同一实施例。此外，具体的特征、结构或特点可以在一个或多个实施例中被组合。

[0021] 除非以其他方式特别指明，从以下讨论中显见，可以理解，贯穿本说明书，利用例如诸如“处理”、“计算”、“选择”、“形成”、“使能”、“抑制”、“识别”、“开始”、“获取”、“托管(host)”、“维护”、“呈现”、“修改”、“接收”、“发送”、“建立”、“确定”等的术语进行的讨论是指可以通过计算平台执行的动作和/或处理，所述计算平台例如是计算机或类似的电子计算装置，其操纵和/或变换在计算平台的处理器、存储器、寄存器和/或其他信息存储、发送、接收和/或显示装置内被表示为物理电子量和/或磁量和/或其他物理量的数据。这样的动作和/或处理可以在被存储在存储介质中的机器可读指令的控制下由计算平台执行。此外，除非以其他方式特别指明，在此参考流程图或其他方式描述的处理也可以全部或部分由这样的计算平台执行和/或控制。

[0022] 图1是能够通过网络104在客户端104和一个或多个服务器106之间建立会话的系统100的示意图。网络104可以采用包括一个或多个数据传输介质的数据链路(未示出)，例如线缆(例如，光纤、双绞线、同轴线)、卫星传输链路或地面无线传输链路。然而，这些仅仅是可用于在网络中传输信息的数据传输介质的示例，并且所要求保护的主题在这方面不受限制。另外，网络104可以采用用于在源和目的地之间传输信息的一个或多个通信协议，例如，互联网协议(IP)、异步传输模式(ATM)和/或适于在无线传输链路上使用的通信协议。然而，这些仅是可以用于在网络中的源和目的地之间传输信息的通信协议的示例，并且所要求保护的主题在这些方面不受限制。在具体实施例中，网络104可以包括一个或多种网络技术和/或配置，例如举几个例子，例如一个或多个局域网、一个或多个企业内部网、互联网、与企业内部网相对的无线载波基础架构互联网和/或无线载波基础架构。尽管在此提供的具体示例在客户端和服务器之间建立会话，但是应当理解，这样的实现方式仅仅是为了说明的目的而提供的示例，并且所要求保护的主题不限于这样的实现方式。例如，也可以通过网络104在对等的装置和/或处理之间建立会话。

[0023] 客户端102可以包括被托管在计算平台上的一个或多个处理，该计算平台通过通信适配器(未示出)可通信地耦接到网络104。类似地，服务器106可以包括被托管在一个或多个计算平台上的处理，这些计算平台通过相关联的通信适配器连接到网络104。在具体实施例中，每个服务器106可以被托管在其自己的关联计算平台上，其自己的关联计算平台与托管其他服务器106的计算平台不同。可替换地，两个或多个服务器可以被托管在单个计算平台上。

[0024] 客户端102可以发起与服务器106的会话,以通过例如建立客户端102和服务器106可经其通过网络104进行通信的会话来提供服务。客户端102可以包括被托管在计算平台上的若干应用程序中的任意一个,例如,电子邮件应用、网络浏览器、文档处理程序、媒体播放器等。这里,应当理解,客户端102可被托管在能够利用网络进行通信的若干装置中的任一装置上和/或在其上被执行,例如,举几个例子,所述装置例如是个人计算机、移动电话、个人数字助理。在具体实施例中,尽管所要求保护的主题在这方面不受限制,但是服务器可以包括适于处理打印任务的打印机、网络服务器、电子邮件服务器、适于(例如通过流传输)分发音频或视频内容的内容服务器。在具体实施例中,服务器106可以通过单个套接字(socket)地址来寻址。套接字地址可以包括例如组合了IP地址和UDP端口号的信息。

[0025] 这里提及的基于爬虫的搜索引擎涉及利用爬虫技术来搜索网络可访问数据源并索引搜索结果的搜索引擎。作为对查询的应答的搜索结果是从被索引的内容中而不是整个网络中提取的。搜索结果可以包括链接的列表,所述链接链接到包含被认为与该查询相关的信息的网页。基于爬虫的搜索引擎可以,在用户已经在查询字段中输入查询并且通过点击“搜索(Search)”、“输入(Enter)”、“开始(Go)”或另一开始输入命令而发出了查询之后,提供搜索结果。这样的查询可被与之前通过爬虫技术收集的搜索结果的索引相比较。以这种方式获得的搜索结果可能包括可能与该查询的目的并不完全相关的大量“命中(hit)”(例如,网页链接)。

[0026] 图2图示出根据一个实施例的如显示装置220上所显示的网络浏览器200。显示装置220例如可以是托管客户端的计算平台的组件,但是不限于此。用户可以经由网络浏览器200在查询字段202中输入构成查询208的字符。可以选定搜索按钮206以将查询208传送给搜索引擎(未示出)以进行处理。这样的搜索引擎例如可以在服务器上实现,但是不限于此。这里,这样的搜索引擎可以基于基于爬虫的被索引的搜索的内容来编制与查询208的主题相关的搜索结果212。然后,搜索引擎可以将搜索结果212传送给显示装置220。在显示装置220上,搜索结果212可被呈现在网络浏览器200的结果页面210中。

[0027] 在一个实施例中,搜索者可以找到他认为对于具体的查询主题提供高度相关的有用搜索结果的数据源。举几个例子,这样的搜索结果可以包括到网站、网页或数据库的链接。搜索者可以创建所谓的“联合搜索”,“联合搜索”可操作用于搜索其中所包含的那些具体数据源。联合搜索可以被保存在电子文件中,该电子文件可用于将来对与该查询主题相关的主题的搜索。

[0028] 由搜索者创建的联合搜索可以不仅包含使用爬虫型搜索结果发现的搜索结果,还可以包括未被爬取的搜索结果。联合搜索的创建者可能有权使用公众不一定能使用的或不一定可通过爬虫搜索或索引的数据库。在一个实施例中,搜索者可以将这些数据库包括在联合搜索中。为了方便对以其它方式不可受访问的数据库具有许可权的搜索者,联合搜索可以提供到那样的数据库的链接。联合搜索还可以包含用于提供对具体的受访问限制的数据库的使用权的相关语法。

[0029] 图3是图示出根据一个实施例的联合搜索库300的示例的示图。联合搜索库300是包含一个或多个联合搜索310A-D的电子库。第一联合搜索310A包括多个数据源,包括网站312和数据库314。作为说明和示例,而非限制于此,第一联合搜索310A可以与钓鱼的主题相关。数据源可以包括与钓鱼的主题有关的网站312并且数据库314可以包括由用户捕到的所

有鱼种的个人日志。其他联合搜索310B-D可以转到与搜索者感兴趣的其他主题。

[0030] 再次参考图2,根据一个实施例,网络浏览器200可被使能以用于搜索联合搜索。用户经由网络浏览器200将构成查询208的字符输入查询字段202中。搜索按钮206可被用来选择针对该搜索所要使用的特定联合搜索。基于查询208的搜索可在联合搜索中所包含的数据源中的任一数据源处被执行。例如,搜索结果212可在显示装置220上被呈现为结果页面210。

[0031] 根据一个实施例,联合搜索的创建者可以公布联合搜索以供其他人使用。用户的联合搜索中的每个联合搜索都可以被上载到例如可由其他用户访问的服务器。这样,联合搜索可被共享,并且提供联合搜索团体库的联合搜索提供者团体可被建立。另外,某些联合搜索创建者可以由于是具体查询主题的专家并且已经在那个领域中提供了特别相关和可信的联合搜索而发展在团体内的信誉。这样,搜索者可以找出针对给定查询主题的具体联合搜索的具体联合搜索创建者。

[0032] 根据一个实施例,联合搜索库可以被网络搜索引擎存储并可由其访问。库内的联合搜索可被网络搜索引擎索引,以变得对搜索者而言更易访问。

[0033] 图4是用于处理联合搜索的方法的实施例400的流程图。每个联合搜索可被提供给搜索引擎402,与一个或多个相关联的关键字一起404。关键字例如可以由联合搜索提供者提供。每个联合搜索可以根据相关联的关键字中的每一个关键字被索引406。

[0034] 根据用于联合搜索的搜索的实现方式,查询可被与一个或多个关键字比较和匹配408。与一个或多个关键字相关联的一个或多个联合搜索可被提供作为针对该查询的搜索结果410。

[0035] 图5是提供用于处理联合搜索的方法的实施例400的示例图形表示的实施例500的示图。每个联合搜索301A-C可以分别与一个或多个关键字520A-C相关联。每个联合搜索310A-C可以被传送并被存储在联合搜索库530中。每个联合搜索可以根据每个关键字被索引到关键字关联532A-C中。每个关键字关联532A-C可以包括关键字和与该关键字相关联的一个或多个联合搜索。

[0036] 查询508可以被传送给搜索引擎(未示出)并且可以与每个关键字关联532A-C中的关键字相比较和相匹配。与匹配的关键字相关联的联合搜索534A-B可以作为搜索结果被提供给结果页面504,以由搜索者用来使用所提供的联合搜索534A-B中的一个或多个来执行联合搜索。

[0037] 再次参考图2,结果页面210可以向用户呈现包括联合搜索534A-D搜索结果104。用户可以选择感兴趣的联合搜索534A-D。在选择联合搜索534A-D之后,联合搜索请求利用查询208被传送给搜索引擎。这里,这样的搜索引擎可以使用该查询来在与所选择的联合搜索534A-D相关联的数据源中执行搜索。搜索引擎可以将联合搜索结果212传送给浏览器200以用于在显示装置220上开始显示。例如,搜索结果212可以被呈现在在浏览器200上的结果页面210中。

[0038] 查询208与关键字的匹配可以使用适合于从联合搜索库中选择相关的联合搜索的任意处理来执行。例如,在一个实施例中,可以利用准确匹配处理,其中该查询必需准确匹配关键字或者其最接近推导。根据另一实施例,查询208可被与关键字的同义词相比较,反之亦然。以下描述其他匹配处理。然而,这些仅仅是示例,并且实现方式不限于此。

[0039] 图6是用于处理联合搜索的方法的实施例600的流程图。每个联合搜索可被提供给搜索引擎602。每个联合搜索可以根据一个或多个索引标准被索引404。索引标准可以包括适于索引联合搜索的任意标准。索引标准可以包括关键字、主题、标题和/或数据源的任意一者或其组合，并且不限于此。

[0040] 此外，与具体索引标准相关联的每个联合搜索可以根据有用性度量被排名606。有用性度量可以包括可用来给每个联合搜索排名的任意度量。有用性度量例如可以与联合搜索对于所意图的主题的关联度和/或质量有关。

[0041] 作为一个示例，关联度度量可以包括可测量的指示符以用于确定联合搜索对于具体索引标准的关联度水平。在索引标准可以是关键字的具体示例中，关联度度量可以包括在与相同或类似关键字有关联的联合搜索中识别的数据源的数目。然而，这仅仅是一个示例，并且实现方式不限于此。

[0042] 作为另一示例，有用性度量可以与用户等级相关联。根据一个实施例，用户等级例如可以包括：针对具体索引标准，基于具体的联合搜索被感知到的有用性而给与该联合搜索的平均等级。这样的用户等级可以提供对具体联合搜索的质量的指示。可以发现某些联合搜索对于处理与具体索引标准相关联的查询特别有用。这些特别有用的联合搜索可以和指示符一起被呈现给用户。然而，这仅仅是一个示例，并且实现方式不限于此。

[0043] 在一个实施例中，根据有用性度量的排名606可与指示符一起被提供给搜索者。排名的指示符可以是呈现相关联的联合搜索的列表的形式，其中在列表的顶部呈现最高排名的联合搜索。在另一个示例中，每个联合搜索设有指示符，例如关联度百分比值和/或星级排名，其中五颗星可以指示高排名而一颗星指示低排名。然而，这些仅仅是示例并且实现方式不限于此。

[0044] 根据用于联合搜索的搜索的实现方式，查询可被与一个或多个索引标准608比较和匹配。与一个或多个索引标准相关联的一个或多个联合搜索可被选为针对该查询的搜索结果610。所选择的联合搜索可以根据一个或多个有用性度量被排名。每个排名的联合搜索可被提供给用户，每个排名的联合搜索具有排名的指示符。

[0045] 图7是提供用于处理联合搜索的方法的实施例600的示例图形表示的实施例700的示意图。每个联合搜索310A-C可以被传送并被存储在联合搜索库730中。每个联合搜索可以根据索引标准被索引到索引关联732A-C中。每个索引关联732A-C可以包括一个或多个索引标准和与这一个或多个索引标准相关联的一个或多个联合搜索。

[0046] 查询708可以被传送给搜索引擎并且可被与每个索引关联732A-C中的索引标准相比较和相匹配。与匹配的索引标准相关联的联合搜索可以由排名引擎740根据一个或多个有用性度量来排名。被索引并被排名的联合搜索734A-C可被提供作为结果页面704中的搜索结果，以由搜索者用来使用所提供的联合搜索734A-C中的一个或多个来执行联合搜索。每个联合搜索734A-C可被提供给用户，每个联合搜索734A-C具有排名的指示符。

[0047] 查询708与索引标准的匹配可以使用适合于索引和匹配联合搜索库730中相关的联合搜索的任意处理。如上所述，索引和匹配可以基于索引标准，所述索引标准包括适合于索引和匹配联合搜索的任意标准。索引标准可以包括关键字、主题、标题和/或数据源的任一者或组合，但是不限于此。作为另一示例，基于爬虫的搜索中找到的域(domain)可被与联合搜索中的域相匹配。作为另一示例，基于爬虫的搜索中找到的域可被与返回分类路径

(category path)的搜索引擎目录中的域相匹配,该分类路径又可被与从基于爬虫的搜索获得的分类路径相匹配。作为又一示例,匹配可以至少部分地基于联合搜索中的域与查询所限定的域的匹配。可以明白,索引标准和匹配可被单独使用或结合使用。然而,这些仅仅是示例,并且实现方式不限于此。

[0048] 如果用户对基于所选择的联合搜索的搜索结果不感兴趣,则用户可以选择具有更低排名的另一联合搜索。如果用户选择与搜索结果不同的联合搜索,则之后,搜索结果被更新来在结果页面中显示与不同的联合搜索有关的搜索结果。

[0049] 图8是用于处理联合搜索的方法的实施例800的流程图。每个联合搜索可被提供给搜索引擎802。搜索可以使用每个联合搜索中所包括的数据源来执行804。数据源的搜索可以是用于从数据源中提取数据的任意合适的方法。作为示例,多个查询可被用来搜索数据源以提取与每个查询有关的数据。然而,这仅仅是示例实现方式,并且所要求保护的主题不限于此。

[0050] 可以从与联合搜索相关联的数据源获取数据806。可以从可能适于索引的这样的数据中解析加权因子。作为示例,但不限于此,加权因子可以包括关键字和/或域。可以定义一个或多个签名,签名将联合搜索和与通过搜索联合搜索的数据源得到的数据相关联的加权因子相关联810。

[0051] 每个签名可以根据一个或多个索引标准被索引812。索引标准可以包括适合于索引联合搜索的任意标准。举几个例子,索引标准可以包括关键字、主题、标题和/或数据源的任意一者或组合。

[0052] 根据用于联合搜索的搜索的实现方式,查询可被与和签名相关联的一个或多个索引标准608相比较814。这样的查询可被与一个或多个签名相匹配816。与一个或多个签名相关联的一个或多个联合搜索可被提供给用户818。

[0053] 可以明白,在上述实施例中,所选择的联合搜索在被呈现给用户之前,可如上所述根据一个或多个有用性度量被排名。

[0054] 图9是提供用于处理联合搜索的方法的实施例800的示例图形表示的实施例900的示图。每个联合搜索910A-C可被传送并被存储在联合搜索库930中。每个联合搜索910A-C的数据源中的每一个可被搜索,提供与联合搜索910A-C相关联的搜索数据932A-C。在一个实施例中,作为示例但不是限制,每个数据源可基于初始查询907被搜索。联合搜索910A-C和相关联的搜索数据932A-C可被传送给搜索结果库950并被存储在其中。从搜索数据932A-C中可提取出加权因子。通过将加权因子与各个联合搜索910A-C相关联可以定义签名942A-n。签名942A-C可以使用签名库940中的索引标准被索引。

[0055] 每个签名942A-C可根据索引标准被索引到索引关联942A-C中。每个索引关联942A-C可以包括一个或多个索引标准以及与这一个或多个索引标准相关联的一个或多个签名942A-C。

[0056] 查询908可被传送给搜索引擎并且可被与每个索引关联934A-C中的索引标准相比较和相匹配。根据一个实施例,与匹配的索引标准相关联的对应签名942A-C可由排名引擎(未示出)根据一个或多个有用性度量来排名。对于与匹配的索引标准相关联的签名942A-C的联合搜索934A-C被确定。联合搜索934A-C可被提供作为结果页面904中的搜索结果,以由搜索者用来使用所提供的联合搜索934A-C中的一个或多个来执行联合搜索。

[0057] 查询908与索引标准的匹配可以使用适合于索引和匹配签名库930中的相关签名和对应联合搜索的任意处理,如上所述。

[0058] 根据一个实施例,除了由搜索引擎提供的基于爬虫的搜索结果以外或作为其替换,搜索引擎可以向搜索者提供相关的联合搜索。如果用户对基于联合搜索的搜索结果不感兴趣,则用户可以选择具有更低排名的另一联合搜索或者可以选择查看通过基于爬虫的搜索提供的搜索结果。如果用户选择列表中的不同联合搜索,则之后,搜索结果被更新以显示与该不同联合搜索有关的搜索结果。

[0059] 图10是提供用于处理联合搜索的系统的实施例1001的示例图形表示的示图。实施例1001包括可操作用于提供基于联合搜索和/或基于爬虫的搜索的搜索结果的浏览器1006。用户可以在客户端1000上的网络浏览器1002中输入查询1012。查询1012可通过网络1050被传送给前端服务器1004。前端服务器1004和网路1050可以不是任意实施例中都需要的元件,而仅仅被图示出来示出可以实现实施例1001的一种途径。网络1050可以是局域网(LAN)、广域网(WAN)或互联网,但不限于此。前端服务器1004可操作用于将查询1012传送给搜索引擎1006,搜索引擎1006可操作用于处理查询1012。

[0060] 搜索引擎1006可以包括基于爬虫的搜索引擎1007和联合搜索引擎1010。基于爬虫的搜索引擎1007可操作用于提供如上所述的基于爬虫的搜索结果。基于爬虫的搜索引擎1007可操作用于将搜索结果传送给前端服务器1004以用于与客户端1000通信。

[0061] 前端服务器1004和搜索引擎1006可以在相同或不同装置上实现。然而,为了说明的目的,它们被示出驻留在不同装置上。

[0062] 联合搜索引擎1010可以包括具有一个或多个被索引的联合搜索的数据库。联合搜索引擎1010可操作用于基于查询1012来确定一组一个或多个被索引的联合搜索。联合搜索引擎1010可操作用于将这一组被索引的联合搜索传送给前端服务器1004。根据一个实施例,除了这组被索引的联合搜索以外,搜索引擎1006还可操作用于传送附加数据,所述附加数据指示这些被索引的联合搜索中的哪些是最相关的。

[0063] 前端服务器1004可操作用于通过网络1050向客户端1000传送相关的被索引的联合搜索的列表和关联度数据,以用于开始在网络浏览器1002上显示。前端服务器1004可操作用于通过网络1050向客户端1000传送基于基于爬虫的搜索的搜索结果以用于在网络浏览器1002上显示。

[0064] 根据一个实施例,可向用户呈现这样的选项,其用于选择基于基于爬虫的搜索或被索引的联合搜索之一的搜索结果。

[0065] 在选择用户感兴趣的被索引的联合搜索之后,网络浏览器1002可操作用于将所选择的被索引的联合搜索通过网络1050传送给前端服务器1004,前端服务器1004可操作用于将所选择的被索引的联合搜索传送给搜索引擎1008。搜索引擎1008可操作用于处理所选择的被索引的联合搜索来获得联合搜索结果。联合搜索结果可被传送给前端服务器1004。前端服务器1004可操作用于将联合搜索结果1018通过网络1050传送给客户端1000以用于开始在网络浏览器1002中显示。

[0066] 以上实现方式的一个优点是:相关的被索引的联合搜索的列表和标准的爬虫搜索结果一起被传送给用户,并且被提供作为标准的爬虫结果的替换。然而,该实现方式也说明了这样的可能性:对于输入网络浏览器1002的查询字段1012的每个查询,搜索请求不得不

在客户端1000和前端服务器1004之间进行两次完整的往返。

[0067] 根据另一实现方式,基于查询的最高排名的被索引的联合搜索的联合搜索结果可被默认地传送给网络浏览器1002。在该实现方式中,用户可以在查询仅进行一次往返的情况下接收相关搜索结果。

[0068] 根据另一实施例,搜索引擎1006可操作用于处理所选择的被索引的联合搜索来获得联合搜索结果。这些联合搜索结果可被传送给前端服务器1004。前端服务器1004可操作用于通过网络1050向客户端1000传送以下两者:联合搜索的列表以及,或者基于被索引的联合搜索或者基于标准的基于爬虫的搜索的搜索结果。

[0069] 在没有前端服务器1004时,网络浏览器1002可操作用于将联合搜索传送给网络1008,随后,万维网1008可操作用于通过网络1050向客户端1000返回联合搜索结果。

[0070] 该实现方式的另一优点在于,客户端1000和前端服务器1004之间的通信(即流量)更少。然而,搜索结果可以和在之前的实施例中一样快地传送给客户端1000,这是因为在联合搜索的列表被传送给客户端1000之前,被索引的联合搜索的列表必需“等待”搜索结果产生并被传送给前端服务器1004。

[0071] 在另一实施例中,如图2中所示,用户界面包括查询字段102、下拉框104、搜索结果页面110和“搜索”按钮106。查询字段102可以操作使得用户可以输入构成查询的字符。下拉框104可操作用于提供一组一个或多个索引的联合搜索的列表。

[0072] 搜索按钮106可以是如果用户对由相应联合搜索提供的初始联合搜索结果不满意,则允许用户选择用户输入的索引的联合搜索的任意机构的形式。下拉框104中列出的联合搜索的列表可以在其他类型的用户界面元件中呈现,这些用户界面元件包括但不限于文本框、列表框、菜单或上下文菜单。用户界面可以使用网络浏览器302来浏览,网络浏览器302例如但不限于是Internet Explorer或Mozilla Firefox。

[0073] 图11是图示出可以实现实施例的计算机系统1100的框图。计算机系统1100包括用于传输信息的总线1102或其他通信机构,和用于处理信息的耦接到总线1102的处理器1104。计算机系统1100还包括用于存储信息和由处理器1104执行的指令的耦接到总线1102的主存储器1106,例如随机存取存储器(RAM)或其他动态存储装置。主存储器1106也可以用于存储在由处理器1104执行的指令的执行期间的临时变量或其他中间信息。计算机系统1100还包括用于存储静态信息和用于处理器1104的指令地耦接到总线1102的只读存储器(ROM)1108或其他静态存储装置。用于存储信息和指令的诸如磁盘或光盘之类的存储装置1110被提供并耦接到总线1102。

[0074] 计算机系统1100可经由总线1102耦接到显示器1112,显示器1112例如是液晶显示器(LCD),用于向计算机用户显示信息。输入装置1114包括数字字母和其它键,其可以耦接到总线1102,以用于向处理器1104传送信息和命令选择。其他类型的用户输入装置可以是光标控制器1116,例如,鼠标、跟踪球或光标方向键,用于向处理器404传送方向信息和命令选择以及用于控制显示器1112上的光标移动。该输入装置通常在允许装置指定平面中的位置的两轴(例如,第一轴(例如,x)和第二轴(例如,y))上具有自由度。

[0075] 本主题与用于实现在此描述的技术的计算机系统1100的使用有关。根据一个实施例,计算机系统1100,响应于处理器1104执行被包含在主存储器1106中的一个或多个指令的一个或多个序列,来执行那些技术。这样的指令可以从机器可读介质,例如存储装置

1110,读入主存储器1106中。主存储器1106中所包含的指令的序列的执行使得处理器1104执行在此描述的处理步骤。在可替换实施例中,可以使用硬连线电路来替换软件指令或与之结合来实现本发明。因此,本发明的实施例不限于硬件电路和软件的任何特定组合。

[0076] 在此使用的术语“机器可读介质”是指参与提供使得机器以特定方式操作的数据的任意介质。在使用计算机系统1100实现的一个实施例中,例如,在向处理器1104提供用于执行的指令时,涉及各种机器可读介质。这样的介质可以采取许多形式,包括但不限于,非易失性介质、易失性介质和传输介质。非易失性介质例如包括光盘或磁盘,例如存储装置1110。易失性介质包括动态储存区,例如主存储器1106。传输介质包括同轴电缆、铜线和光纤,包括包含总线1102的电线。传输介质还可以采取声波或光波的形式,例如在无线电和红外数据通信期间生成的那些。

[0077] 机器可读介质的通常形式例如包括闪盘、软盘、硬盘、磁盘或任意其它磁介质、CD-ROM、任意其它光介质、打孔卡、纸带、带有孔的图案的任意其它物理介质、RAM、PROM和EPROM、FLASH-EPROM、任意其它存储芯片或存储盒、如下所述的载波或计算机可读的任意其它介质。

[0078] 在向处理器1104携带用于执行的一个或多个指令的一个或多个序列时,涉及机器可读介质的各种形式。例如,指令可能一开始被携带在远程计算机的磁盘上。远程计算机可以将指令装载到其动态储存器中并使用调制解调器通过电话线发送指令。计算机系统1100本地的调制解调器可以接收电话线上的数据并使用红外发射器将数据转换成红外信号。红外检测器可以接收红外信号中所携带的数据并且合适的电路可以将这些数据置于总线1102上。总线1102将这些数据携带至主存储器1106,处理器1104可以从主存储器1106获取指令并执行指令。可选地,由主存储器1106接收到的指令可以在处理器1104的执行之前或之后被存储在存储装置1110上。

[0079] 计算机系统1100还包括耦接到总线1102的通信接口418。通信接口1118提供耦接到与本地网络1122连接的网络链路1120的双向数据通信。例如,通信接口1118可以是用于提供到对应类型的电话线的数据通信连接的综合业务数字网络(ISDN)卡或调制解调器。作为另一示例,通信接口1118可以是用于提供到可兼容的LAN的数据通信连接的局域网(LAN)。也可以实现无线链路。在任意这样的实现方式中,通信接口1118发送和接收携带表示各种信息的数字数据流的光的、电磁的或光学的信号。

[0080] 网络链路1120通常通过一个或多个网络来提供到其他数据装置的数据通信。例如,网络链路1120可以通过本地网络1122来提供到托管计算机1124或到由互联网服务提供商(ISP)1126操作的数据设备的连接。ISP 1126进而通过全球分组数据通信网络(现在常称为“互联网”)1128来提供数据通信服务。本地网络1122和互联网1128两者都使用携带数字数据流的电磁信号或光信号。通过各种网络的信号以及在网络链路1120上和通过通信接口1118的信号携带来自和去往计算机系统1100的数字数据,并且是传送信息的载波的示例形式。

[0081] 计算机系统1100可以通过(一个或多个)网络、网络链路1120和通信接口1118来发送消息和接收数据,包括程序代码。在互联网示例中,服务器1130可以通过互联网1128、ISP 1126、本地网络1122和通信接口1118来传送所请求的应用程序代码。

[0082] 接收的代码在其被接收到时可由处理器404执行,和/或被存储在存储装置1110或

其他非易失性存储器中以用于稍后执行。以这种方式，计算机系统1100可以获得载波形式的应用代码。

[0083] 尽管已经图示和描述了当前所想到的示例实施例，但是本领域技术人员应当理解，在不偏离所要求保护的主题的情况下，可以做出各种其他修改例，并且可以取代等同例。另外，在不偏离在此描述的中心思想的情况下，可以进行许多修改来使得特定情形适应所要求保护的主题的教导。因此，所要求保护的主题不希望被限制于所公开的具体实施例，而是，这样的所要求保护的主题也可以包括落在所附权利要求的范围内的所有实施例，以及其等同例。

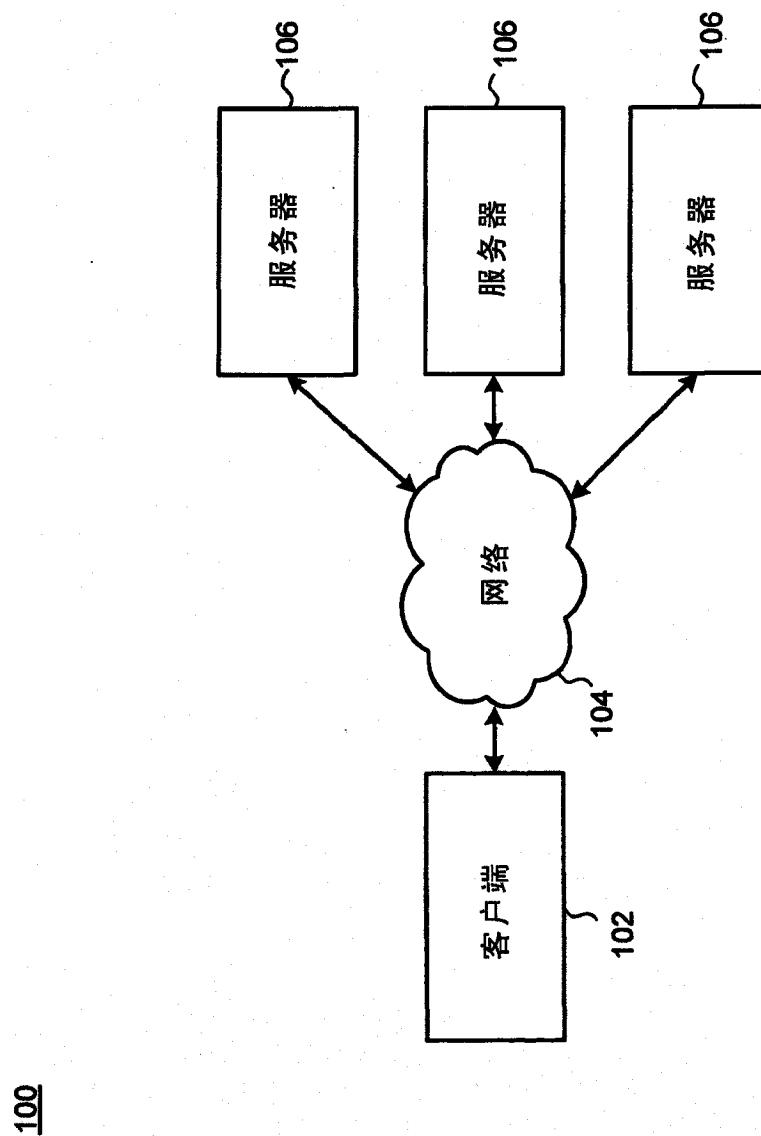


图1

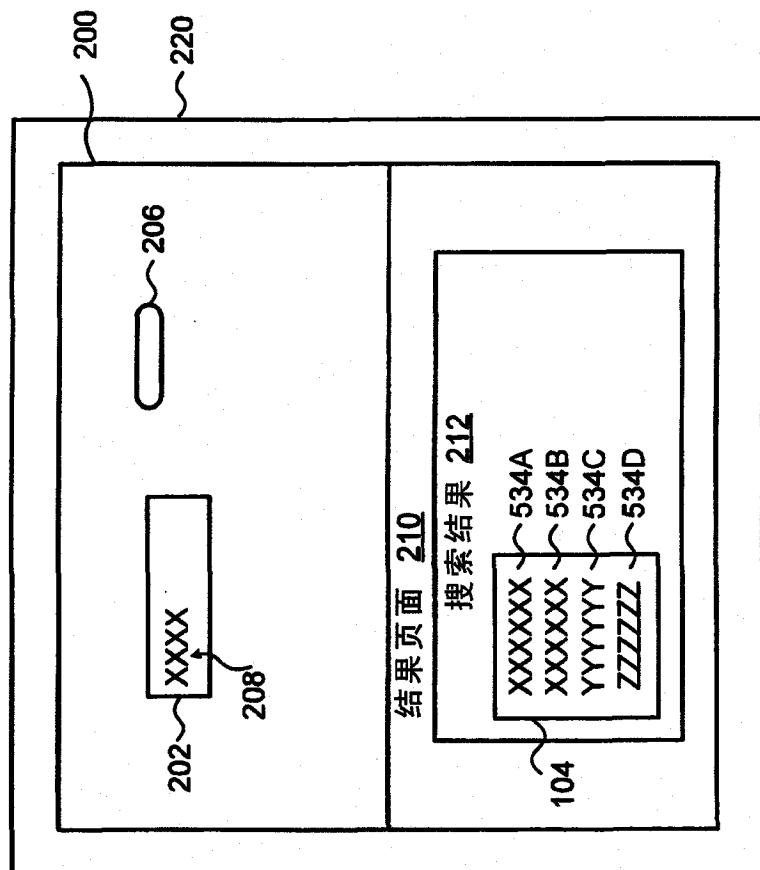


图2

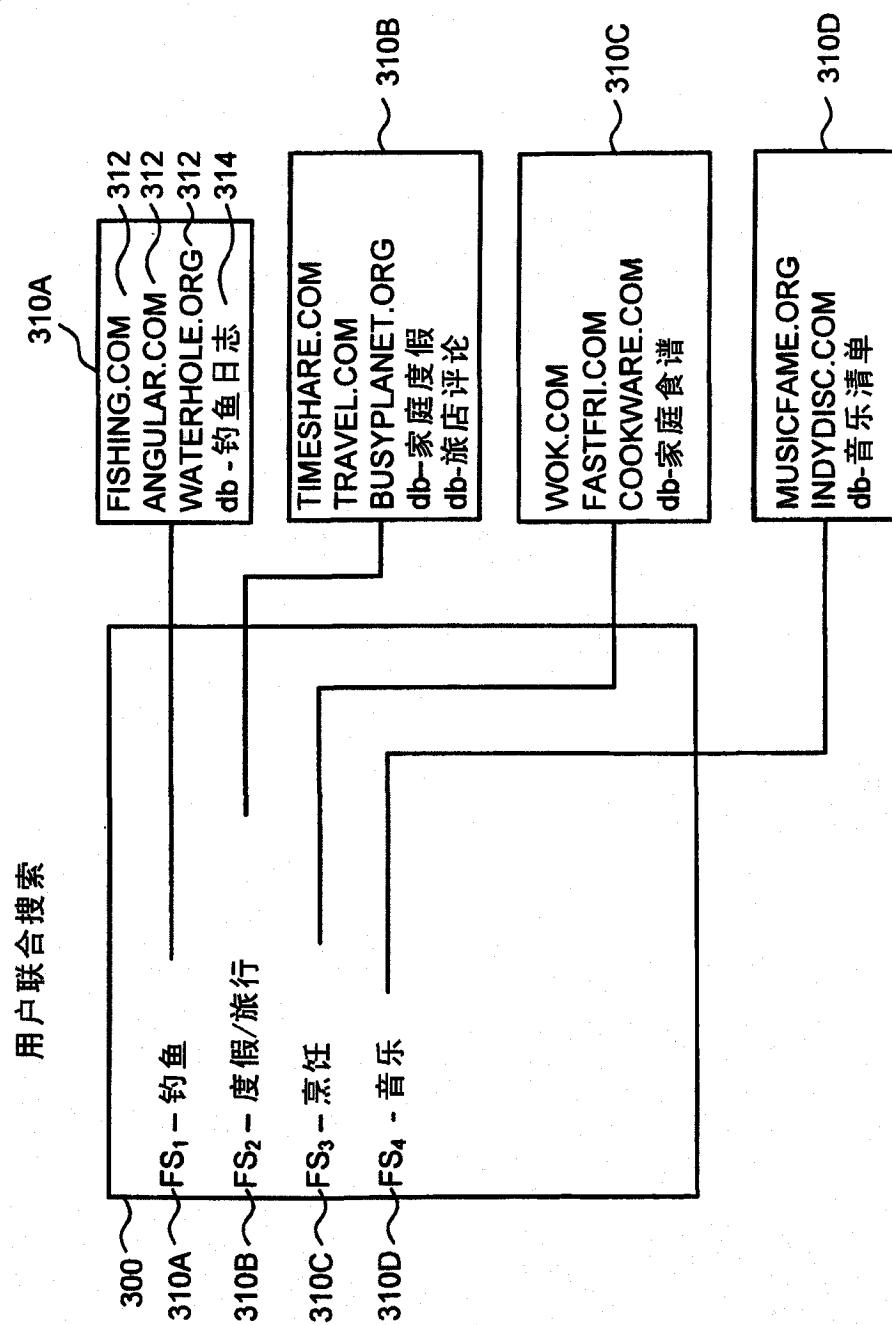


图3

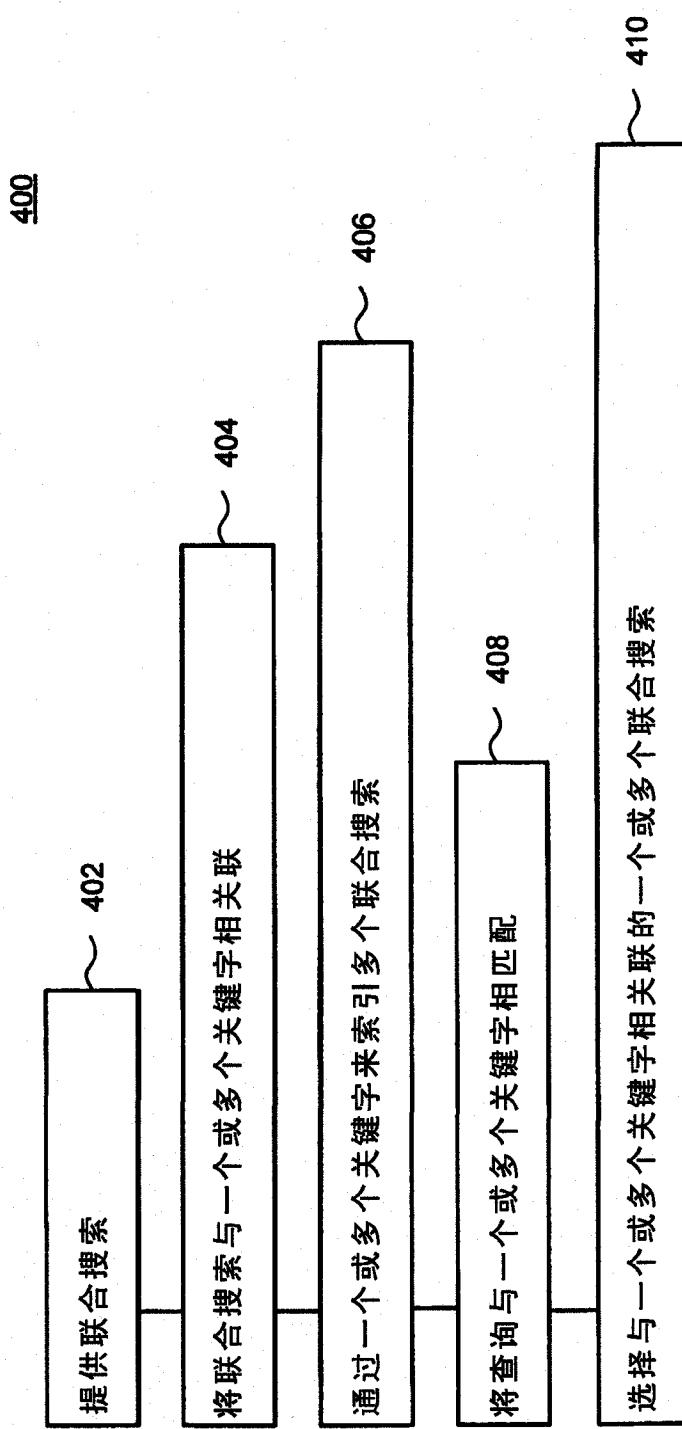


图4

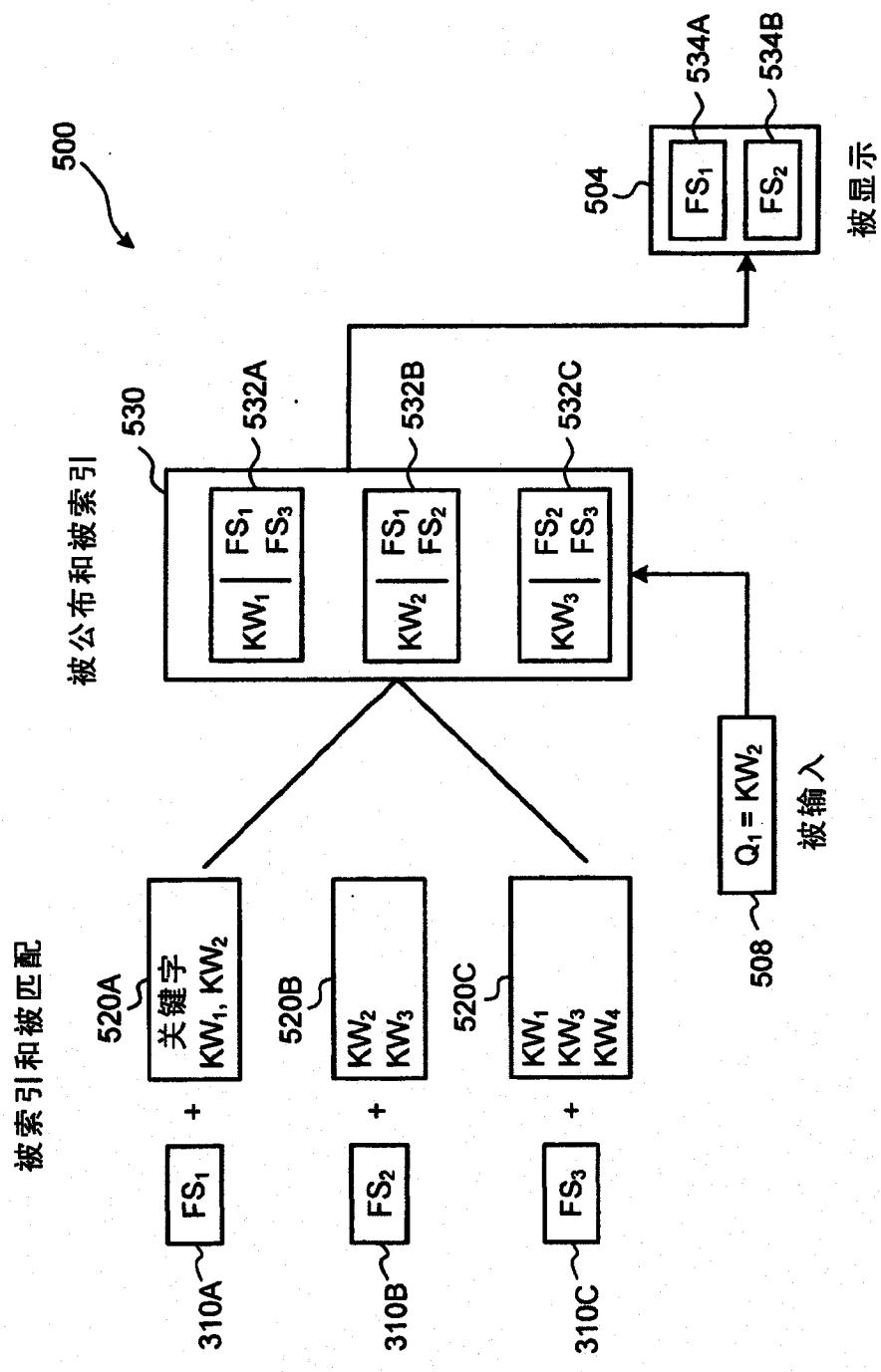


图 5

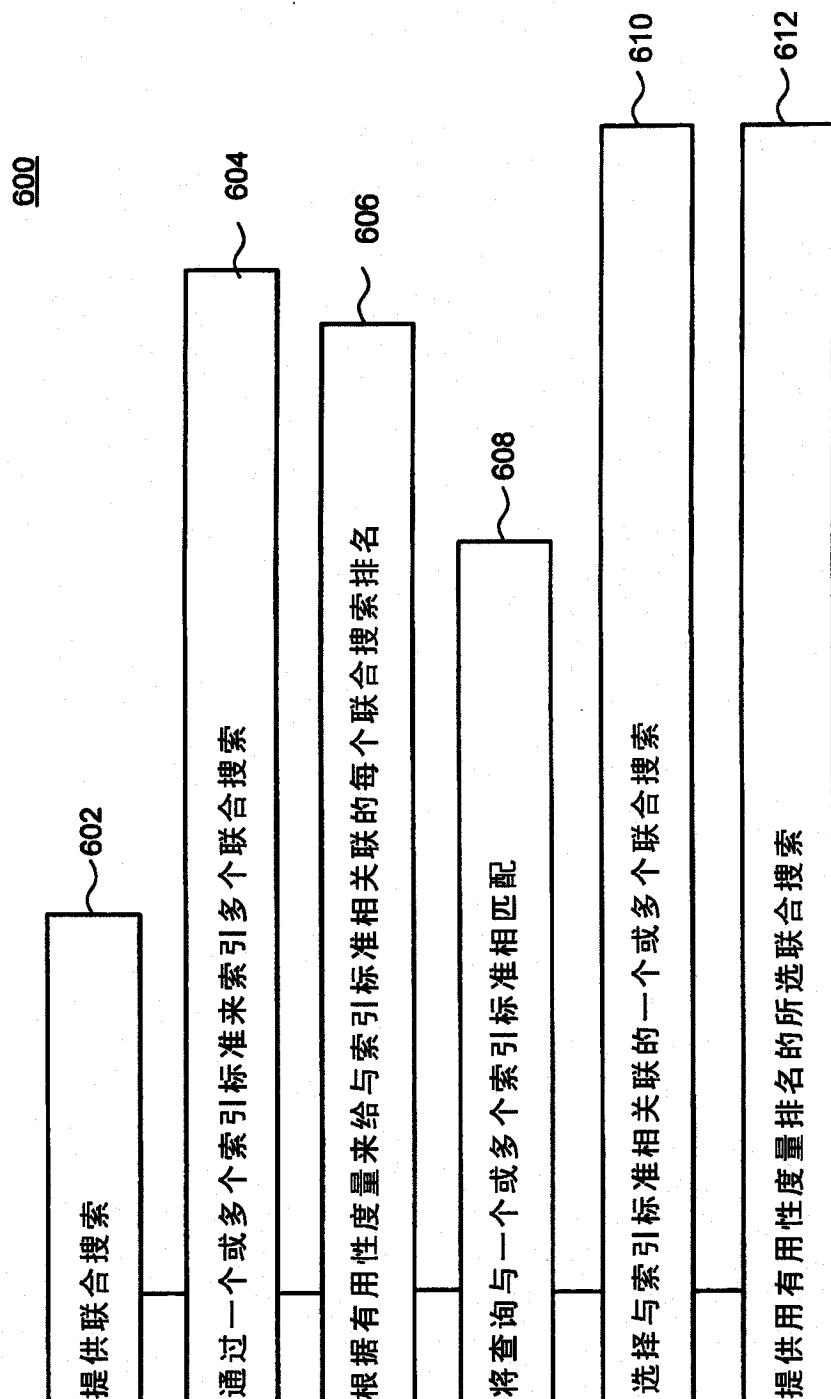


图6

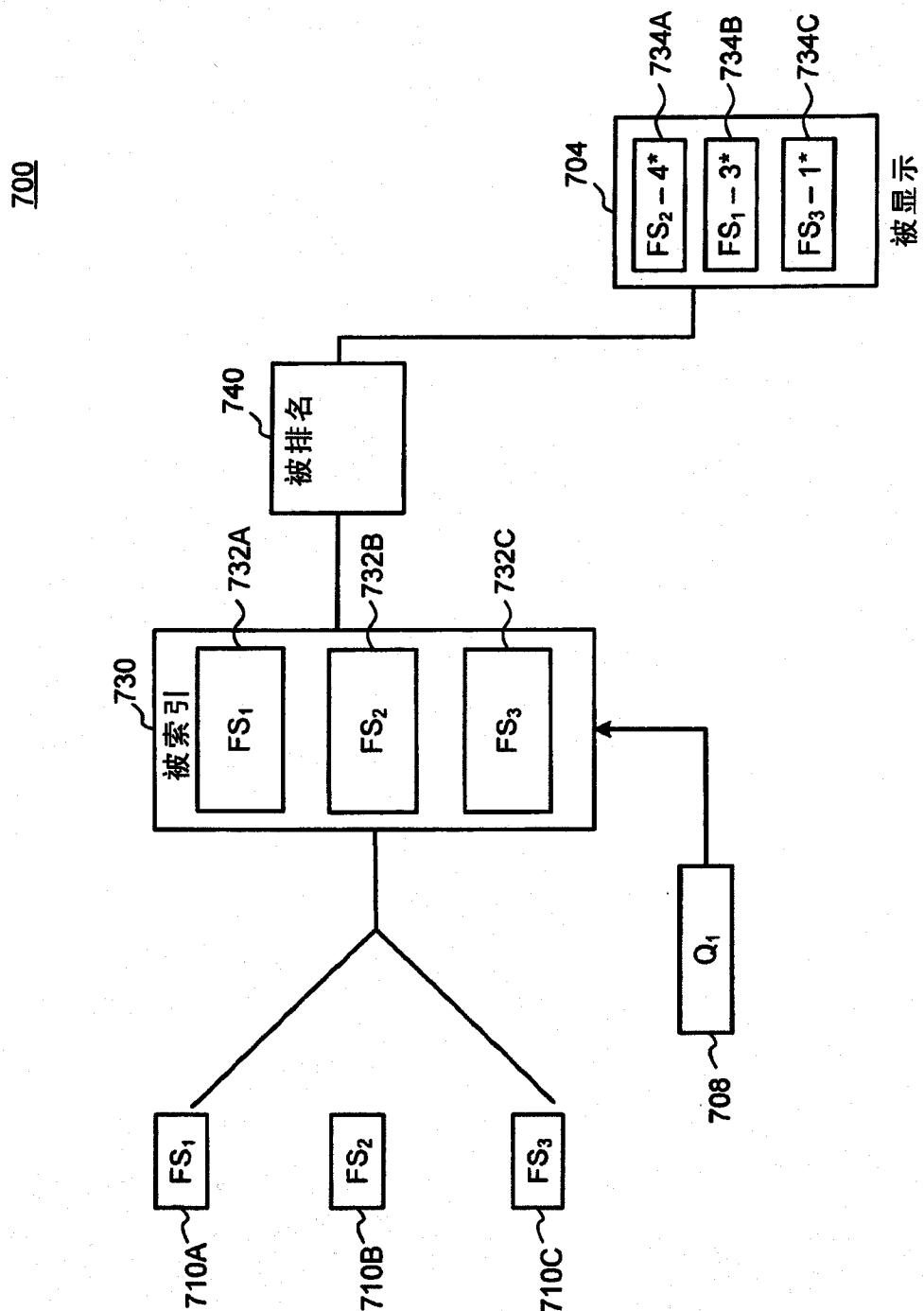


图7

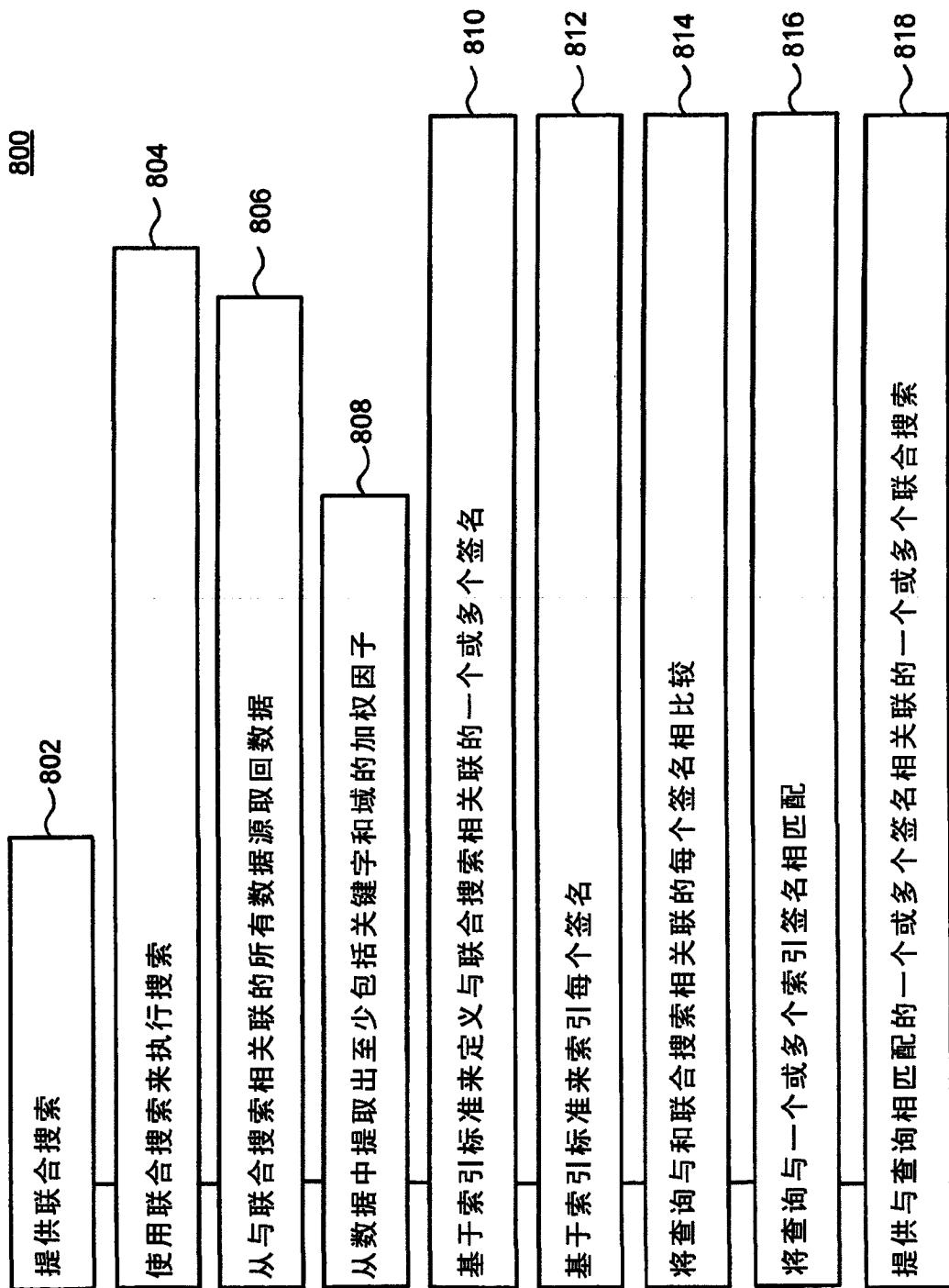


图8

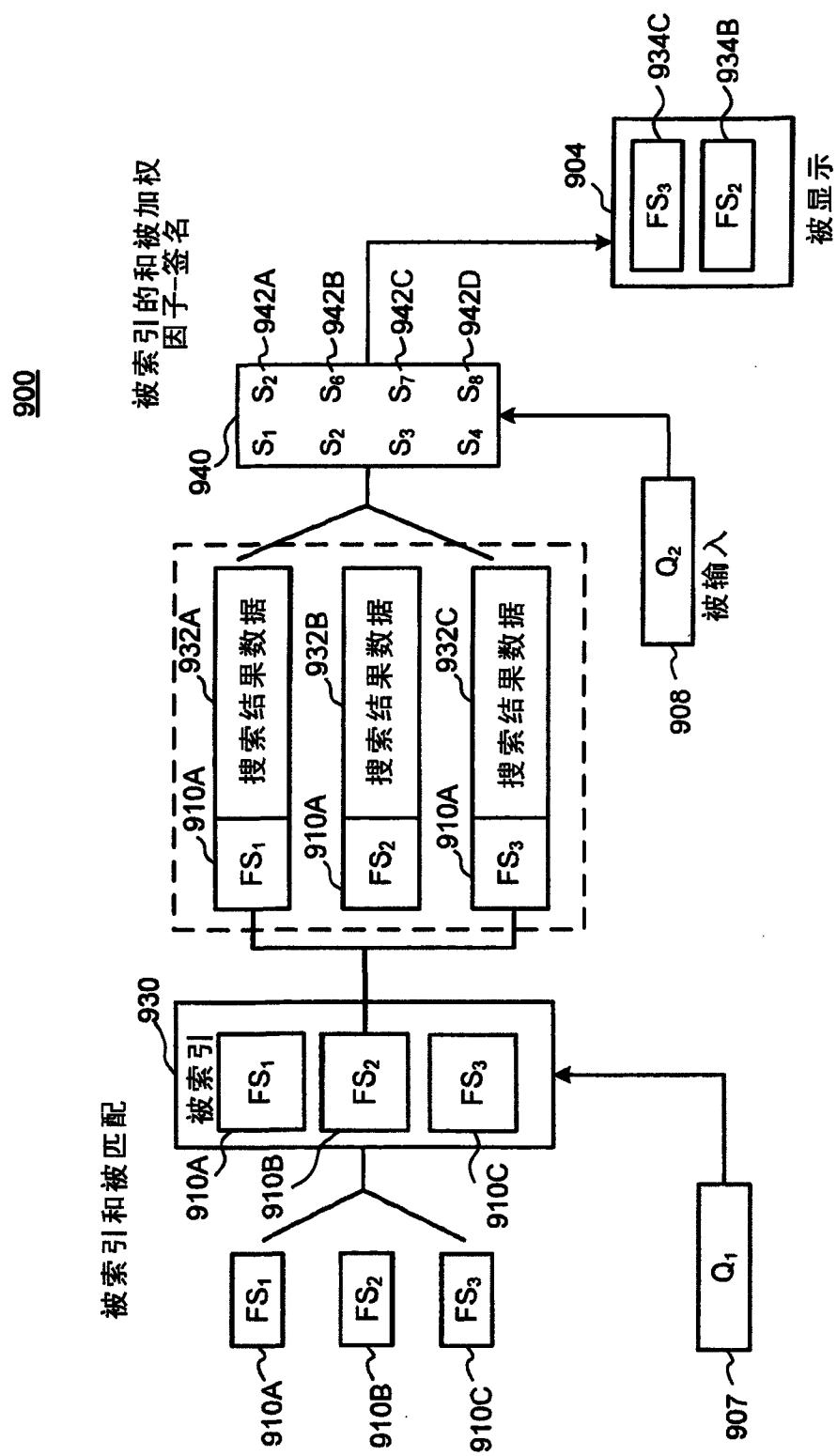


图9

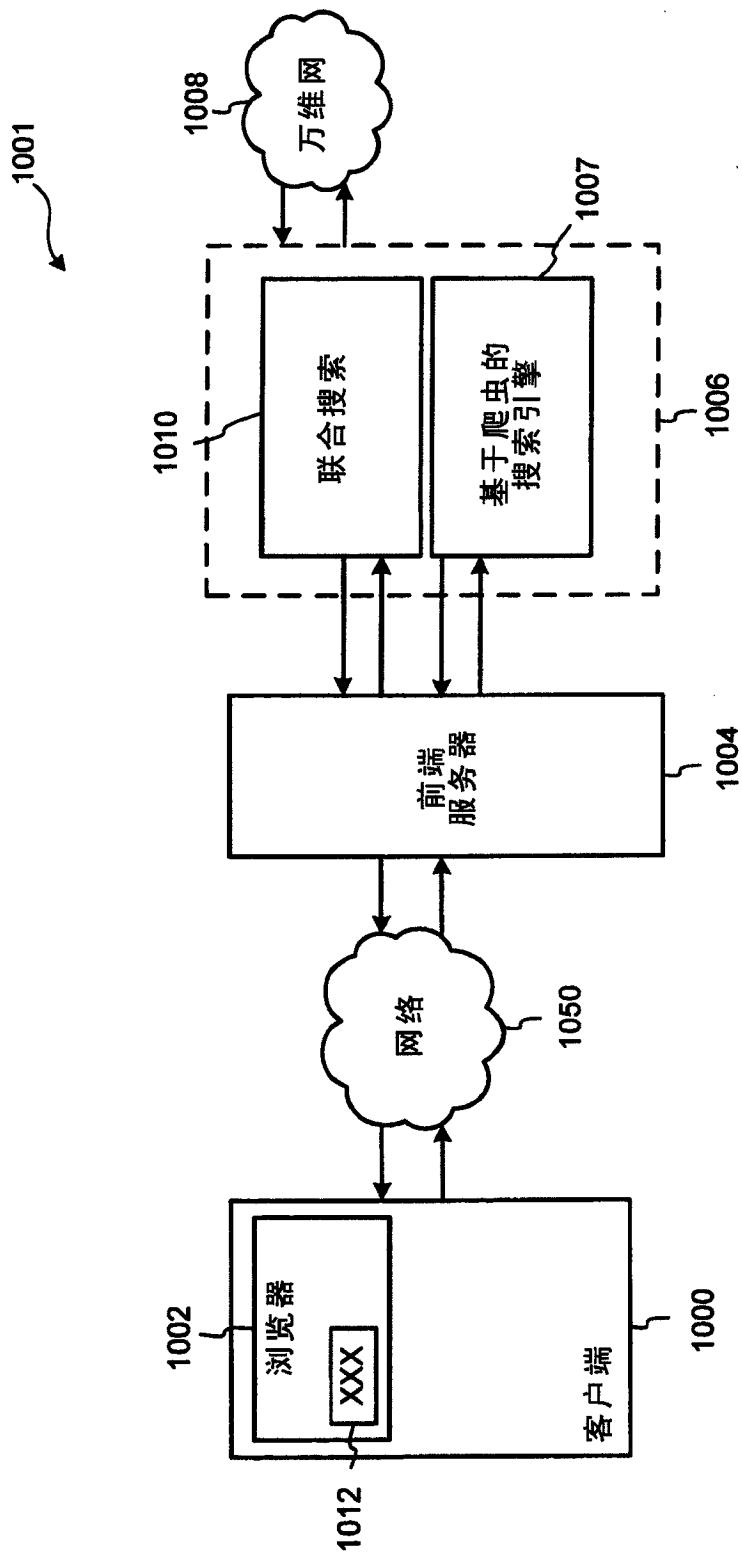


图10

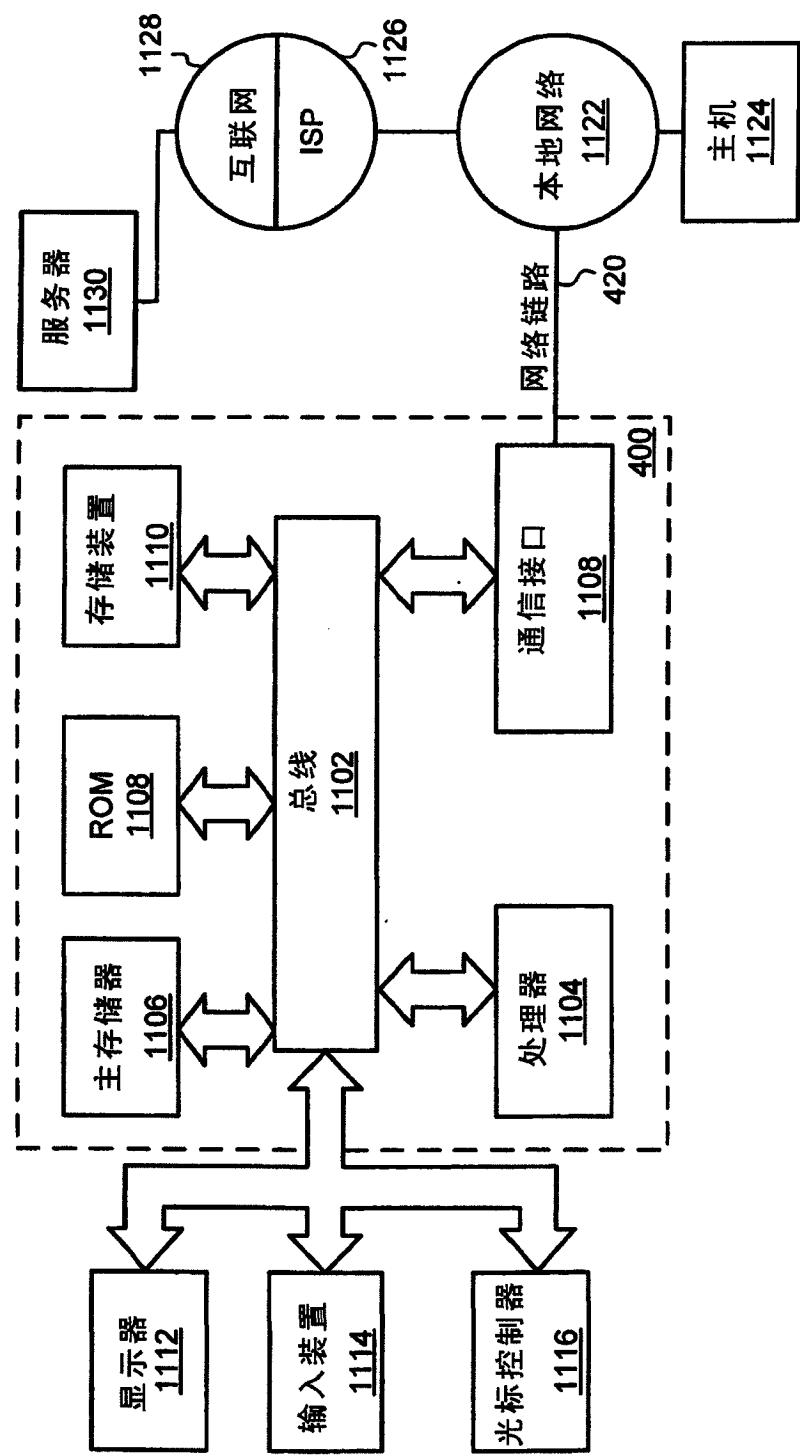


图11