



(12)发明专利申请

(10)申请公布号 CN 108713186 A

(43)申请公布日 2018. 10. 26

(21)申请号 201680079782.X

(22)申请日 2016.11.04

(30)优先权数据

14/949803 2015.11.23 US

(85)PCT国际申请进入国家阶段日

2018.07.23

(86)PCT国际申请的申请数据

PCT/US2016/060488 2016.11.04

(87)PCT国际申请的公布数据

WO2017/091340 EN 2017.06.01

(71)申请人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 K.肯塔帕迪 H.V.黄 Z.M.罗思

(74)专利代理机构 中国专利代理(香港)有限公司 72001

代理人 刘红 陈岚

(51)Int.Cl.

G06F 7/00(2006.01)

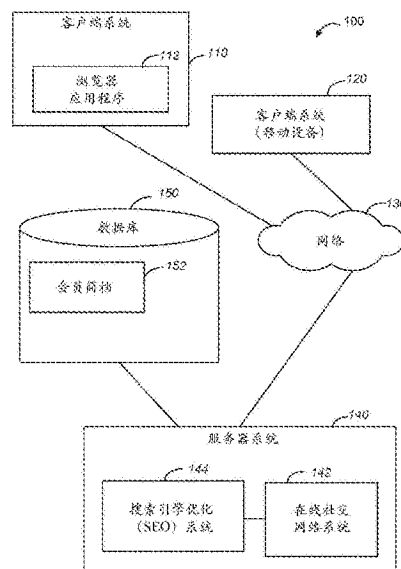
权利要求书3页 说明书12页 附图5页

(54)发明名称

优先化表示位置的搜索词

(57)摘要

一种搜索引擎优化系统设置有在线社交网络系统。在线社交网络系统包括搜索引擎优化(SEO)系统或与其通信,该搜索引擎优化系统被配置成基于它们各自对用户的预测值来优先化表示地理位置的搜索词(潜在搜索词)。职位相关搜索词的值表示为分配给该搜索词的优先级分数。SEO系统使用概率模型为不同的搜索词生成优先级分数,该概率模型考虑了表示搜索词有多可能包括在搜索查询中的值,以及指示由搜索词表示的位置的相对重要性的其它信号。



1. 一种计算机实现的方法,包括:
 - 访问搜索词,所述搜索词包括表示地理位置的位置标识;
 - 确定与所述位置标识相关联的在线社交网络系统的会员数量;
 - 使用至少一个处理器来确定所述搜索词的重要性值,所述重要性值反映了职位相关搜索请求多频繁地包括所述搜索词,并且还反映了与所述位置标识相关联的所述在线社交网络系统的所述会员数量;
 - 利用所述重要性值为所述搜索词生成优先级分数;
 - 基于所述搜索词的所述优先级分数,在由所述在线社交网络系统生成的网页中选择性地包括表示所述地理位置的词;以及
 - 引起在显示设备上呈现所述网页。
2. 根据权利要求1所述的方法,其中除了所述位置标识之外,所述搜索词还包括另一关键字。
3. 根据权利要求1所述的方法,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量是所述在线社交网络系统中的会员简档数量,其包括用于存储当前就业信息的所述会员简档的字段中的所述位置标识。
4. 根据权利要求1所述的方法,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量是所述在线社交网络系统中的会员简档数量,其包括用于存储当前就业信息的所述会员简档的字段中或用于存储过去就业信息的所述会员简档的字段中的所述位置标识。
5. 根据权利要求1所述的方法,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量反映了由所述位置标识表示的地理位置处的所述会员数量的变化。
6. 根据权利要求1所述的方法,其中生成所述重要性分数包括利用一个或多个外部信号。
7. 根据权利要求6所述的方法,其中所述外部信号包括由所述位置标识表示的地理位置的人口规模。
8. 根据权利要求1所述的方法,其中所述职位相关搜索请求针对一个或多个搜索引擎,所述一个或多个搜索引擎包括第三方搜索引擎和由在线社交网络系统提供的搜索引擎,所述第三方搜索引擎和所述在线社交网络系统由不同实体提供。
9. 根据权利要求1所述的方法,其中为所述搜索词生成所述优先级分数包括除了所述重要性值之外还使用为所述搜索词生成的相关性分数,所述相关性分数表示包括所述搜索词的搜索请求有多可能产生相关结果。
10. 根据权利要求1所述的方法,其中所述网页是求职目录页面或表示由所述在线社交网络系统维护的职位发布的网页。
11. 一种计算机实现的系统,包括:
 - 搜索词访问模块,其使用至少一个处理器来实现,以访问搜索词,所述搜索词包括表示地理位置的位置标识;
 - 位置强度评估器,其使用至少一个处理器来实现,以确定与所述位置标识相关联的在线社交网络系统的会员数量;
 - 重要性值生成器,其使用至少一个处理器来实现,以确定所述搜索词的重要性值,所述

重要性值反映了职位相关搜索请求多频繁地包括所述搜索词,并且还反映了与所述位置标识相关联的所述在线社交网络系统的所述会员数量;

优先级分数生成器,其使用至少一个处理器来实现,以利用所述重要性值为所述搜索词生成优先级分数;

网页生成器,其使用至少一个处理器来实现,以在所述在线社交网络系统中生成网页,并基于所述搜索词的所述优先级分数,在所述网页中选择性地包括表示所述地理位置的词;以及

呈现模块,其使用至少一个处理器来实现,以引起在显示设备上呈现所述网页。

12. 根据权利要求11所述的系统,其中除了所述位置标识之外,所述搜索词还包括另一关键字。

13. 根据权利要求11所述的系统,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量是所述在线社交网络系统中的会员简档数量,其包括用于存储当前就业信息的所述会员简档的字段中的所述位置标识。

14. 根据权利要求11所述的系统,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量是所述在线社交网络系统中的会员简档数量,其包括用于存储当前就业信息的所述会员简档的字段中或用于存储过去就业信息的所述会员简档的字段中的所述位置标识。

15. 根据权利要求11所述的系统,其中与所述位置标识相关联的所述在线社交网络系统的所述会员数量反映了由所述位置标识表示的地理位置处的所述会员数量的变化。

16. 根据权利要求11所述的系统,其中生成所述重要性分数包括利用一个或多个外部信号。

17. 根据权利要求16所述的系统,其中所述外部信号包括由所述位置标识表示的地理位置的人口规模。

18. 根据权利要求11所述的系统,其中所述职位相关搜索请求针对一个或多个搜索引擎,所述一个或多个搜索引擎包括第三方搜索引擎和由在线社交网络系统提供的搜索引擎,所述第三方搜索引擎和所述在线社交网络系统由不同实体提供。

19. 根据权利要求11所述的系统,其中所述优先级分数生成器用于为所述搜索词生成所述优先级分数包括除了所述重要性值之外还使用为所述搜索词生成的相关性分数,所述相关性分数表示包括所述搜索词的搜索请求有多可能产生相关结果。

20. 一种机器可读的非暂时性存储介质,其具有可由机器执行以使所述机器执行以下操作的指令数据,所述操作包括:

访问搜索词,所述搜索词包括表示地理位置的位置标识;

确定与所述位置标识相关联的在线社交网络系统的会员数量;

确定所述搜索词的重要性值,所述重要性值反映了职位相关搜索请求多频繁地包括所述搜索词,并且还反映了与所述位置标识相关联的所述在线社交网络系统的所述会员数量;

利用所述重要性值为所述搜索词生成优先级分数;

基于所述搜索词的所述优先级分数,在由所述在线社交网络系统生成的网页中选择性地包括表示所述地理位置的词;以及

引起在显示设备上呈现所述网页。

优先化表示位置的搜索词

[0001] 本PCT申请要求于2015年11月23日提交的标题为《优先化表示位置的搜索词(PRIORITIZING SEARCH TERMS REPRESENTING locations)》的美国专利申请序列号14/949,803的申请日的权益,其全部内容以引用方式并入本文中。

技术领域

[0002] 本申请涉及软件和/或硬件技术的技术领域,并且在一个示例实施例中,涉及用于在在线社交网络系统背景下优先化表示相应组织机构的搜索词的系统和方法。

背景技术

[0003] 在线社交网络可被视为连接虚拟空间中的人的平台。在线社交网络可以是基于网络的平台,诸如例如社交联网网站,并且可以由用户经由网络浏览器或经由设置在移动电话、平板计算机等等上的移动应用程序进行访问。在线社交网络可以是专门为企业界设计的面向企业的社交网络,其中注册会员建立并记录他们认识并信任的专业人员的网络。每个注册会员可以由会员简档表示。会员简档可以由一个或多个网页表示,或呈XML(可扩展标记语言)、JSON(JavaScript对象表示法)或类似格式的会员信息的结构化表示。社交联网网站会员的会员简档网页可能会强调相关会员的就业经历和教育。在线社交网络可以存储包括一个或多个利于会员以及非会员的职位相关搜索的组件。

附图说明

[0004] 在附图中以示例而非限制性地示出本发明的实施例,其中相同的附图标记表示相似的元件,并且其中:

图1是网络环境的图形表示,在该网络环境内可以实现用于在在线社交网络系统中优先化表示相应地理位置的搜索词的示例方法和系统;

图2是根据一个示例实施例的用于在在线社交网络系统中优先化表示相应地理位置的搜索词的系统的方框图;

图3是示出根据示例实施例的在在线社交网络系统中优先化表示相应地理位置的搜索词的方法的流程图;

图4是用于按位置导航求职目录的用户界面的示例表示;以及

图5是呈计算机系统的形式的示例机器的图形表示,该计算机系统内可以执行用于使机器执行本文讨论的任何一个或多个方法的一组指令。

具体实施方式

[0005] 描述了用于在在线社交网络系统中优先化表示相应地理位置的搜索词的方法和系统。在以下描述中,为了解释的目的,阐述了许多具体细节以便提供对本发明的实施例的透彻理解。然而,对于本领域技术人员来说显而易见的是,可以在没有这些具体细节的情况下实践本发明。

[0006] 如本文所使用,术语“或”可以被解释为包括性或排它性意义。类似地,术语“示例性”仅仅意味着某些事物的示例或范例,并不一定是实现目标的优选或理想手段。另外,虽然下面讨论的各种示例性实施例可以利用基于Java的服务器和相关环境,但是本发明中仅仅为了清楚起见而给出这些实施例。因此,包括各种系统体系结构的任何类型的服务器环境都可以采用本文描述的以应用为中心的资源系统和方法的各种实施例,并且被认为是在本发明的范围内。

[0007] 为了本说明书的目的,短语“在线社交联网应用程序”和“在线社交网络系统”可以被称为短语“在线社交网络”并且可以与其互换使用,或者仅仅是“社交网络”。还应该注意的,在线社交网络可以是任何类型的在线社交网络,诸如例如专业网络、基于兴趣的网络或允许用户作为注册会员加入的任何在线社交网络。为了描述的目的,在线社交网络的注册会员可以简称为会员。

[0008] 在线社交网络的每个会员由会员简档(也称为会员简档或简称简档)表示。会员简档可以与指示该会员与社交网络的其他会员的人脉关系的社交链接相关联。会员简档还可以包括来自在线社交网络的其他会员的评论或推荐或者与其相关联,并与其它网络资源(诸如,例如出版物等)链接。如上所述,在线社交网络系统可以设计成允许注册会员建立和记录他们认识和信任的专业人员的网络。社交网络中的任何两个会员可以指示他们相互愿意在社交网络的背景下“连接”,其中他们可以查看彼此的简档、简档推荐和对彼此的认可,或者经由社交网络进行联系。在社交网络的背景下连接的会员可以被称为彼此的“人脉”,并且他们相应的简档与指示连接这两个简档的相应人脉链接相关联。

[0009] 社交网络会员的简档信息可以包括各种信息,诸如例如会员的名字、会员的当前和先前的地理位置、会员的当前和先前的就业信息、与会员的教育有关的信息、关于会员的专业造诣的信息、出版物、专利等。社交网络会员的简档信息还可以包括关于会员的专业技能的信息。可以存在于简档中的特定类型的信息(诸如,例如公司、行业、工作职位等)被称为简档属性。特定会员简档的简档属性可以具有一个或多个值。例如,简档属性可以表示公司,并被称为公司属性。特定简档中的公司属性可以具有表示已经雇用相关会员的公司的相应标识的值。简档属性的其它示例是行业属性和区域属性。会员简档中的行业属性和区域属性的相应值可以指示相关会员受雇于旧金山湾区的银行业。

[0010] 在线社交网络系统不仅可以维护会员的相应简档,还可以维护诸如公司、大学等的组织机构的简档。例如,公司的简档可以与公司网页相关联,并且包括关于公司的信息。在线社交网络系统中的公司网页可以包括视觉控件,例如用户可以点击以指示他们想要“关注”公司简档的“关注”按钮。表示“关注”公司简档的会员的会员简档可以包括指示该会员简档与公司简档之间的这种关系的链接。这种关系可以在在线社交网络系统的背景下表示为例如关于公司职位空缺、公司组织机构中的变化、管理团队中的新会员等等的新闻和通知可以例如经由会员的新闻馈送网页等传递给与会员简档相关联的会员。

[0011] 在线社交网络系统还维护关于职位发布的信息。为了本说明书的目的,也称为“职位”的职位发布是电子存储的实体,其包括雇主可以关于职位空缺发布的信息。职位发布中的信息可以包括例如行业、公司、职位、需求和/或期望的技能、职位的地理位置等。在线社交网络系统可以被配置成将会员简档与职位发布进行匹配,使得使用例如所谓的职位推荐系统将那些已被标识为由特定会员简档表示的会员可能感兴趣的职位发布在显示设备上

呈现给会员进行查看。职位推荐系统将某些职位发布标识为会员可能感兴趣,并按照与相关联的会员简档相关的顺序将这些职位发布呈现给会员。会员可以通过在搜索框中输入搜索词并检查返回的搜索结果来访问职位发布。搜索词可以包括一个或多个表示压制性、职称、专业技能、公司名称、地理位置等的关键字或短语。访问职位发布的另一种方式是导航到表示求职目录的网页并点击(或以其它方式参与)对应于感兴趣的搜索词(例如,公司名称或地理位置)的链接,这会引发呈现对包含该搜索词的职位发布的引用。在图4中示出了用于按位置导航求职目录的用户界面400的示例表示。

[0012] 虽然在线社交网络系统可以有益地用于帮助其会员进行他们的求职,但是可以被认为是主动求职者的人可能不一定是在线社交网络系统的会员。与此同时,即使他们还不是会员,主动求职者也可能在使用在线搜索引擎进行搜索而作为结果返回在线社交网络系统维护的职位发布时受益。在线社交网络系统可以被配置成向用户提供丰富的求职经历,而不管他们是否有在线社交网络系统的会员身份,其中源自在线社交网络系统的JSERP(求职结果页面)排名在搜索结果的顶部。在一个实施例中,在线社交网络系统被配置成基于它们各自对JSERP排名的预测贡献来优先化搜索词(潜在搜索词)。职位相关搜索词的值可以表示为分配给该搜索词的优先级分数。术语搜索词将被理解为意指单词或由多于一个单词组成的短语。被优先化的搜索词可以表示例如专业或职称,可以提供就业的组织机构,诸如例如公司、律师事务所、大学等,以及地理位置。搜索词的一些示例是“护士”、“电气工程师”、“产品经理”、“社交网络公司”、“旧金山湾区”等。将注意的是,虽然可以提供就业的组织机构可以是公司以外的实体,但为了本说明书的目的,术语“公司”将用于指可以提供就业的任何组织机构。

[0013] 考虑到与潜在在工作场所相关联的大量地理位置,理解JSERP相对于彼此的值是有益的-换句话说,针对列出了在其它位置可到职的职位的JSERP,确定列出了在特定位置可到职的职位的JSERP的相对优先级。能够确定不同<位置,关键字>对的相对优先级可能是有益的。例如,与包括<印度班加罗尔,石油>对(列出了在班加罗尔的石油行业的职位)的JSERP相比,用户可能对包括<德克萨斯州,石油>对(列出了德克萨斯州石油行业的职位)的JSERP(求职结果页面)更感兴趣。再举另一个示例,旧金山湾区的地理位置可能在科技行业的关键字上排名高。

[0014] 用于优先化与位置相关联的JSERP的现有解决方案完全基于职位计数。例如,在我们的职位目录中,在列出头等位置时,一直包括返回最多职位数量的位置。问题在于,仅仅因为一个位置具有很多职位空缺并不意味着对应的JSERP很有价值,或者很大一部分客人关心位置。针对优先化与不同关键字相关联的位置,这个问题仍然存在。

[0015] 优先化搜索词的一种方法是基于广告一个位置的职位的数量。这种方法本身可能并不总是最优的,因为它可能只是因为他们有很多职位空缺而导致人员配置公司的高排名。然而,仅仅因为特定的地理位置具有很多职位空缺并不意味着对应的JSERP很有价值,或者很大一部分求职者关心位置。在一些情况下,在优先化与用作搜索词的不同关键字相关联的位置时可能还需要解决这个问题。

[0016] 在一个示例实施例中,在线社交网络系统包括搜索引擎优化(SEO)系统或与其通信,该搜索引擎优化系统被配置成计算特定搜索词的相应优先级分数并且使用这些优先级分数来增强用户的在线求职经历。可以自动或手动选择要评分的一组搜索词,并将其作为

一组搜索词存储在数据库中。SEO系统可以被配置成使用概率模型来为不同搜索词生成优先级分数,该概率模型考虑了表示搜索词有多可能包括在搜索查询中的值和表示包括搜索词的搜索有多可能产生相关的结果的值。表示搜索词有多可能包括在搜索查询中的值可以被称为流行度分数。表示包括搜索词的搜索查询有多可能产生相关结果的值可被称为相关性分数。概率模型可以有益地用于表示地理位置的搜索词以及用于包括地理位置和关键字的组的搜索词。

[0017] 表示可以提供就业的位置的搜索词 w 可以被称为位置搜索词。在一个实施例中,SEO系统可以被配置成为位置搜索词 w 生成重要性值 $\text{Imp}(w)$ 。位置搜索词 w 的重要性值可以通过组合来自对应于以下方面的数据源的信号来生成:流行度、强度和外部信号。SEO系统可以被配置成通过检查关于在线社交网络系统内的搜索的搜索量以及通过检查关于一个或多个第三方搜索引擎内的搜索的搜索量来利用关于位置搜索词有多可能发布在搜索查询中的信息来为位置搜索词生成流行度值,如在说明书中稍后进一步详细描述。

[0018] 由SEO系统用于为位置搜索词生成重要性值的强度信号包括以下中的一个或多个:由位置搜索词表示的位置处的当前雇员的数量,由位置搜索词表示的位置处的在线社交网络系统的会员数量(在线社交网络系统中的会员简档的数量,其包括用于存储当前就业信息的会员简档的字段中的位置标识),在该位置工作过的会员数量,在特定时间段内(例如,在过去一年内或过去18个月内)在该位置工作过的会员数量。在该位置工作过的会员数量可以被确定为在线社交网络系统中的会员简档的数量,其包括在用于存储当前就业信息的会员简档的字段中,或者在用于存储具有晚于预定日期的过去就业结束日期的过去就业信息的会员简档的字段中的位置标识。SEO系统用于为位置搜索词生成重要性值的另一个强度信号是某个时间段内(例如,在过去一年内或过去18个月内)该位置的会员数量的变化(增长或下降)。由SEO系统用来为位置搜索词生成重要性值的外部信号包括例如该位置的人口规模以及特定时间段内(例如,在过去一年内或过去18个月内)该位置处人口规模的变化(增长或下降)。表示强度信号和外部信号的信息可以从各种来源获得,例如公共和私人数据库,以及由在线社交网络系统存储的数据。

[0019] SEO系统可以被配置成生成(位置,关键字)对的重要性值 $\text{Imp}(l, w)$ 。(位置,关键字)对的重要性值可以通过组合来自对应于以下方面的数据源的信号来生成:流行度、强度和外部来源。SEO系统可以被配置成通过检查关于在线社交网络系统内的搜索的搜索量以及通过检查关于一个或多个第三方搜索引擎内的搜索的搜索量来利用关于位置搜索词有多可能发布在搜索查询中的信息来为(位置,关键字)对生成流行度值,如在说明书中稍后进一步详细描述。

[0020] 由SEO系统用来为(位置,关键字)对生成重要性值的强度信号包括以下中的一个或多个:在该位置居住或工作以及在他们的会员简档中列出技能或职称内的关键字的在线社交网络的当前会员数量,在该位置居住或工作过的在线社交网络系统的会员以及在他们的会员简档中列出技能或职称内的关键字的人数,在特定时间段内(例如,在过去一年或过去18个月内)在该位置工作过以及在他们的会员简档中列出技能或职称内的关键字的在线社交网络系统的会员数量,以及在特定时间段内(例如,在过去一年或过去18个月内)在该位置工作过以及在他们的会员简档中列出技能或职称内的关键字的在线社交网络系统会员数量的变化(增长或下降)。

[0021] 由SEO系统用来为(位置,关键字)对生成重要性值的外部信号可以包括该位置的人口规模,以及在特定时间段内(例如,在过去一年内或过去18个月内)该位置的人口规模的变化(增长或下降)。

[0022] 为位置搜索词生成的相应重要性值可以用于生成相应优先级分数。在一些实施例中,位置搜索词的优先级分数通过例如使用下面示出的等式1将其相关性分数乘以其重要性分数来生成。

[0023] 等式(1):

$$\text{PriorityScore}(l) = \text{Pr}(\text{RELEVANT} \& l) = \text{Imp}(l) * \text{Pr}(\text{RELEVANT}/l),$$

其中 w 是搜索词, $\text{Imp}(l)$ 是表示由搜索词 w 表示的关键字的重要性的值,而 $\text{Pr}(\text{RELEVANT}/l)$ 是表示搜索词 l 的相关性分数的概率。

[0024] 对于(位置,关键字)对 (l, w) ,其优先级分数可以通过例如使用下面示出的等式2将其相关性分数乘以其重要性分数来生成。

[0025] 等式(2):

$\text{PriorityScore}(l, w) = \text{Pr}(\text{RELEVANT} \& l, w) = \text{Imp}(l, w) * \text{Pr}(\text{RELEVANT}/l, w)$,其中 (l, w) 是(位置,关键字)对, $\text{Imp}(l, w)$ 是表示(位置,关键词)对的重要性的值,而 $\text{Pr}(\text{RELEVANT}/l, w)$ 是表示(位置,关键词)对的相关性分数的概率。

[0026] 使用为位置搜索词生成的相应优先级分数来确定在职位目录中要突出显示哪些位置(例如,确定哪些位置要包括在每个字母表下),确定哪些JSERP登录页面(列出某些位置的职位)要包括在职位目录中,以及确定哪些JSERP登录页面包括到提交给一个或多个第三方搜索引擎(诸如,例如Google®或Bing®)的站点地图中。

[0027] 当为(位置,关键字)对生成优先级分数时,使用它们相应的优先级分数来确定在职位目录中要突出显示哪些(位置,关键字)对,对应于(位置,关键字)对的哪些JSERP登录页面要包括在职位目录中,以及确定将哪些JSERP登录页面要包括到提交给一个或多个第三方搜索引擎的站点地图中。

[0028] 返回到用于为搜索词生成流行度分数并且计算搜索词有多可能包括在搜索查询中的概率的过程的讨论,以便为特定的搜索词 w (也称为主题搜索词或仅仅是搜索词,并且可以是位置搜索词或任何其它搜索词)生成流行度分数 $\text{Pr}(w)$,SEO系统监测包括主题搜索词的职位相关搜索。在一个实施例中,SEO系统在一段时间内监测由一个或多个特定目标第三方搜索引擎(例如,Google®、Yahoo!®)执行的所有职位相关搜索,并且在一些实施例中监测在在线社交网络系统内执行的相关搜索。这些来源中的每一个关于主题搜索词 w 的监测结果用于生成相应的间歇流行度值 $P_j(w)$,其中 j 是来自 k 个数据源的第 j 个数据源。例如,Google®数据源的 $P_j(w)$ 可以基于包括搜索词 w 的职位相关搜索的百分比来确定。

[0029] 当在线社交网络系统被用作用于确定 $P_j(w)$ 的数据源时,SEO系统将每个搜索请求视为职位相关搜索。当第三方搜索引擎被用作用于确定 $P_j(w)$ 的数据源时,SEO系统可以首先确定搜索的意图是否与求职有关并且仅考虑那些已经被标识为职位相关的搜索,而忽略那些尚未被标识为职位相关的搜索。将针对第三方搜索引擎的求职标识为职位相关可以通过在搜索请求中检测存在已经被标识为意图指标的附加词(例如,单词“职位”或“职业”)来实现。

[0030] 因为基于从不同来源获得的数据生成的流行度值可以是不同的尺度,所以SEO系

统可以被配置成首先对给定搜索词 w 的间歇流行度值 $P_j(w)$ 进行归一化,然后将归一化的流行度值聚合来得出流行度分数 $Pr(w)$ 。这种方法可以由下面示出的等式(3)表示。

[0031] 等式(3)

$$Pr(w) = \text{popularityAggregateFunction}(\text{normFunction}_1(P_1(w)), \text{normFunction}_2(P_2(w)), \dots, \text{normFunction}_k(P_k(w)))$$

[0032] 在一个实施例中,针对每个间歇流行度值使用不同的归一化函数(针对 $P_1(w)$ 的 normFunction_1 ,针对 $P_2(w)$ 的 normFunction_2 等)。在上面的等式(2)中表示为 $\text{popularityAggregateFunction}$ 的聚合函数可被选择为从特定百分位范围(例如,从第20百分位到第80百分位)选择的一组归一化流行度值中的最大值、中值、平均值、均值中的一个。在一些实施例中,聚合函数可以是通过地面实况数据学习的机器学习模型(诸如逻辑回归)的输出。归一化函数 $\text{normFunction}_j(P_j(w))$ 将每个间歇流行度值 $P_j(w)$ 映射到相同的区间。

[0033] 例如,归一化函数 $\text{scale}(P_j(w))$ 可以将每个间歇流行度值 $P_j(w)$ 映射到区间 $[0, 1]$ 并利用三个百分位值-下阈值(α -百分位值)、中值(50百分位值)和上阈值(β 百分位值)。归一化函数执行从间歇流行度值到 $[0, 1]$ 的分段线性映射。如果间歇流行度值小于下阈值,则将其映射为0。对大于或等于下阈值且小于或等于中值的间歇流行度值执行线性缩放至 $[0, 0.5]$ 。对大于或等于中值且小于或等于上阈值的间歇流行度值执行线性缩放至 $[0.5, 1]$ 。如果间歇流行度值大于上阈值,则将其映射为1。然后可以使用来自该组归一化流行度值的最大值作为聚合函数: $\max(\text{尺度}(P_1(w)), \text{尺度}(P_2(w)), \dots, \text{尺度}(P_k(w)))$ 。应用于每个间歇流行度值的缩放可能不同,因为对于每个间歇流行度类型,百分位值可能不同。

[0034] 在一些实施例中,SEO系统可以被配置成使用位置搜索词的重要性值作为该搜索词的优先级分数。然而在其它实施例中,如上所述,可以使用为位置搜索词生成的相应重要性值来导出相应的对应优先级分数,例如,通过将表示重要性值的值乘以表示为位置搜索词生成的相关性分数的值,如以上等式(1)所表示。

[0035] 如上所述,表示包括搜索词的搜索有多可能产生相关结果的值可以称为相关性分数。在一个实施例中,SEO系统可以被配置成使用多个相关性指标指标来确定搜索词 w 的相关性分数 $Pr(\text{RELEVANT}/w)$ 。

[0036] 搜索词的相关性的指标的一个示例是响应于包括搜索词并且源自于在线社交网络系统的查询而返回的搜索结果的数量。搜索词的相关性的另一个指标可能与分配给返回结果的相应质量分数有关。例如,第三方搜索引擎响应于包括搜索词的查询返回搜索结果。返回的结果每个都有由搜索引擎分配给它的质量分数。源自于在线社交网络系统的那些返回的搜索结果的质量分数的总和可以被SEO系统用作该搜索词的相关性的指标中的一个。可以基于监测关于响应于包括搜索词的查询返回并且源自于在线社交网络系统的搜索结果的用户参与信号来获取搜索词的相关性的又一个指标。例如,关于响应于包括搜索词的查询返回并且源自于在线社交网络系统的搜索结果,SEO系统可以监测并记录信号,诸如特定数量的最高职位结果的点击率(CTR)。这些信号可以聚合在单个职位结果(JSERP)上,以获取该JSERP的组合用户参与分数。例如,SEO可以利用与搜索词相关联的JSERP的总CTR作为搜索词的相关性的另一个指标。而且,SEO可以利用特定数量的最高职位结果的平均停留时间(在转到不同页面或结束会话之前查看职位说明/细节所花费的时间)以及与搜索词相关联的JSERP的总停留时间作为搜索词的相关性的其它指标。然后可以利用该用户参与分数

来导出搜索词的相关性分数。通过检查在线社交网络系统中的会员简档可以获取搜索词的相关性的另一个指标。例如,SEO系统可以确定在会员简档中多频繁地使用搜索词,例如指定当前或过去的就业职位。

[0037] 使用关于特定搜索词 w 的不同指标来生成相应的间歇相关性值 $P_j(\text{RELEVANT}/w)$,其中 j 是来自 k 个数据源的第 j 个数据源。因为基于从不同来源获取的数据生成的相关性值可以是不同的尺度,所以SEO系统可以被配置成首先对给定搜索词 w 的间歇相关性值 $P_j(\text{RELEVANT}/w)$ 进行归一化,然后将归一化的相关性值聚合来得出相关性分数 $Pr(\text{RELEVANT}/w)$ 。这种方法可以由下面示出的等式(4)表示。

[0038] 等式(4)

$$Pr(\text{RELEVANT}/w) = \text{relevanceAggregateFunction}(\text{normFunction}_1(P_1(\text{RELEVANT}/w)), \text{normFunction}_2(P_2(\text{RELEVANT}/w)), \dots, \text{normFunction}_n(P_n(\text{RELEVANT}/w)))$$

[0039] 针对每个间歇相关性值可以使用不同的归一化函数(针对 $P_1(\text{RELEVANT}/w)$ 的 normFunction_1 ,针对 $P_2(\text{RELEVANT}/w)$ 的 normFunction_2 等)。此外,在一些实施例中,这些归一化函数还与用于相关性分数计算的函数不同。在上面的等式(3)中表示为 $\text{relevanceAggregateFunction}$ 的聚合函数可被选择为从特定百分位范围(例如,从第20百分位到第80百分位)选择的一组归一化流行度值中的最大值、中值、平均值、均值中的一个。在一些实施例中,聚合函数可以通过地面实况数据学习的机器学习模型(诸如逻辑回归)的输出。在一些实施例中,归一化函数 $\text{normFunction}_j(P_j(\text{RELEVANT}/w))$ 将每个间歇相关性值 $P_j(\text{RELEVANT}/w)$ 映射到相同的区间并利用两个阈值-下阈值(ϵ_1)和上阈值(ϵ_2)。

[0040] 例如,关于间歇 $P_j(\text{RELEVANT}/w)$ 是响应于包括搜索词的查询而返回并且源自于在线社交网络系统的搜索结果的数量,归一化函数 $\text{scale}(P_j(\text{RELEVANT}/w))$ 使用阶梯函数将职位结果计数映射到 $[0, 1]$:如果职位结果计数小于下阈值,则为0;如果职位结果计数大于上阈值,则为1。如果职位结果计数大于下阈值并小于上阈值,则其归一化值如下面等式(5)中所示进行计算。

[0041] 等式(5)

$$\text{scale}(P_j(\text{RELEVANT}/w)) = (P_j(\text{RELEVANT}/w) - \epsilon_1) / (\epsilon_2 - \epsilon_1)$$

[0042] 在另一个示例中,在间歇 $P_j(\text{RELEVANT}/w)$ 是源自于在线社交网络系统的那些返回的搜索结果的质量分数的总和的情况下,页面和搜索词 w 的组合质量分数是使用诸如特定百分位(例如,从第20百分位到第80百分位)之间的值的最大值、中值、平均值、均值等的聚合函数导出的。聚合函数也可以考虑位置折扣,即为顶部位置的求职结果提供更大的权重。如上所解释,在一些实施例中,为职位相关搜索词生成的相应相关性分数可以用于例如通过将表示搜索词的相关性分数的值乘以表示该同一搜索词的流行度分数的值来导出相应优先级分数,如上面等式(1)表示。

[0043] 在确定(位置,关键字)对 (l, w) 的优先级分数时,除了重要性值之外,SEO系统还使用针对该(位置,关键字)对计算的相关性分数,如由上面的等式(2)表示。SEO系统使用多个相关性指标来确定(位置,关键字)对 (l, w) 的相关性得分 $Pr(\text{RELEVANT}/l, w)$ 。(位置,关键字)对的相关性指标的一个示例是响应于包括关键字和位置作为搜索词并且源自于在线社交网络系统的查询而返回的搜索结果的数量。(位置,关键字)对的相关性的另一个指标可能与分配给返回结果的相应质量分数有关。例如,如上所述,第三方搜索引擎响应于包括关

键字和位置作为搜索词的查询返回搜索结果。返回的结果每个都有由搜索引擎分配给它的质量分数。源自于在线社交网络系统的那些返回的搜索结果的质量分数的总和可以被SEO系统用作该(位置,关键词)对的相关性指标中的一个。可以基于监测关于响应于包括关键字和位置作为搜索词的查询返回并且源自于在线社交网络系统的搜索结果的用户参与信号来获取(位置,关键字)对的相关性的又一个指标。监测和记录的参与信号可能与特定数量的最高职位结果的点击率(CTR)有关。这些信号可以聚合在单个职位结果(JSERP)上,以获取该JSERP的组合用户参与分数。通过监测用户对搜索结果的参与而导出的其它信号可能包括与该(位置,关键字)对相关性的JSERP的总CTR,特定数量的最高职位结果的平均停留时间,以及与该(位置,关键字)对相关性的JSERP的总停留时间。通过检查在线社交网络系统中的会员简档可以获取该(位置,关键字)对的相关性的另一个指标。例如,SEO系统可以确定在会员简档中多频繁地使用搜索词,例如指定雇主位置的当前或过去位置。

[0044] 因为使用不同的相关性指标关于(位置,关键字)对 (l, w) 产生相应的间歇相关性值 $P_j(\text{RELEVANT}/l, w)$,其中 j 是来自 k 个数据源的第 j 个数据源,所以将这些间歇相关性值进行归一化,然后进行聚合来得出相关性分数 $P_r(\text{RELEVANT}/l, w)$ 。关于为位置搜索词生成的间歇相关性值进行归一化和聚合,可以如上所述使用归一化和聚合方法。

[0045] 示例性搜索词优先化系统可以在图1中示出的网络环境100的背景下实现。如图1中所示,网络环境100可以包括客户端系统110和120以及服务器系统140。客户端系统120可以是移动设备,诸如,例如移动电话或平板计算机。在一个示例实施例中,服务器系统140可以托管在线社交网络系统142。如上所解释,在线社交网络的每个会员由包含关于该会员的个人和专业信息的会员简档表示,并且该会员简档可以与指示该会员与在线社交网络中的其他会员简档的连接的社交链接相关联。会员简档和相关信息可以作为会员简档152存储在数据库150中。

[0046] 客户端系统110和120能够利用例如在客户端系统110上执行的浏览器应用程序112或在客户端系统120上执行的移动应用程序,经由通信网络130访问服务器系统140。通信网络130可以是公共网络(例如,互联网、移动通信网络或能够传递数字数据的任何其它网络)。如图1中所示,服务器系统140还托管搜索引擎优化(SEO)系统144。如上所解释,SEO系统144可以被配置成基于它们对JSERP排名的相应预测贡献来区分搜索词的优先级。职位相关搜索词的值表示为分配给该搜索词的优先级分数。在不同的实施例中,SEO系统144使用概率模型为搜索词生成优先级分数,该概率模型考虑了表示搜索词有多可能包括在搜索查询中的值和/或表示包括搜索词的搜索有多可能产生相关的结果的值,以及其它信号,如上所述。在图2中示出了对应于SEO系统144的示例性搜索词优先化系统。

[0047] 图2是用于在图1的在线社交网络系统142中优先化搜索词的系统200的方框图。如图2中所示,系统200包括搜索词访问模块210、位置强度评估器220、重要性值生成器230以及优先级分数生成器240。搜索词访问模块210可以被配置成访问包括表示地理位置的位置标识的搜索词。在一些实施例中,除了位置标识的搜索词之外,搜索词还包括另一关键字。

[0048] 位置强度评估器220可以被配置成确定与位置标识相关联的在线社交网络系统的会员数量。与位置标识相关联的会员数量可以被确定为在线社交网络系统142中的会员简档数量,其包括用于存储当前就业信息的会员简档字段中的位置标识,被确定为在线社交网络系统中的会员简档数量,其包括用于存储当前就业信息的会员简档字段中或用于存储

过去就业信息的会员简档字段中的位置标识,或者被确定为反映了由位置标识表示的地理位置处的会员数量的变化的值。

[0049] 该重要性值生成器230使用该信息来为搜索词生成重要性值。重要性值反映了职位相关搜索请求多频繁地包括搜索词。在一个实施例中,重要性值生成器230使用反映了职位相关搜索请求多频繁地包括搜索词的重要性值以及反映了与位置标识相关联的在线社交网络系统的会员数量的值来确定搜索词的重要性值。除了位置标识之外,如果搜索词还包括另一关键字,则搜索词的重要性值反映了关联(位置,关键字)对的重要性。

[0050] 在确定搜索词的重要性值时考虑的职位相关搜索请求可以是一个或多个搜索引擎的那些请求,诸如由第三方提供的搜索引擎以及由在线社交网络系统142提供的搜索引擎。为了为搜索词生成重要性值,重要性值生成器230使用如上所述的其它信号,诸如由位置标识所标识的地理位置的人口规模。

[0051] 优先级分数生成器240可以被配置成利用重要性值来为搜索词生成优先级分数。除了重要性值之外,优先级分数生成器240还可以使用为搜索词生成的相关性分数。相关性分数表示包括搜索词的搜索请求有多可能产生相关结果。例如,可以通过计算重要性值和相关性分数的乘积来生成搜索词的优先级分数。优先级分数生成器240还可以被配置成基于由在线社交网络142维护的会员简档的特定字段(例如,过去的或当前就业字段)中的主题搜索词的出现频率来调整优先级分数。除了位置标识之外,如果搜索词还包括另一关键字,则优先级分数生成器240为搜索词生成的优先级分数是关联(位置,关键字)对的优先级值。

[0052] 还在图2中示出了网页生成器250和呈现模块260。网页生成器250可以被配置成在在线社交网络系统142中生成网页并且基于搜索词的优先级分数在网页中选择性地包括表示由位置标识所标识的地理位置的词。例如,具有较高优先级分数的搜索词可以包括到表示求职目录的网页中,而具有较低优先级分数的搜索词可以从该网页中省略。作为另一示例,包括具有较高优先级分数的一个或多个搜索词的那些职位发布可以包括在提交给一个或多个第三方搜索引擎的站点地图中,而不包括任何较高评分的搜索词的那些职位发布可以从这样的站点地图中省略。呈现模块260可以被配置成引起在显示设备上呈现各种网页(例如,表示会员简档的网页或表示求职目录的网页)。可以参考图3描述由系统200执行的一些操作。

[0053] 图3是用于在图1的在线社交网络系统142中优先化搜索词的方法300的流程图。方法300可以由处理逻辑来执行,该处理逻辑可以包括硬件(例如,专用逻辑、可编程逻辑、微代码等)、软件(诸如在通用计算机系统或专用机器上运行)或两者的组合。在一个示例实施例中,处理逻辑驻留在图1的服务器系统140处,并且具体地,驻留在图2中所示的系统200处。

[0054] 如图3中所示,方法300在操作310处开始,此时图2的搜索词访问模块210访问包括表示地理位置的位置标识的搜索词。图2的位置强度评估器220在操作320处确定与位置标识相关联的在线社交网络系统142的会员数量。在操作330处,图2的重要性值生成器230确定搜索词的重要性值。重要性值反映了职位相关搜索请求多频繁地包括搜索词,并且还反映了与位置标识相关联的在线社交网络系统的会员数量。

[0055] 在操作340处,图2的优先级分数生成器240利用重要性值为搜索词生成优先级分

数。如上所述,除了其重要性值之外,还可以使用其相关性分数来生成搜索词的优先级分数。可以使用上述方法来生成搜索词的相关性分数。在操作350处,图2的网页生成器在在线社交网络系统142中生成网页,并且基于由优先级分数生成器240为搜索词生成的优先级分数而在网页中选择性地包括表示地理位置的词。图2的呈现模块250然后引起在显示设备上呈现网页。

[0056] 本文描述的示例方法的各种操作可以至少部分地由暂时配置(例如,通过软件)或永久配置成执行相关操作的一个或多个处理器来执行。无论是暂时配置还是永久配置,这类处理器都可以构成用于执行一个或多个操作或功能的处理器实现的模块。在一些示例实施例中,本文涉及的模块可以包括处理器实现的模块。

[0057] 类似地,本文描述的方法可以至少部分地由处理器实现。例如,方法的至少一些操作可以由一个或多个处理器或处理器实现的模块执行。某些操作的执行可以分布在一个或多个处理器中,不仅驻留在单个机器内,而且跨多个机器部署。在一些示例实施例中,一个或多个处理器可以位于单个位置(例如,在家庭环境、办公室环境内或作为服务器场),而在其它实施例中,处理器可以分布在多个位置。

[0058] 图5是呈计算机系统500的示例形式的机器的图形表示,其内可以执行用于使机器执行本文讨论的任何一个或多个方法的一组指令。在替代实施例中,机器作为独立设备运行或者可以连接(例如,联网)到其它机器。在联网部署中,机器可以在服务器-客户端网络环境中以服务器或客户端机器的资格进行操作,或者在对等(或分布式)网络环境中作为对等机器进行操作。该机器可以是个人计算机(PC)、平板电脑、机顶盒(STB)、个人数字助理(PDA)、蜂窝电话、网络器械、网络路由器、交换机或桥接器,或能够执行指定要由该机器采取的动作的一组指令(按顺序或以另外方式)的任何机器。此外,虽然仅示出单个机器,但术语“机器”也应被理解为包括单独或联合执行一组(或多组)指令以执行本文讨论的任何一种或多种方法的任何机器集合。

[0059] 示例计算机系统500包括经由处理器502(例如,中央处理单元(CPU)、图形处理单元(GPU)或两者)、主存储器504和静态存储器506,其经由总线505彼此通信。计算机系统500还可以包括视频显示单元510(例如,液晶显示器(LCD)或阴极射线管(CRT))。计算机系统500还包括字母数字输入设备512(例如,键盘)、用户界面(UI)导航设备514(例如,光标控制设备)、磁盘驱动单元516、信号生成设备518(例如,扬声器)和网络接口设备520。

[0060] 磁盘驱动单元516包括机器可读介质522,在该机器可读介质上存储具体实施本文描述的方法或功能中的任何一个或多个或由其利用的一组或多组指令和数据结构(例如,软件524)。在由计算机系统500执行期间,软件524还可以完全或至少部分地驻留在主存储器504内和/或处理器502内,其中主存储器504和处理器502也构成机器可读介质。

[0061] 软件524可以进一步经由网络接口设备520利用多种公知的传输协议(例如,超文本传输协议(HTTP))中的任何一种在网络526上传输或接收。

[0062] 虽然机器可读介质522在示例实施例中被示出为单个介质,但术语“机器可读介质”应被理解为包括单个介质或多个介质(例如,集中式或分布式数据库,和/或关联的高速缓存和服务器),其存储一组或多组指令。术语“机器可读介质”还应被理解为包括任何介质,其能够存储和编码由机器执行的一组指令并且使得机器执行本发明的实施例的任何一个或多个方法,或者能够存储和编码由这样的一组指令利用或与其相关联的数据结构。因

此,术语“机器可读介质”应被理解为包括但不限于固态存储器、光学介质和磁性介质。这类介质还可以包括但不限于硬盘、软盘、闪存卡、数字视频盘、随机存取存储器(RAM)、只读存储器(ROM)等。

[0063] 本文描述的实施例可以在包括安装在计算机上的软件、硬件或者软件和硬件的组合的操作环境中实现。如果实际上公开了多于一个,则本发明主题的这类实施例在本文可以单独地或共同地通过术语“发明”来引用,这仅仅是为了方便,并且不旨在自发地将本申请的范围限制为任何单个发明或发明构思。

[0064] 模块、组件和逻辑

本文将某些实施例描述为包括逻辑或多个组件、模块或机构。模块可以构成软件模块(例如(1)在非暂时性机器可读介质上或(2)在传输信号中具体实施的代码)或硬件实现的模块。硬件实现的模块是能够执行某些操作并且可以以某种方式配置或布置的有形单元。在示例实施例中,一个或多个计算机系统(例如,独立的客户端或服务器计算机系统)或一个或多个处理器可以由软件(例如,应用程序或应用程序部分)配置成硬件实现的模块,该硬件实现的模块用于执行如本文所述的某些操作。

[0065] 在各种实施例中,硬件实现的模块可以机械地或电子地实现。例如,硬件实现的模块可以包括被永久配置成执行某些操作的专用电路或逻辑(例如,作为专用处理器,诸如现场可编程门阵列(FPGA)或专用集成电路(ASIC))。硬件实现的模块还可以包括由软件临时配置以执行某些操作的可编程逻辑或电路(例如,包括在通用处理器或其它可编程处理器内)。将理解,在专用和永久配置的电路中或者在临时配置的电路中(例如,由软件配置的)机械地实现硬件实现的模块的决定可以由成本和时间考虑因素来驱使。

[0066] 因此,术语“硬件实现的模块”应该被理解为包括有形实体,即物理构造、永久配置(例如,硬连线)或临时或暂时配置(例如,编程)成以某种方式操作和/或执行本文所述的某些操作的实体。考虑到其中硬件实现的模块被临时配置(例如,编程)的实施例,每个硬件实现的模块不需要在任何一个时刻配置或例示。例如,在硬件实现的模块包括使用软件配置的通用处理器的情况下,通用处理器可以在不同的时间被配置成各自不同的硬件实现的模块。软件可以相应地将处理器配置成例如在一个时刻构成特定的硬件实现的模块并且在不同的时刻构成不同的硬件实现的模块。

[0067] 硬件实现的模块可以向其它硬件实现的模块提供信息并从其接收信息。因此,所描述的硬件实现的模块可以被视为通信地耦合。在同时存在多个这类硬件实现的模块的情况下,可以通过连接硬件实现的模块的信号传输(例如,通过适当的电路和总线)来实现通信。在其中多个硬件实现的模块在不同时间被配置或例示的实施例中,可以例如通过存储和检索多个硬件实现的模块可访问的存储器结构中的信息来实现这类硬件实现的模块之间的通信访问。例如,一个硬件实现的模块可以执行操作,并且将该操作的输出存储在其通信地耦合到的存储器设备中。然后另一个硬件实现的模块可以在稍后时间访问存储器设备以检索和处理存储的输出。硬件实现的模块也可以启动与输入或输出设备的通信,并且可以对资源(例如,信息的集合)进行操作。

[0068] 本文描述的示例方法的各种操作可以至少部分地由暂时配置(例如,通过软件)或永久配置成执行相关操作的一个或多个处理器来执行。无论是暂时配置还是永久配置,这类处理器都可以构成用于执行一个或多个操作或功能的处理器实现的模块。在一些示例实

施例中,本文涉及的模块可以包括处理器实现的模块。

[0069] 类似地,本文描述的方法可以至少部分地由处理器实现。例如,方法的至少一些操作可以由一个或多个处理器或处理器实现的模块执行。某些操作的执行可以分布在一个或多个处理器中,不仅驻留在单个机器内,而且跨多个机器部署。在一些示例实施例中,一个或多个处理器可以位于单个位置(例如,在家庭环境、办公室环境内或作为服务器场),而在其它实施例中,处理器可以分布在多个位置。

[0070] 一个或多个处理器还可以操作以支持“云计算”环境中的相关操作的执行或作为“软件即服务”(SaaS)。例如,至少一些操作可以由一组计算机(作为包括处理器的机器的示例)执行,这些操作可以经由网络(例如,互联网)并且经由一个或多个适当的接口(例如,应用程序编程接口(API))进行访问。

[0071] 因此,已经描述了在在线社交网络系统中优先化表示位置的搜索词的方法和系统。虽然已经参考具体示例实施例描述了实施例,但将显而易见的是,在不脱离本发明主题的更广范围的情况下,可以对这些实施例进行各种修改和改变。因此,说明书和附图应被认为是说明性的而不是限制性的。

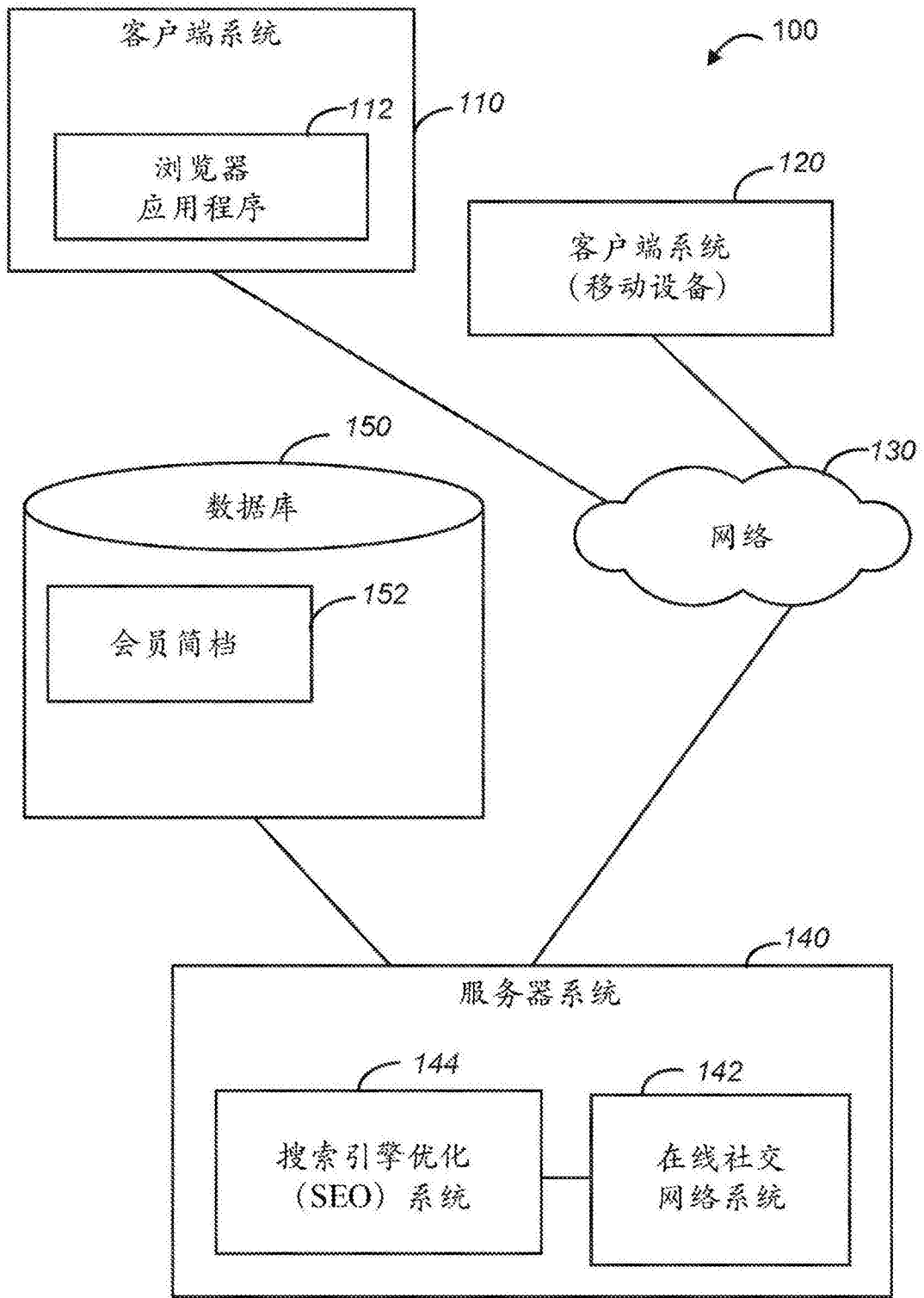


图 1

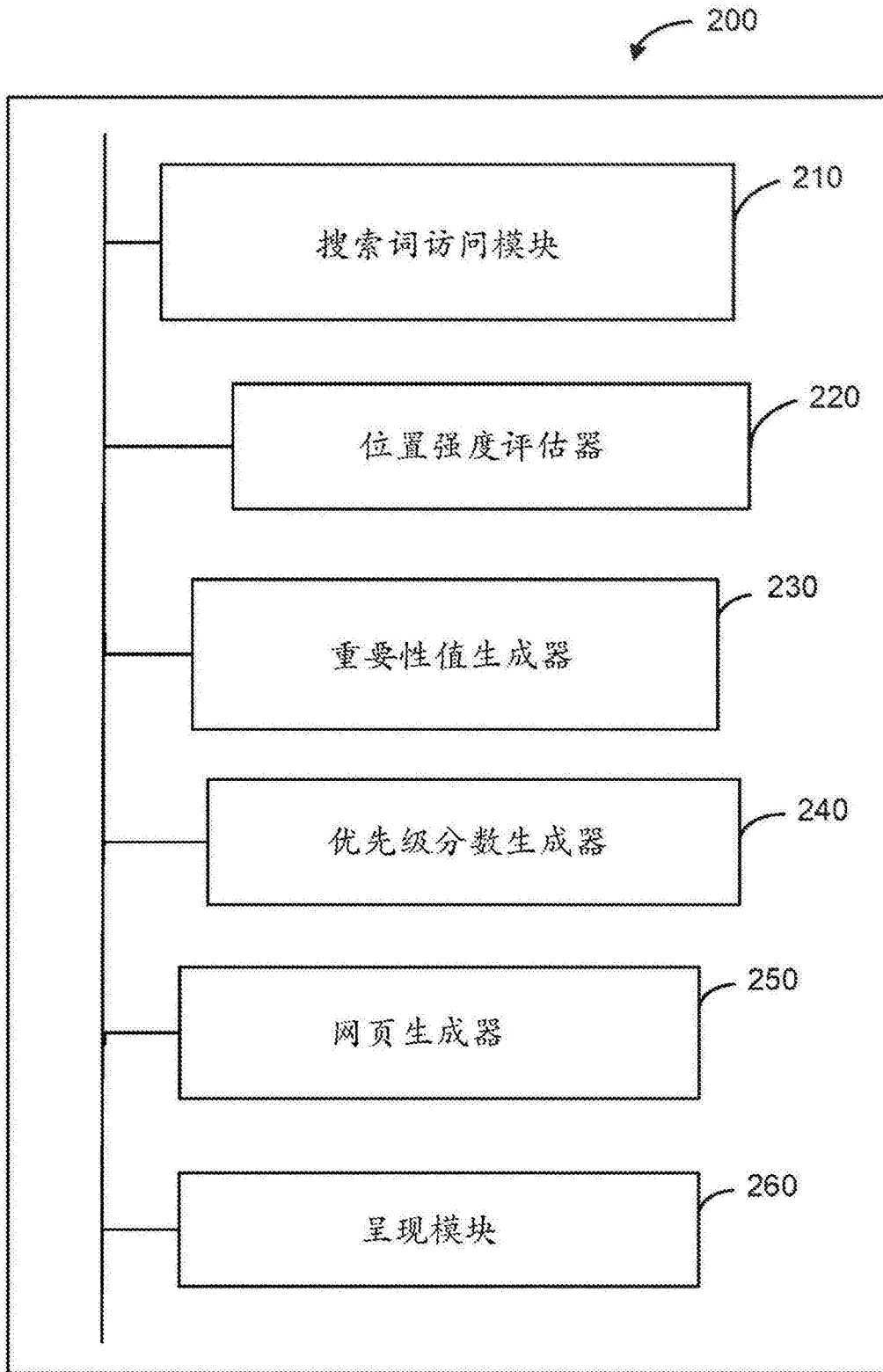


图 2

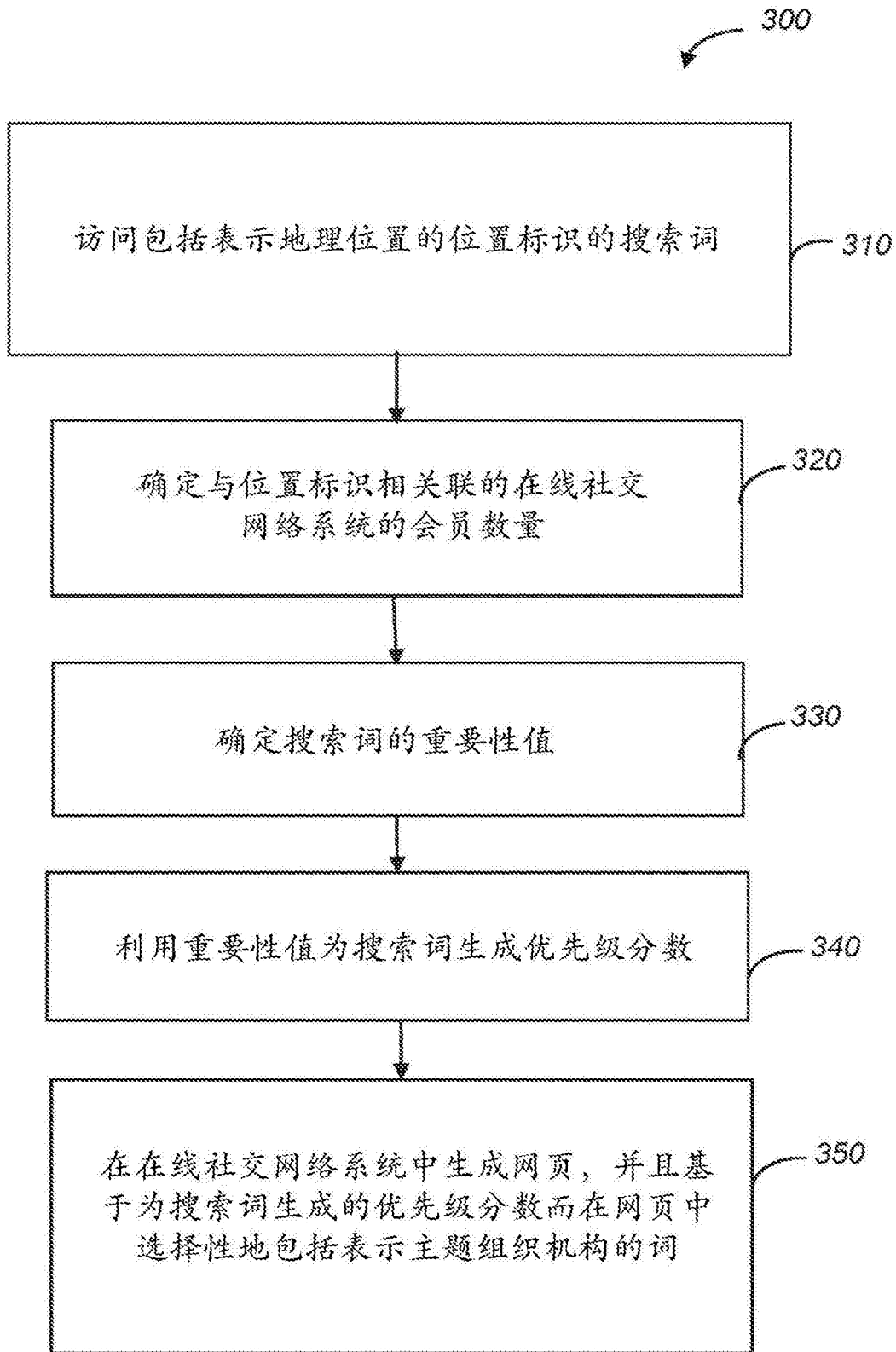


图 3

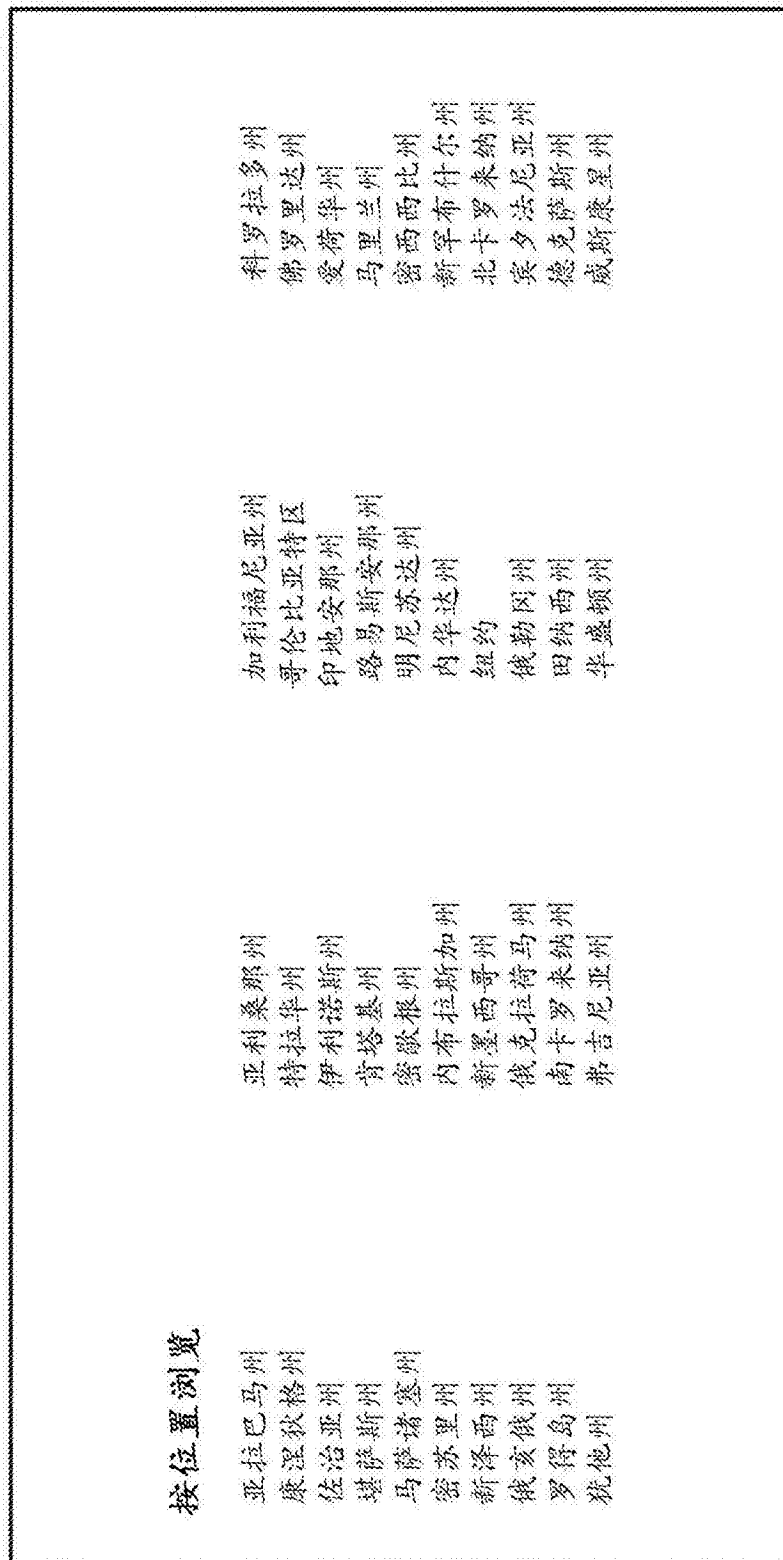


图 4

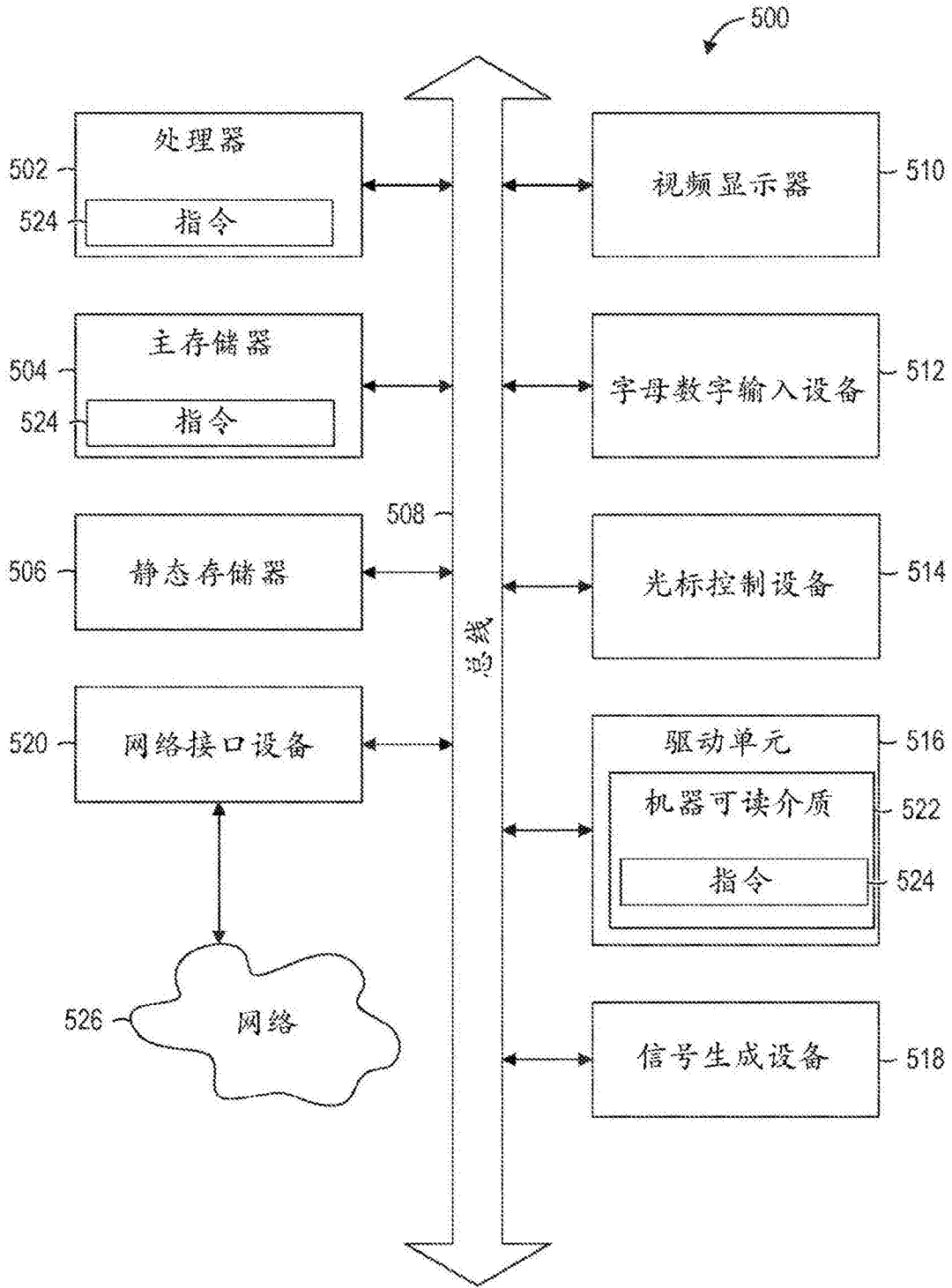


图 5