



[12] 发明专利申请公布说明书

[21] 申请号 200680007639.6

[43] 公开日 2008年3月5日

[11] 公开号 CN 101137985A

[22] 申请日 2006.3.10

[21] 申请号 200680007639.6

[30] 优先权

[32] 2005.3.10 [33] US [31] 11/078,685

[86] 国际申请 PCT/US2006/008961 2006.3.10

[87] 国际公布 WO2006/099331 英 2006.9.21

[85] 进入国家阶段日期 2007.9.10

[71] 申请人 雅虎公司

地址 美国加利福尼亚州

[72] 发明人 希亚姆·卡普尔 吉格纳舒·帕里克

[74] 专利代理机构 北京康信知识产权代理有限责任
公司

代理人 余刚 尚志峰

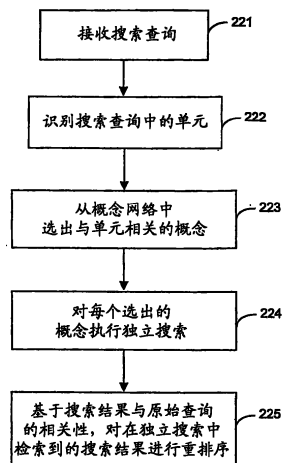
权利要求书5页 说明书12页 附图3页

[54] 发明名称

重排序和提高搜索结果的相关性

[57] 摘要

本发明提出了一种用于重排序和提高互联网搜索的结果的相关性的方法。首先将搜索查询分解为多个独立的单元。每个单元对应于一个或多个表示自然概念的词。对概念网络进行分析，以定位与搜索查询中的单元相关的概念。从概念网络选出特定概念。对每个选出的概念执行独立的互联网搜索。将从这些搜索得出的搜索结果与原始搜索查询中的单元进行比较，并根据它们与原始搜索查询的相关性对搜索结果进行分级。



1. 一种提高在搜索中检索到的搜索结果的相关性的方法，所述方法包括以下步骤：

识别搜索查询中的单元，其中，每个所述单元对应于所述搜索查询中的一个概念；

选出概念网络中的与所述搜索查询中的所述单元相关的一个或多个概念；

基于从所述概念网络中选出的所述相关概念中的至少一个执行搜索；以及

基于所述搜索结果与所述搜索查询的相关性，对在所述搜索中检索到的搜索结果进行分类。

2. 根据权利要求1所述的方法，其中，所述搜索为互联网搜索。
3. 根据权利要求1所述的方法，其中，从所述概念网络选出一个或多个所述相关概念进一步包括以下步骤：

确定所述相关概念在之前提交的搜索查询中出现的频繁程度；

选出相对于其他相关概念而言，在所述之前提交的搜索查询中出现较为频繁的相关概念。

4. 根据权利要求1所述的方法，其中，从所述概念网络选出一个或多个所述相关概念进一步包括以下步骤：

从所述概念网络中的概念中选出预定数目的与所述单元最为相关的相关概念。

5. 根据权利要求1所述的方法,其中,从所述概念网络选出一个或多个所述相关概念进一步包括以下步骤:

从所述概念网络中选出直接连接至所述单元的所有相关概念。

6. 根据权利要求1所述的方法,其中,从所述概念网络选出一个或多个所述相关概念进一步包括以下步骤:

从所述概念网络中选出与所述单元是同义词的相关概念。

7. 根据权利要求1所述的方法,其中,对在所述搜索中检索到的所述搜索结果进行分类进一步包括以下步骤:

基于所述搜索查询中的所述单元在每个所述搜索结果中出现的频繁程度,为每个所述搜索结果分配重排序分数。

8. 根据权利要求7所述的方法,其中,对在所述搜索中检索到的所述搜索结果进行分类进一步包括以下步骤:

基于所述搜索查询中的所述单元、联合单元、及扩展单元在每个所述搜索结果中出现的频繁程度,为每个所述搜索结果分配重排序分数。

9. 根据权利要求7所述的方法,其中,每个所述重排序分数包括用于对每个所述相关概念执行搜索的搜索引擎分配的层级分数的基值。

10. 根据权利要求8所述的方法,其中,对在所述搜索中检索到的所述搜索结果进行分类进一步包括以下步骤:

基于所述重排序分数的值,对所述搜索结果进行分类。

11. 根据权利要求 10 所述的方法，其中，对在所述搜索中检索到的所述搜索结果进行分类进一步包括以下步骤：

为导航查询分配高于在其它类型的查询中检索到的搜索结果的层级。

12. 一种提高在网络搜索中检索到的搜索结果的相关性的计算机系统，所述计算机系统包括：

用于识别搜索查询中的单元的代码，其中，每个所述单元对应于所述搜索查询中的一个概念；

用于选出概念网络中的一个或多个与所述单元相关的概念的代码；

用于执行至少一次搜索以定位搜索结果的代码，其中，所述搜索结果与从所述概念网络中选出的至少一个所述相关概念相匹配；

用于基于所述搜索结果与所述搜索查询的相关性，来对在至少一次搜索中检索到的搜索结果进行重排序的代码；以及
计算机可读介质，用于存储所述代码。

13. 根据权利要求 12 所述的计算机系统，其中，用于选出所述概念网络中的一个或多个所述相关概念的代码进一步包括：

用于确定所述相关概念在之前提交的搜索查询中出现的频繁程度的代码。

14. 根据权利要求 13 所述的计算机系统，其中，用于选出所述概念网络中的一个或多个所述相关概念的代码进一步包括：

用于选出相对于其他相关概念而言，在所述之前提交的搜索查询中出现最为频繁的相关概念的代码。

15. 根据权利要求 12 所述的计算机系统，其中，用于选出所述概念网络中的一个或多个所述相关概念的代码进一步包括：

用于从所述概念网络中的概念中选出预定数目的与所述单元最为相关的相关概念的代码。

16. 根据权利要求 12 所述的计算机系统，其中，用于选出所述概念网络中的一个或多个所述相关概念的代码进一步包括：

用于在所述概念网络中选出与所述单元直接连接的所有相关概念的代码。

17. 根据权利要求 12 所述的计算机系统，其中，用于选出所述概念网络中的一个或多个所述相关概念的代码进一步包括：

用于选出所述概念网络中的与所述单元是同义词的相关概念的代码。

18. 根据权利要求 12 所述的计算机系统，其中，用于对在至少一次搜索中检索到的所述搜索结果进行重排序的代码进一步包括：

用于基于所述搜索查询中的所述单元在每个所述搜索结果中出现的频繁程度，为每个所述搜索结果分配重排序分数的代码。

19. 根据权利要求 12 所述的计算机系统，其中，用于对在至少一次搜索中检索到的所述搜索结果进行重排序的代码进一步包括：

用于基于所述搜索查询中的所述单元、联合单元、及单元扩展在每个所述搜索结果中出现的频繁程度，为每个所述搜索结构分配重排序分数的代码。

-
20. 根据权利要求 19 所述的计算机系统, 其中, 对在至少一次搜索中检索到的所述搜索结果进行重排序的代码进一步包括:
 用于基于重排序分数的值对所述搜索结果进行分类的代码。
 21. 根据权利要求 20 所述的计算机系统, 其中, 每个所述重排序分数包括由用于对至少一个相关概念执行搜索的搜索引擎分配的层级分数的基值。
 22. 一种计算机可读介质, 用于携带在由一个或多个处理器执行时使一个或多个处理器执行权利要求 1-11 中的任意一个所述的步骤的指令。

重排序和提高搜索结果的相关性

背景技术

本发明涉及重排序和提高互联网搜索结果的相关性的技术，更具体地涉及使用概念网络（concept network）提高互联网搜索结果的相关性并对结果进行重排序的技术。

随着互联网的出现和用户可以通过万维网（网络）访问的网页和媒体内容的增多，需要为用户提供一种从网络过滤和获取所需信息的改进方法。已开发出了多种搜索系统和程序来满足用户获取所需信息的需求。通过 Yahoo!、Google 等网站可以接触到这类技术的实例。通常，用户输入查询，搜索程序返回一个或多个与查询相关的搜索结果（链接）。返回的搜索结果可能与用户实际上所要查找的内容非常相关或完全无关。搜索结果与查询的相关性部分地取决于实际输入的查询和所使用的搜索系统（基础收集系统）的稳健性。

人类并不能自然地按照查询的方式来思考。它们部分地是由查询搜索引擎和查找图书馆目录的需求强加给我们的。人类也不能自然地按照单个词的方式来思考。人类是按照自然概念的方式来进行思考的。

搜索查询通常包括定义一个或多个概念的几个词。通常，搜索查询中的一些词比其它词与定义概念更为相关。搜索引擎无法获知搜索查询中的哪些词与用户的意图最为相关。结果，搜索引擎常常会找到许多与用户的意图无关的搜索结果。

搜索引擎通常根据搜索结果与搜索查询的相关性来对搜索结果进行排序。搜索查询经常包括多个会导致搜索引擎返回无关的搜索结果的冗余项。搜索引擎经常会将这些无关的搜索结果排列在搜索结果列表靠前的位置。

因此，需要提出一种提高搜索结果与用户的最初意图的相关性的互联网搜索方法。

发明内容

本发明提供了一种提高互联网搜索结果与用户意图的相关性的方法。本发明还提供了一种通过确定搜索结果与搜索查询中的单元、单元联合、及单元扩展的关联来对搜索查询的搜索结果进行重排序的方法。

首先，将搜索查询分解为多个独立的单元。每个单元对应于一个或多个表示自然概念的词。联合单元是同时出现在搜索查询中、但不足够相关以形成新单元的两个或多个单元。扩展单元是同时出现在搜索查询中、并足够相关以形成新单元的两个或多个单元。

本发明对概念网络进行分析，以定位与搜索查询中的单元相关的概念。概念网络将彼此相关的概念链接起来。根据特定概念与查询中的单元的关系，从概念网络中选出特定概念。

根据一个实施例，基于概念在之前提交的搜索查询中出现的频繁程度从概念网络选出概念。从选自概念网络的概念中选择在之前的搜索查询中出现频率较高的概念。对从概念网络中选出的一个或多个相关概念执行单独的互联网搜索。

将每次单独搜索得到的搜索结果与搜索查询进行比较，并根据它们与查询的相关性进行分类。可以基于搜索结果与最初的搜索查询中的单元、联合单元、及扩展单元的相关性对搜索结果进行分类。

通过下面的详细描述和附图，本发明的其它目的、特征、和优点将变得显而易见，其中，贯穿所有的附图，相同的参考标号表示相似的特征。

附图说明

图 1A 是能够实现本发明实施例的互联网通信系统的示意图；

图 1B 是能够实现本发明实施例的互联网搜索系统的示意图；

图 2 是示出根据本发明实施例的提高互联网搜索结果的相关性的方法的流程图；以及

图 3A 至 3D 示出了根据本发明的能够用于处理搜索查询的概念网络的实例。

具体实施方式

图 1A 示出了根据本发明实施例的包括客户系统 120 的信息检索和通信网络 100 的总体视图。在计算机网络 100 中，客户系统 120 可以通过互联网 140 或其它通信网络（例如，通过任意的 LAN 或 WAN 连接）与多个服务器系统 150₁ 至 150_N 通信。例如，客户系统 120 可以与搜索结果服务器 160 通信。如本文所述，根据本发明的客户系统 120 用于与服务器系统 150₁ 至 150_N 和 160 中的任意一个通信，以访问、接收、检索、以及显示媒体内容和诸如网页和网站的其他信息。

图 1A 所示系统中的多个元件包括不必在这里详细说明的传统公知元件。例如，客户系统 120 可以包括桌上型个人计算机、工作站、膝上型电脑、PDA、行动电话、或任何 WAP 使能 (WAP-enabled) 装置、或任何其它的能够直接或间接地连接互联网的计算机装置。客户系统 120 通常运行允许客户系统 120 的用户访问、处理、及浏览其能够通过互联网 140 从服务器系统 150_i 至 150_N 获取的信息和页面的 HTTP 客户程序，例如，微软的 Internet Explorer™ 浏览器、Netscape Navigator™ 浏览器、Mozilla™ 浏览器、Opera 浏览器，或移动电话、PDA、或其它无线装置中的 WAP 使能浏览器等浏览程序。

客户系统 120 通常还包括一个或多个诸如键盘、鼠标、触摸屏、笔等的用户接口装置 122，用于与由显示器（例如，监视屏、LCD 显示器等）上的浏览器提供的图形用户界面 (GUI)、以及由服务器系统 150_i 至 150_N 及其它服务器提供的页面、表格、及其它信息进行交互。本发明适用于互联网（称为特定全球互联网）。但是，应该理解，可以使用其它网络来代替互联网或将其他网络和互联网一起使用，例如，内联网、外联网、虚拟专用网络 (VPN)、非基于 TCP/IP 的网络、以及任何 LAN 或 WAN 等。

根据一个实施例，客户系统 120 及其所有的组成部分都是能够使用包括计算机代码的应用程序进行配置的运算符。其中，可以使用诸如 Intel Pentium™ 处理器、AMD Athlon™ 处理器等、或多个处理器的中央处理单元来运行计算机代码。优选地，将用于操作和配置本文中描述的用于通讯、处理、及显示数据和媒体内容的客户系统 120 的计算机代码下载并存储在硬盘上，但是也可以将整个程序代码或其多个部分存储在任何公知的其它易失性、或非易失性存储介质或装置（例如，ROM 或 RAM）中，或设置在其它任何可以存

储程序代码的介质（例如，压缩磁盘（CD）介质、数字多功能磁盘（DVD）介质、软盘等）上。

另外，可以使用任何公知的通信介质和协议（例如，TCP/IP、HTTP、HTTPS、以太网等），通过公知的互联网或其他传统的网络连接（例如，外联网、VPN、LAN等）将整个程序代码或其多个部分从软件源（例如，服务器系统 150_1 至 150_N ）传送并下载至客户系统 **120**。还应理解，用于实现本发明的多个方面的计算机代码可以通过任何能够在客户系统上执行的程序语言实现，例如，通过C、C+、HTML、XML、Java、JavaScript、任何 scripting 语言（例如VBScript）实现。在一些实施例中，不向客户系统 **120** 下载任何代码，而是通过服务器来执行所需的代码，或执行客户系统 **120** 已经存在的代码。

根据一个实施例，在客户系统 **120** 上执行的客户应用程序（由模块 **125** 表示）包括用于控制客户系统 **120** 及其组件与服务器系统 150_1 至 150_N 和 **160** 进行通信以及处理和显示从服务器系统接收到的数据内容的指令。另外，客户应用程序模块 **125** 包括用于处理数据和媒体内容的各种软件模块。例如，应用程序模块 **125** 可以包括：一个或多个搜索模块 **126**，用于处理搜索请求和搜索结果数据；用户界面模块 **127**，用于以文本、数据帧、以及活动窗口（例如，浏览器窗口和对话框）的形式呈现数据和媒体内容；以及应用程序接口模块 **128**，用于与客户系统 **120** 上执行的各种应用程序连接和通信。另外，界面模块 **127** 可以包括浏览器，例如，客户系统 **120** 上设置的默认浏览器或其他浏览器。

根据一个实施例，搜索结果服务器 **160** 用于向客户系统 **120** 提供搜索结果数据和媒体内容，服务器系统 **150** 用于（例如）响应于从服务器系统 **160** 提供的搜索结果页面中选出的链接而向客户系统 **120** 提供诸如网页的数据和媒体内容。一个实施例中的服务器系统

160 涉及用于从万维网收集信息和用于将一个或多个索引与(例如)一个或多个页面以及页面的链接等连接起来的收集方法。该收集方法包括自动网页爬虫(web crawler)、蜘蛛(spider)等,也包括手动或半自动的用于在分级结构中对网页进行分类和排序的分类算法和界面。在某些方面,服务器**160**还配置有用于对网页进行处理和排序的搜索相关算法,例如,Google的PageRank算法。服务器**160**还优选地用于以查询日志文件的形式记录搜索查询。

一方面,服务器**160**用于响应于从客户系统(特别是搜索模块**126**)接收的各种搜索请求而提供数据。服务器系统**150**和**160**可以是单个组织的一部分,例如,Yahoo! Inc.提供给用户的分布式服务器系统(distributed server system),或者它们可以是完全不同的组织的部分。服务器系统**150**和服务器系统**160**分别包括至少一个服务器和相关的数据库系统,并且可以包括多个服务器和相关的数据库系统,并且虽然以单个框示出,但是在地理上可以是分开的。例如,服务器系统**160**的所有服务器可以彼此紧邻(例如,位于一个单独的建筑物或校园内的服务器区中),或者可以彼此分布得很远(例如,位于城市A中的一个或多个服务器以及位于城市B中的一个或多个服务器)。本文中使用的术语“服务器系统”通常包括本地分布的或越过一个或多个地理位置的一个或多个逻辑上和/或物理上连接的服务器。另外,术语“服务器”通常包括本领域公知的计算机系统、相关的存储系统、以及数据库应用程序。本文中的术语“服务器”和“服务器系统”可以交换使用。

根据一个实施例,服务器**160**包括响应于从客户系统**120**接收的搜索查询向用户提供搜索结果的算法。根据本发明的实施例,服务器系统**160**用于提高从客户系统**120**接收的搜索查询的相关性(以下将详细讨论)。

图 1B 示出了实现本发明实施例的互联网搜索系统的视图。将搜索查询 170 传送至搜索引擎 175 以启动互联网搜索（例如，网络搜索）。搜索引擎 175 从搜索语料库（web corpus）190 定位匹配搜索查询 170 的网络内容。搜索语料库 190 表示可以通过万维网、互联网、内联网、本地网、及广域网访问的内容。

搜索引擎 175 从搜索语料库 190 检索匹配搜索查询 170 的内容，并将匹配的内容（即，搜索结果）传送至页面汇编程序 180。页面汇编程序 180 根据搜索结果与搜索查询的相关性对搜索结果进行分类，并以便于向用户显示的顺序对结果进行汇编。在搜索结果显示屏幕 185 中向用户显示最相关的搜索结果。

本发明提供了提高互联网搜索结果与用户意图的相关性的方法。图 2 示出了根据本发明实施例的方法的实例。应该理解，图 2 所示的具体步骤并不用于限制本发明的范围。对图 2 所示方法的各种修改都处于本发明的范围之内。

用户可以通过输入搜索查询启动互联网搜索（例如，网络搜索）。如图 2 所示，本发明的系统在步骤 221 接收来自用户的搜索查询。在步骤 222，将搜索查询分解成多个单元。

可以将搜索查询分解为多个组成部分（称作单元）。查询处理引擎使用统计学方法将搜索查询分解为多个单元。单元是一个或多个词的序列，通常对应于自然概念，如“纽约市”或“捕食的鸟”。从搜索查询生成概念单元的方法的进一步细节在未决的、共同转让给（co-pending and commonly-assigned）Shyam Kapur 的、于 2003 年 11 月 12 日提交的美国专利申请 10/713,576 中讨论，其内容结合于此作为参考。

概念网络是用于说明相关概念之间的关系的结构。搜索查询中的每个单元都位于概念网络中。概念网络用于识别与搜索查询单元相关的概念。当搜索查询单元已经位于概念网络中之后，在步骤 223，选出与该单元相关的概念网络中的概念。

概念网络可以使用许多方法来连接相关概念。根据本发明的一个实施例，概念网络对作为同义词的概念、具有更具体意义的概念、具有更一般意义的概念、概念的具体实际生活实例、以及发音类似于概念或使用一些相同词的公知术语或名字进行连接。

使用图 3A 所示的实例，如果搜索查询中的单元是摩天楼，则系统在概念网络中定位概念“摩天楼”并识别相关概念。图 3A 示出了用于“摩天楼”的概念网络的实例。在这个概念网络中，将概念“摩天楼”连接至诸如“建筑”和“大厦”的更一般的术语。还将“摩天楼”连接至相似的术语“高层建筑”和摩天楼的著名实例“帝国大厦”。

根据另一个实施例，分析先前提提交的搜索查询以确定概念网络中的相关概念在先前提提交的搜索查询中同时出现的频繁程度。可以通过连接在先前提提交的搜索查询中同时出现的概念来构建概念网络。在步骤 223，从概念网络中选出在之前提交的搜索查询中同时出现最频繁的相关概念。将所有之前提交的搜索查询都存储在存储器中，用作分析。

概念网络可以基于同时出现在所有用户提交的查询中的概念。作为另一个实例，概念网络或其任何部分可以是连接一个特定用户（或一组用户）提交的搜索查询中同时出现的概念的基于会话（session）的概念网络。在步骤 223，选出在之前提交的搜索查询中同时出现最频繁的相关概念。

图 3B 示出了基于会话的概念网络的实例。在图 3B 中，由于特定用户在过去已经提交了将“美洲虎”连接至汽车相关概念的查询，所以将主概念“美洲虎”连接至相关概念“豪华汽车”、“XYZ 汽车公司”、及“车赛”。不同的用户（例如）可能提交了表示他对美洲虎动物感兴趣的在先查询。对于该用户，本发明创建将“美洲虎”连接至诸如猫、动物园、或狩猎旅行的动物相关概念的不同概念网络。

根据本发明的另一个实施例，概念网络可以将一个或多个用户在特定时间内递交的先前查询中同时出现最频繁的概念连接起来。图 3C 示出了时间限制概念网络的实例。在该实例中，将概念“Jane Doe”连接至相关概念“Jane Doe 现场演出”、“Jane Doe 音乐 CD”、及“乐器音乐”。这些相关概念是在特定时间间隔内在先前的搜索查询中与“Jane Doe”同时出现最频繁的概念。特定时间间隔可以是（例如）过去的 24 小时、过去的一周、或过去的一个月。

在图 3C 的实例中，概念网络是以基于在过去 24 小时中最流行的搜索查询的与名为 Joe Doe 的歌手相关的概念为基础的。在接下来的 24 小时内，可以将包括“Jane Doe”的最流行的搜索查询与具有相同名字的政治家相关联。图 3D 示出了如何将用于“Jane Doe”的概念网络变为包括到相关概念“Jane Doe 美国参议员”和“Doe 立法提案权”的连接的概念网络。更新概念网络，以使其包括在最近的具有单元“Jane Doe”的查询中出现最频繁的概念。

根据本发明的实施例，从概念网络中选出最紧密相关的概念。最紧密相关的概念可以是（例如）直接连接至概念网络中的主概念的所有概念。其它概念可以通过直接连接的概念之一间接连接至主概念。图 3C 示出了通过“乐器音乐”连接的概念“Jane Doe”和“小提琴”之间的间接连接的实例。

图 3A 至图 3D 仅示出了连接概念网络中的概念的关系的几个实例。提供的这些实例仅用于阐述的目的，而不用于限制本发明的范围。概念网络中的多个概念之间也可能存在许多其它的关系。

在步骤 223 执行的选择处理可以基于任意标准。例如，可以在步骤 223 从概念网络中选出出现最频繁的前 5 个相关概念。在另一个实例中，在步骤 223 选出出现最频繁的前 50% 或前 25% 的相关概念。根据本发明，可以使用许多其它的选择方法。本文中讨论的实例用于阐述本发明的原理，而不用于限制本发明的范围。

在步骤 224，对在步骤 223 选出的相关概念中的一个或多个执行独立的互联网搜索（例如，网络搜索）。例如，如果搜索查询中存在 4 个单元，并且在步骤 223 为每个单元选出了一个相关概念，则在步骤 224 执行 4 个独立的互联网搜索。根据一个实施例，如果在步骤 223 选出了大量的相关概念，则只对这些概念的子集执行互联网搜索。例如，如果在步骤 223 选出了 20 个概念，则只对与搜索查询中的所有单元相关的前 5 个概念执行互联网搜索。

搜索引擎 175 可以对在步骤 223 选出的概念执行独立的互联网搜索。可以使用任何公知的互联网搜索方法（例如，使用 Google 或 Yahoo! 搜索方法）执行互联网搜索。

对搜索引擎 175 在步骤 224 执行的单独互联网搜索中的每一个检索单独的搜索结果组。搜索引擎 175 通常根据搜索结果与每个相关概念的相关性来对每个互联网搜索的搜索结果进行分类。

在步骤 225，本发明对在步骤 224 执行的互联网搜索中检索到的搜索结果进行重排序。可以将互联网搜索中检索到的搜索结果与对整个原始搜索查询执行的互联网搜索中检索到的搜索结果组合在一起。

将每个搜索结果与原始搜索查询中的单元、单元联合、及单元扩展进行比较。基于搜索结果与原始搜索查询的相关性，为每个搜索结果分配层级或分数。

通过对搜索结果和搜索查询中的单元、联合单元、及扩展单元进行比较，可以确定搜索结果与原始搜索查询的相关性。联合单元是同时出现在搜索查询中、但不足够相关而形成新单元的两个或多个单元。扩展单元是同时出现在搜索查询中、且足够相关而形成新单元的两个或多个单元。

分析搜索结果，以确定来自搜索查询的单元、单元联合、及单元扩展在搜索结果中出现的频率。基于单元、联合单元、和/或扩展单元的实例在搜索结果中出现的频率（或相对频率），为搜索结果分配新的分数。给包括较多的单元、联合单元、和/或扩展单元的实例的搜索结果分配较高的分数。

根据本发明的实施例，对在互联网搜索中检索到的搜索结果进行重排序。根据重排序处理，将重排序分数分配给每个搜索结果。例如，可以通过将搜索引擎 **175** 分配的原始层级分数与基于搜索结果中的查询单元、联合单元、及扩展单元的频率而计算出的新分数相乘，来计算重排列分数。然后基于重排序分数对搜索结果进行分类。

在步骤 225 中，给从某些类型的搜索查询中接收的搜索结果分配较高的分数。例如，可以对在导航查询（navigational query）中检索到的搜索结果分配高于在其它类型的查询中检索到的搜索结果的层级。给导航查询更高的分数是基于导航查询通常能够检索到更为相关的搜索结果的共识。

一旦基于搜索结果与原始搜索查询的相关性而为每个搜索结果分配了重排序分数，则在步骤 225 根据从最高的重排序分数到最低的重排序分数的顺序对搜索结果进行分类。最高的重排序分数表示该内容与原始搜索查询最为相关，最低的重排序分数表示该内容与原始搜索查询最不相关。

本发明通过对与搜索查询中的单元相关的概念匹配的内容进行定位，提高了在互联网搜索中检索到的搜索结果的相关性。如上所述，从概念网络中选出多个概念。然后，将搜索结果与来自基于整个搜索查询的标准互联网搜索的搜索结果相结合，并根据搜索结果与搜索查询的相关性对搜索结果进行分类。

通过使用概念网络提高搜索结果的数量，使得至少一部分搜索结果很可能与搜索查询以及用户意图高度相关，而不论用户意图是什么。因为基于搜索结果与查询的相关性来对搜索结果进行分类，所以最先显示最相关的结果。使用这些方法，本发明能够识别出大量的相关搜索结果。

尽管此处参考本发明的具体实施例来描述本发明，但是可以对本发明进行大量的改变、变化、及替换。在一些实例中，在不偏离上述的本发明的范围的条件下，可以不使用相应特征而实现本发明的一些特征。因此，在不偏离本发明的实质范围 and 精神的条件下，可以对本文中公开的具体配置和方法进行多种改变。本发明不限于所公开的具体实施例，而可以包括所有落入权利要求范围内的实施例和等价物。

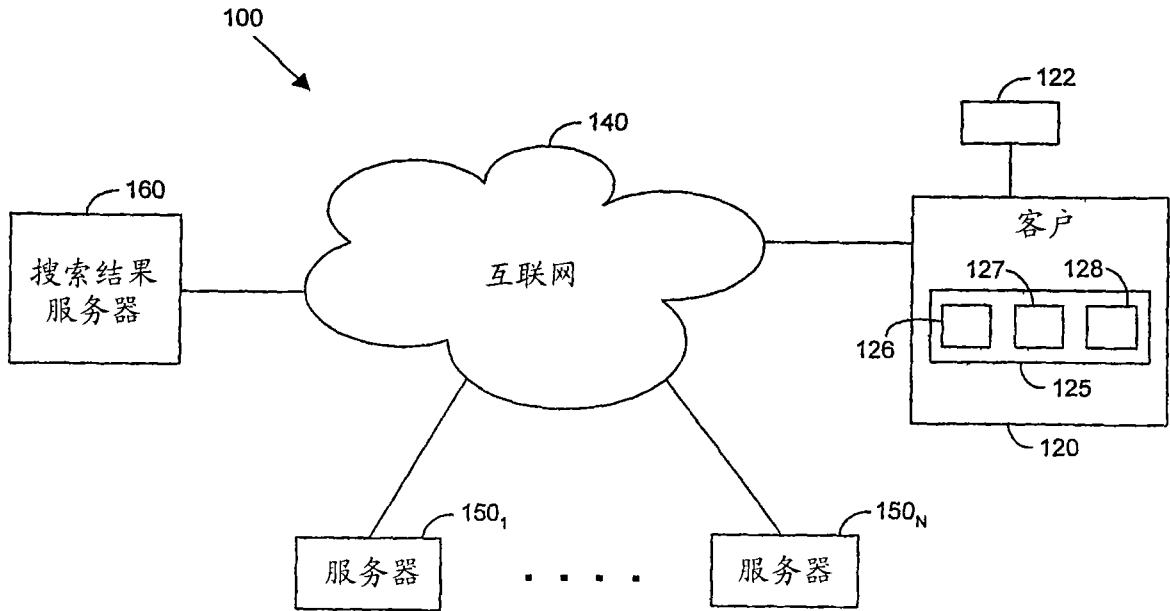


图 1A

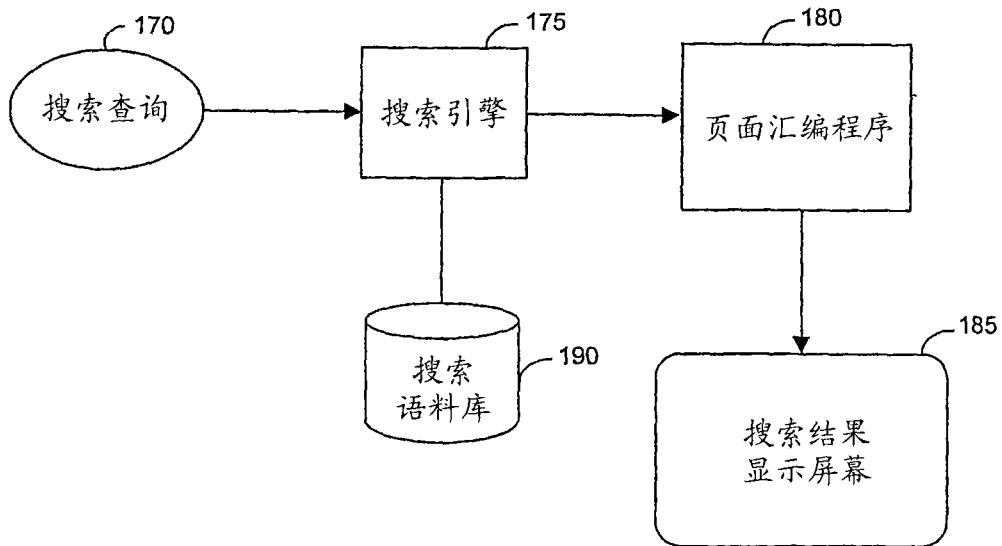


图 1B

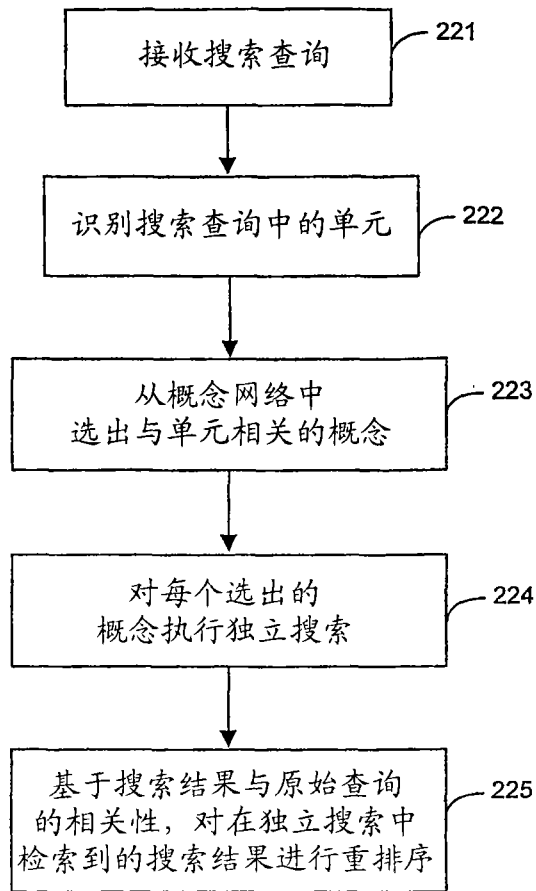


图 2

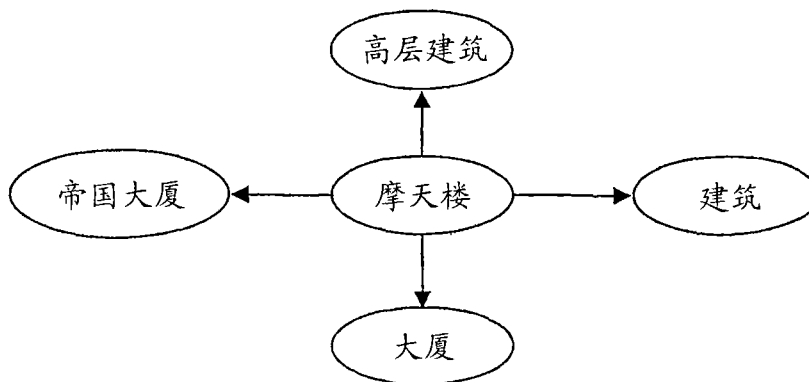


图 3A

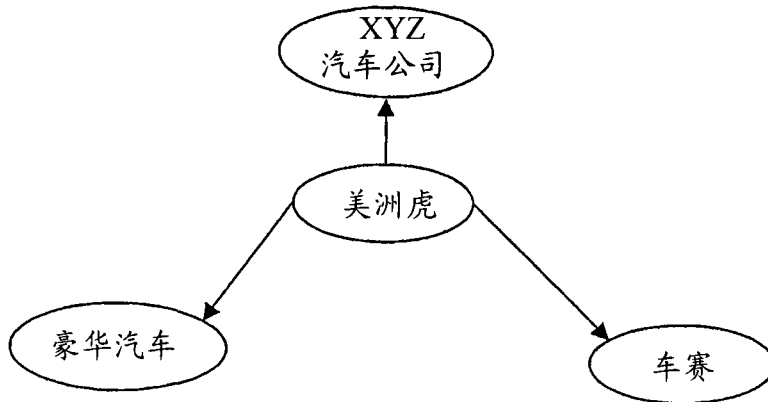


图 3B

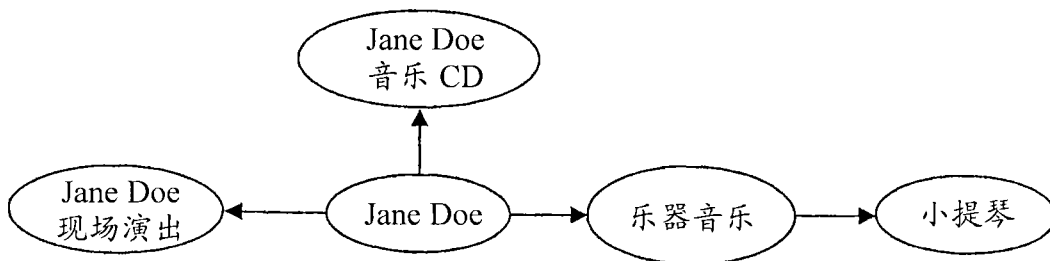


图 3C



图 3D