

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2018/223993 A1

(43) 国际公布日

2018 年 12 月 13 日 (13.12.2018)

WIPO | PCT

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2018/090150

(22) 国际申请日: 2018 年 6 月 6 日 (06.06.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710431235.X 2017年6月8日 (08.06.2017) CN

(71) 申请人: 广州优视网络科技有限公司 (GUANGZHOU UC NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省广州市天河区黄埔大道中路146-150号首层自编07房, Guangdong 510627 (CN)。

(72) 发明人: 潘岸腾 (PAN, Anteng); 中国广东省广州市天河区黄埔大道中路146-150号首层自编07房, Guangdong 510627 (CN)。

(74) 代理人: 北京博雅睿泉专利代理事务所 (特殊普通合伙) (BEYOND TALENT PATENT AGENT FIRM); 中国北京市朝阳区朝阳门外大街10号昆泰大厦1201单元, Beijing 100020 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: APPLICATION SEARCH METHOD, DEVICE AND SERVER

(54) 发明名称: 一种应用软件搜索方法、装置及服务器

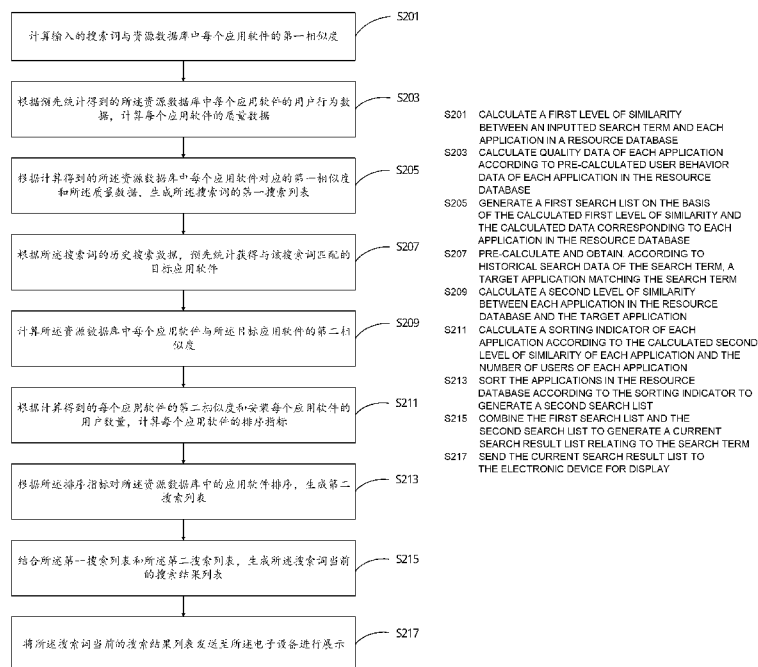


图 7

(57) Abstract: An application search method, device and server related to the technical field of the Internet. One implementation of the application search method comprises: calculating a first level of similarity between an inputted search term and each application in a resource database (S201); calculating quality data of each application according to pre-calculated user behavior data of each application in the resource database (S203); and generating a first search list relating to the search term on the basis of the calculated first level of similarity and the calculated quality data corresponding to each application in the resource database (S205). The application search

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

method comprehensively considers the level of similarity between a search term and an application as well as quality data of the application, and provides a result that significantly enhances user experience of downloading. A user can obtain from a search result a desired or preferred application that at the same time ensures user experience.

(57) 摘要: 一种应用软件搜索方法、装置及服务器, 属于互联网技术领域。所述应用软件搜索方法的一种实现方式包括: 计算输入的搜索词与资源数据库中每个应用软件的第一相似度 (S201), 然后根据预先统计得到的资源数据库中每个应用软件的用户行为数据, 计算每个应用软件的质量数据 (S203), 最后再根据计算得到的资源数据库中每个应用软件对应的第一相似度和质量数据, 生成搜索词的第一搜索列表 (S205)。这种综合考虑了搜索词与应用软件的相似度以及应用软件的质量数据的应用软件搜索方法, 其搜索出的结果可以明显提升用户的下载使用体验, 用户可以从搜索结果中得到自己想要使用或者喜欢使用的并且有使用体验保障的应用软件。

一种应用软件搜索方法、装置及服务器

技术领域

本发明涉及互联网技术领域，具体而言，涉及一种应用软件搜索方法、装置及服务器。

背景技术

现有技术中，用户可以在应用软件商店（如 AppStore、AndroidMarket 等）中免费下载或付费下载应用软件，以便于娱乐、学习或工作。众所周知，应用软件商店中同一类型的应用软件通常具有多款（如相机类型的应用软件包括相机 360、美颜相机、Poco 相机等等），若事先没有既定的下载对象，用户通过应用软件商店获取想要下载的某一类型的应用软件时，通常是在应用软件商店的界面上先输入搜索词进行搜索，再在搜索结果中自行选择。这种应用软件搜索方式无法为用户提供智能优化的下载服务，展示出的搜索结果给用户的体验不佳，经常会出现没有用户想要下载的应用软件或者搜索出的应用软件使用体验欠佳，不能很好的满足用户的下载使用需求。

发明内容

有鉴于此，本发明的目的在于提供一种应用软件搜索方法、装置及服务器，以改善上述问题。

本发明实施例提供一种应用软件搜索方法，该方法包括：计算输入的搜索词与资源数据库中每个应用软件的第一相似度；根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据；根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库中的至少一个应用软件。

本发明另一实施例提供一种应用软件搜索装置，该装置包括：第一相似度计算模块，用于计算输入的搜索词与资源数据库中每个应用软件的第一相似度；质量数据计算模块，用于根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据；第一搜索列表生成模块，用于根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库中的至少一个应用软件。

本发明另一实施例提供一种服务器，所述服务器包括：存储器；处理器；及应用软件搜索装置，该应用软件搜索装置安装于所述存储器中并包括一个或多个由所述处理器执行的软件功能模块，该应用软件搜索装置包括：第一相似度计算模块，用于计算输入的搜索词与资源数据库中每个应用软件的第一相似度；质量数据计算模块，用于根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据；第一搜索列表生成模块，用于根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库中的至少一个应用软件。

本发明另一实施例提供一种服务器，包括：存储器，用于存储可执行指

令；处理器，用于根据所述可执行指令的控制，运行服务器执行本发明其他实施例提供的应用软件搜索方法。

本发明另一实施例提供一种计算机可读取存储介质，其中，所述计算机可读取存储介质中存储有计算机软件产品，所述计算机软件产品包括若干指令，用以使得一台计算机设备执行本发明其他实施例提供的应用软件搜索方法。

本发明实施例提供的应用软件搜索方法、装置及服务器，计算输入的搜索词与资源数据库中每个应用软件的第一相似度，然后根据预先统计得到的资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据，最后再根据计算得到的资源数据库中每个应用软件对应的第一相似度和质量数据，生成搜索词的第一搜索列表。这种综合考虑了搜索词与应用软件的相似度以及应用软件的质量数据的应用软件搜索方法，其搜索出的结果可以明显提升用户的下载使用体验，用户可以从搜索结果中得到自己想要使用或者喜欢使用的并且有质量保障的应用软件。

附图说明

为了更清楚地说明本发明实施例的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，应当理解，以下附图仅示出了本发明的某些实施例，因此不应被看作是对范围的限定，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他相关的附图。

图 1 为本发明实施例提供的一种服务器与至少一个电子设备的通信交互示意图；

图 2 为本发明实施例提供的一种服务器的方框示意图；

图 3 为本发明实施例提供的一种应用软件搜索方法的流程图；

图 4 为本发明实施例提供的一示例中 App Store 的搜索界面示意图；

图 5 为本发明实施例提供的根据图 4 中搜索界面输入的搜索词“相机”进行搜索后得到的搜索结果展示界面示意图；

5 图 6 为本发明实施例提供的点击图 5 中所示的“相机 360”后进入的浏览界面示意图；

图 7 为本发明实施例提供的另一种应用软件搜索方法的流程图；

图 8 为本发明实施例提供的一种应用软件搜索装置的功能模块框图。

图标：100-服务器；200-电子设备；110-应用软件搜索装置；120-存储器；
10 130-处理器；140-通信单元；150-资源数据库；1102-第一相似度计算模块；1104-质量数据计算模块；1106-第一搜索列表生成模块；1108-目标应用软件获取模块；1110-第二相似度计算模块；1112-排序指标计算模块；1114-第二搜索列表生成模块；1116-搜索结果列表生成模块。

具体实施方式

15 为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本发明实施例的组件可以以各种不同的配置来布置和设计。应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项
20 在一个附图中被定义，则在随后的附图中不需要对其进行进一步定义和解释。

因此,以下对在附图中提供的本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围,而是仅仅表示本发明的选定实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

5 请参阅图 1,是本发明实施例提供的服务器 100 与至少一个电子设备 200 进行通信的交互示意图。所述电子设备 200 可通过网络与服务器 100 建立通信连接,以进行数据传输或交互。所述服务器 100 可以提供至少一个适用于各种操作系统的应用(Application, APP)的安装包供所述电子设备 200 下载。所述
10 电子设备 200 可通过所述网络访问所述服务器 100 以下载适用于该电子设备 200 的操作系统的应用的安装包,并将所述应用安装到所述电子设备 200 中。

本实施例中,所述应用可以是应用软件商店,如 App Store、AndroidMarket、MobileMarket 等。当安装于所述电子设备 200 上的应用软件商店在前台打开时,用户可以在打开的界面上浏览各款应用软件,或者在界面上通过输入搜索词搜索自身想要浏览或下载使用的应用软件。

15 本实施例中,所述服务器 100 可以是,但不限于,网络服务器、文件传输协议服务器等。所述电子设备 200 可以是,但不限于,智能手机、平板电脑、个人数字助理、移动上网设备等。所述网络可以是,但不限于,有线网络或无线网络。所述电子设备 200 的操作系统可以是,但不限于,安卓(Android)系统、IOS 系统、Windows phone 系统等。

20 请参阅图 2,是本发明实施例提供的所述服务器 100 的方框示意图。所述服务器 100 包括应用软件搜索装置 110、存储器 120、处理器 130、通信单元 140 以及资源数据库 150。

所述存储器 120、处理器 130 以及通信单元 140 之间直接或间接的电性连接，以实现数据的传输或交互。所述应用软件搜索装置 110 包括至少一个可以软件或固件的形式存储于所述存储器 120 中或固化在所述服务器 100 的操作系统中的软件功能模块。所述处理器 130 用于执行所述存储器 120 中存储的可执行模块，例如所述应用软件搜索装置 110 所包括的软件功能模块及计算机程序等。

所述通信单元 140 用于通过所述网络建立所述服务器 100 与电子设备 200 之间的通信连接，并用于通过所述网络收发数据。

所述资源数据库 150 可以用于存储例如应用软件商店中各种应用软件的信息。可选地，资源数据库 150 可以存储于所述存储器 120 中，以供所述处理器 130 在执行例如所述应用软件搜索装置 110 所包括的软件功能模块时访问。

应用于本发明实施例中，虽然服务器 100 示出了多个装置，但是服务器 100 可以仅涉及其中几个装置。例如，服务器 100 可以仅涉及存储器 120 和处理器 130，存储器 120 用于存储可执行指令，处理器 130 用于根据存储器 120 中存储的可执行指令的控制，运行服务器 100 执行本发明实施提供的应用软件搜索方法。

请参阅图 3，是本发明实施例提供的一种应用软件搜索方法的流程图。本实施例中的方法适用于所述服务器 100。所应说明的是，本方法不以图 3 及下述的具体顺序为限制。下面将对图 3 所示的各步骤进行详细阐述。

步骤 S101，计算输入的搜索词与资源数据库中每个应用软件的第一相似度。

作为一种实施方式，当应用软件商店(如 App Store、华为应用商城、小米应用商店等)在电子设备 200 的前台打开时，用户可以在应用界面上输入搜索

词，然后通过电子设备 200 将所述搜索词发送到服务器 100，以使服务器 100 根据接收到的搜索词在资源数据库 150 中进行应用软件搜索。

下面以所述资源数据库 150 中的一个目标应用软件为例，示例性地阐述所述服务器 100 计算所述搜索词与目标应用软件的第一相似度的实现过程。可以理解，计算该第一相似度的方式可以不限于此。

预先获取描述该目标应用软件的文本信息。所述文本信息可以包括目标应用软件的名称或者该目标应用软件的介绍文字。所述文本信息还可以是该目标应用软件的名称和介绍文字的组合。作为一种可能的实现方式，目标应用软件的名称或者介绍文字可以是目标应用软件上传者或管理者手动设置并在上传服务器 100 时一并发送到服务器 100 的，也可以通过网络从特定网页中去抓取。若上述名称或文字来源于网页，可以将抓取内容中附带的例如 html 元素的网页信息剔除后得到相应的文本信息。

然后，使用预设的分词算法，例如最大匹配法、逆向匹配法、双向匹配法等，对该目标应用软件的文本信息及搜索词进行分词处理。在本实施例中，在进行分词处理时可以对分词结果进行过滤以保留能反映文本内容的分词。例如，可以将分词结果中的标点符号（如：句号、顿号及逗号等）和不表达实质内容的分词（如：“无需”、“没有”、“的”及“又”等连接词）过滤掉。

计算出每个分词在所述目标应用软件的文本信息、所述搜索词中出现的词频(term frequency)。词频(term frequency)用于表示分词在其对应文本内容中出现的频率。本实施例中，可以采用以下计算式计算每个分词在目标应用软件的文本信息中的词频：

$$tf_i = \frac{n_i}{\sum_k n_k}$$

其中， n_i 表示分词 t_i 在目标应用软件的文本信息中出现的次数， $\sum_k n_k$ 表示在目标应用软件的文本信息中所有分词出现的次数之和， tf_i 表示分词 t_i 在目标应用软件的文本信息中的词频。

标应用软件的文本信息中出现的频率。

类似地，本实施例中，可以采用以下计算式计算每个分词在搜索词中的词频：

$$tf'_i = \frac{n'_i}{\sum_k n'_k}$$

5 其中， n'_i 表示分词 t_i 在搜索词中出现的次数， $\sum_k n'_k$ 表示在所述搜索词中所有分词出现的次数之和， tf'_i 表示分词 t_i 在所述搜索词中出现的频率。

再根据所述资源数据库 150 中的所有目标应用软件对应的文本信息，计算每个分词的逆文档频率(inverse document frequency)。以 J 表示资源数据库 150 中所有目标应用软件的集合， $isContain_{i,j}$ 表示分词 t_i 是否在目标应用软件 j 10 对应的文本信息中出现，若等于 1 则表示出现了，若等于 0 则表示没有出现， idf_i 表示分词 t_i 的逆文档频率，具体计算公式可以为：

$$idf_i = \frac{1}{\log_2 \left(1 + \sum_{j \in J} isContain_{i,j} \right)}$$

然后，根据每个分词的逆文档频率和在目标应用软件的文本信息中的词频，计算每个分词在该目标应用软件的文本信息中的 $tf-idf$ (term frequency- 15 inverse document frequency)值。具体地，计算分词 t_i 在该目标应用软件的文本信息中的 $tf-idf$ 值的方式可以是：

$$tf-idf_i = tf_i \times idf_i$$

其中， idf_i 表示分词 t_i 的逆文档频率， tf_i 表示分词 t_i 在该目标应用软件的文本信息中出现的频率， $tf-idf_i$ 表示分词 t_i 在该目标应用软件的文本信息中的 $tf-idf$ 20 值。

以及，根据每个分词的逆文档频率和在搜索词中的词频，计算每个分词在搜索词中的 $tf-idf$ 值。具体地，计算分词 t_i 在搜索词中的 $tf-idf$ 值的方式可以是：

$$tf-idf'_i = tf'_i \times idf_i$$

其中, idf_i 表示分词 t_i 的逆文档频率, tf'_i 表示分词 t_i 在搜索词中出现的频率, $tf-idf'_i$ 表示分词 t_i 在搜索词中的 $tf-idf$ 值。

最后,按照上述方法计算出每个分词在目标应用程序的文本信息中的 $tf-idf$ 值后,得到对应该目标应用程序的 $tf-idf$ 向量:

$$5 \quad tdf = (tf-idf_1, tf-idf_2, \dots, tf-idf_i, \dots)$$

其中, tdf 表示该目标应用程序的 $tf-idf$ 向量。

以及,按照上述方法计算出每个分词在搜索词中的 $tf-idf$ 值后,得到该搜索词的 $tf-idf$ 向量:

$$tdf' = (tf-idf'_1, tf-idf'_2, \dots, tf-idf'_i, \dots)$$

10 其中, tdf' 表示该搜索词的 $tf-idf$ 向量。

再根据向量 tdf 和 tdf' 计算出搜索词与目标应用程序的第一相似度,具体计算式可以是:

$$sim = \cos \langle tdf, tdf' \rangle$$

15 即,通过所述搜索词与目标应用程序的 $tf-idf$ 向量的余弦值表示第一相似度。

按照上述示例中所述的方法可以计算出输入的搜索词与资源数据库 150 中每个应用程序的第一相似度。

步骤 S103,根据预先统计得到的所述资源数据库 150 中每个应用程序的用户行为数据,计算每个应用程序的质量数据。

20 本实施例中,应用程序的用户行为数据可以与在预设时段内平均浏览该应用程序的用户数和平均下载该应用程序的用户数相关。其中,预设时段可以是当前时刻之前的一定时段,例如,可以是当前时刻之前的 N 天或小时,或者是该应用程序自在应用程序商店上线发布开始至当前的时段。另外,本实施例中,

所述平均浏览该应用软件的用户数可以是指平均每天浏览该应用软件的用户数，也可以是指平均每小时浏览该应用软件的用户数。类似地，平均下载该应用软件的用户数可以是指平均每天下载该应用软件的用户数，也可以是指平均每小时下载该应用软件的用户数。

5 这里需要说明的是，本实施例中，所述“浏览”应用软件主要包括用户在例如应用软件商店的展示界面上看到该应用软件，并点击进入到介绍应用软件的浏览界面进行浏览的操作。以 App Store 为例，在一个示例中，如图 4 所示，用户在 App Store 搜索界面的搜索框内输入“相机”后，展示界面展示出如图 5 所示的搜索结果（仅是示意性的）。以其中的“相机 360”为例，用户可以点击
10 展示界面上的“相机 360”选项后，进入到如图 6 所示的介绍“相机 360”基本信息的浏览界面，在该浏览界面上用户可以浏览“相机 360”的介绍信息，此为一次“浏览”操作。

作为一种实施方式，计算每个应用软件的质量数据的方式可以是：

根据计算式： $Qul = \frac{num_{br} + num_{dl}}{Num_{sum}}$ 计算每个应用软件的质量数据。其中，

15 Num_{sum} 表示在预设时段内平均曝光该应用软件的用户数； num_{br} 表示在预设时段内平均浏览该应用软件的用户数； num_{dl} 表示在预设时段内平均下载该应用软件的用户数。

这里需要说明的是，上述“曝光”应用软件主要包括在例如应用软件商店的展示界面上该应用软件被展示给用户观看的操作。也就是说，在预设时段
20 内平均曝光该应用软件的用户数可以是指在预设时段内平均观看到该应用软件的用户数，例如在预设时段内平均每天观看到该应用软件的用户数，或平均每小时观看到该应用软件的用户数，等等。

以上仅仅是所述应用软件质量数据的一种计算方式，在其它实施例方式中，还可以加入用于反映用户对所述应用软件的其它用户行为数据来计算所述质量数据。

步骤 S105 根据计算得到的资源数据库 150 中每个应用软件对应的第一相似度和质量数据，生成搜索词的第一搜索列表。

本实施例中，所述第一搜索列表包括资源数据库 150 中的至少一个应用软件。

作为一种实施方式，生成所述第一搜索列表的过程可以是：先根据计算出的第一相似度对资源数据库 150 中的应用软件进行排序，例如可以按照第一相似度的降序进行排序，得到候选搜索列表；再将该候选搜索列表中的应用软件按序分为多组，每组包括第一预设数目个应用软件；然后，分别计算每组中的第一预设数目个应用软件的质量数据，并根据计算出的质量数据对该第一预设数目个应用软件进行排序，例如按照质量数据进行降序排列；最后，在每组的第一预设数目个应用软件中选择第二预设数目个应用软件加入预先初始化的一空列表中，得到所述第一搜索列表。优选地，第一搜索列表中包括的应用软件可以是在按照降序排列时质量数据排列靠前的第二预设数目个应用软件。

本实施方式中，所述第一预设数目可以为一预设的固定数值或为一预设的数值区间内的任意值。例如，该第一预设数目可以是预设的固定数值 25，或者可以是预设的数值区间[10,30]内的任意值。也就是说，上述将候选搜索列表中的应用软件按序分为多组时，每组包括的应用软件数目可以相同(例如都是 25 个)，也可以不同(例如各组的应用软件数目在 10~30 之间变化)。例如，可以将所述第一预设数目设定为一固定值，分组时最后一组的应用软件

数目可以不足于第一预设数目。所述第二预设数目与所述第一预设数目之比可以为预定比例，如 0.6~0.8 之间任意比例。

下面给出一具体示例，以更好的阐述本实施例中生成所述第一搜索列表的过程，但可以是理解的是，该示例仅是示意性的。本示例中设定所述第一
5 预设数目为固定数值 15，设定所述第二预设数目为 12，即，两者比值为 0.8。

步骤 S1，对资源数据库 150 中的所有应用软件，应用步骤 S101 中所述的方式计算出输入的搜索词与每个应用软件的第一相似度。提取对应相似度大于预设值（如大于 0）的所有应用软件，然后按照第一相似度进行降序排列，得到候选搜索列表。

10 步骤 S2，从所述候选搜索列表中取出排在前面的第一预设数目个应用软件，计算该第一预设数目个应用软件的质量数据，然后按照计算出的质量数据的重新降序排列，最后截取质量数据排在前面的第二预设数目个应用软件添加进预先初始化的一目标列表中。该目标列表初始时空列表。

步骤 S3，判断所述候选搜索列表是否为空，如果不为空，则返回步骤
15 S2，若此时候选搜索列表中剩余应用软件数目小于所述第一预设数目，可以默认继续执行，并适应性的按照所述预定比例（本示例中为 0.8）改变所述第二预设数目的值；如果所述候选搜索列表为空，则将当前的所述目标列表作为所述第一搜索列表。

本示例中，作为一种实施方式，每次将应用软件添加进所述目标列表时，
20 可以按照在所述目标列表中从后至前的顺序依次添加。

本实施例中得到的所述第一搜索列表可以直接作为搜索词本次被搜索的搜索结果列表返回到所述电子设备 200，以通过所述电子设备 200 展示给用户。用户可以在该搜索结果列表中选择自己想要浏览或者下载的应用软

件，由于本次搜索结果综合考虑了应用软件的质量数据和与搜索词的第一相似度，所以可以较好的提升用户的下载使用体验。

请参阅图 7，是本发明实施例提供的另一应用软件搜索方法的流程图。

如图 7 所示，本实施例的方法中包括如下步骤：步骤 S201-步骤 S217。

5 步骤 S201，计算输入的搜索词与资源数据库 150 中每个应用软件的第一相似度。

步骤 S203，根据预先统计得到的所述资源数据库 150 中每个应用软件的用户行为数据，计算每个应用软件的质量数据。

10 步骤 S205，根据计算得到的所述资源数据库 150 中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表。

本实施例中的步骤 S201~步骤 S205 的具体实现方式可以参照图 3 所示实施例中步骤 S101~步骤 S105 的详细阐述，在此不再赘述。

步骤 S207，根据所述搜索词的历史搜索数据，预先统计获得与该搜索词匹配的目标应用软件。

15 本实施例中，可以根据预设时段内该搜索词的历史搜索反馈数据，统计是否有超过一半的用户搜索该搜索词后选择下载同一个应用软件，如果有，那么该应用软件即为与该搜索词匹配的目标应用软件。可以理解，所述预设时段可以是任意长度的时间段，例如可以是应用软件商店自上线发布至当前时刻的时段，或者自该搜索词第一次被搜索至当前时刻的时段等等。

20 步骤 S209，计算所述资源数据库 150 中每个应用软件与所述目标应用软件的第二相似度。

作为一种实施方式，所述第二相似度可以是指协同过滤 Jaccard 相似性。相应地，计算资源数据库 150 中的每个应用软件与目标应用软件的所述第二相似度的方式可以是：

根据计算式：

$$Sim_{a,i} = \begin{cases} \frac{|UserInstalled_i \cap UserInstalled_a|}{|UserInstalled_i \cup UserInstalled_a|}, & T_i \cap T_a \neq \Phi \\ 0, & T_i \cap T_a = \Phi \end{cases}$$

计算所述资源数据库 150 中每个应用软件与该目标应用软件的相似度。其中， i 表示资源数据库 150 中的任一应用软件， a 表示所述目标应用软件； $Sim_{a,i}$ 表示应用软件 i 与目标应用软件 a 的第二相似度； $UserInstalled_i$ 表示资源数据库 150 的所有用户（例如，或者说，应用软件商店的所有用户）中安装了应用软件 i 的用户子集； $UserInstalled_a$ 表示资源数据库 150 的所有用户中安装了目标应用软件 a 的用户子集； T_i 表示应用软件 i 的标签集合， T_a 表示应用软件 a 的标签集合。

上述的应用软件的标签集合是指例如由运营手工打上的标识应用软件类型的一个或多个标签的集合。例如，应用软件“微信”对应的标签集合可以是{社交、通信}。

步骤 S211，根据计算得到的每个应用软件的所述第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标。

作为一种实施方式，可以根据计算式：

$$Or_i = \alpha \frac{Sim_{a,i}}{\text{avg}_{j \in I} (Sim_{a,j})} + \beta \frac{|NumUserInstalled_i|}{\text{avg}_{j \in I} (|NumUserInstalled_j|)}$$

计算每个应用软件的排序指标。其中， Or_i 表示资源数据库 150 中任一应用软件 i 的排序指标； I 表示资源数据库 150 中所有应用软件的集合； $\text{avg}_{j \in I} (Sim_{a,j})$

表示资源数据库 150 中所有应用软件对应的第二相似度的平均值；
 $|NumUserInstalled_i|$ 表示应用软件 i 的安装用户的用户数量；
 $avg_{j \in I}(|NumUserInstalled_j|)$ 表示资源数据库 150 中所有应用软件的安装用户的
用户数量的平均值； α 、 β 是预设的权重参数，两者均大于 0，且和为 1。

5 步骤 S213，根据所述排序指标对所述资源数据库 150 中的应用软件排序，生成第二搜索列表。

作为一种实施方式，对资源数据库 150 中的所有应用软件，根据排序指标的降序进行排序，生成第二搜索列表。

10 步骤 S215，结合所述第一搜索列表和所述第二搜索列表，生成所述搜索词当前的搜索结果列表。

本实施例中，通过所述第一搜索列表和第二搜索列表生成该搜索词本次被搜索的搜索结果列表可以通过以下步骤 S01-S03 实现。

步骤 S01，初始化一目标列表，该目标列表初始为空列表。

步骤 S02，向该目标列表中添加应用软件。

15 其中，向该目标列表中添加应用软件的步骤包括从所述第一搜索列表中按顺序(例如从前至后)获取第三预设数目个应用软件加入所述目标列表中，以及从所述第二搜索列表中按顺序(例如从前至后)获取第四预设数目个应用软件加入所述目标列表中。

20 所述第三预设数目和所述第四预设数目根据用于显示所述搜索结果列表的设备每屏能够显示的应用软件总数目预先设定。例如，所述第三预设数目可以设定为每屏能够显示的应用软件总数目的 30%，所述第四预设数目可以设定为每屏能够显示的应用软件总数目的 70%。假设所述电子设备 200 每屏可显示 6 个应用软件，则所述第三预设数目可以设定为 2，所述第四预设

数目可以设定为 4。

步骤 S03，判断所述第一搜索列表和所述第二搜索列表是否为空。

若第一搜索列表、第二搜索列表都不为空，则返回上述向该目标列表中
添加应用程序的步骤，即，步骤 S02。

5 若第一搜索列表、第二搜索列表中出现一个为空，则将另一个非空的搜索列表中剩余的所有应用程序添加进上述目标列表中。

若第一搜索列表、第二搜索列表都为空，则将当前的所述目标列表作为该搜索词当前的搜索结果列表并输出。

下面给出一具体示例，以便更好的阐述本实施例中通过所述第一搜索列表
10 和第二搜索列表生成该搜索词本次被搜索的搜索结果列表的过程。可以理解的是，本示例仅是示意性的。

步骤 S001，初始化一目标列表，该目标列表初始时空。

步骤 S002，判断所述第一搜索列表中应用程序数目是否大于等于所述第三预设数目，若大于等于，则取该第一搜索列表中排在前面的第三预设数目
15 个应用程序添加进所述目标列表中，若小于，则将该第一搜索列表中所有应用程序一起添加进所述目标列表中；再判断所述第二搜索列表中应用程序数目是否大于等于所述第四预设数目，若大于等于，则取该第二搜索列表中排在前面的第四预设数目个应用程序添加进所述目标列表中，若小于，则将该第二搜索列表中所有应用程序一起添加进所述目标列表中。

20 步骤 S003，判断所述第一搜索列表和所述第二搜索列表是否都为空。若全部为空，则将所述目标列表输出作为所述搜索词本次的搜索结果列表；若其中一个为空，则将另一个非空的搜索列表中的剩余所有应用程序一起添加进所述目标列表中后，再将该目标列表作为所述搜索词本次的搜索结果列表

输出；若两个都不为空，则返回上述步骤 S002。

本示例中，从所述第一搜索列表和第二搜索列表中取出应用软件向所述目标列表中添加时，可以是按照在所述目标列表中从后至前的顺序依次添加。

5 步骤 S217，将所述搜索词当前的搜索结果列表发送至所述电子设备 200 进行展示。

本实施例中提供的应用软件搜索方法，既考虑了资源数据库 150 中的每个应用软件的质量数据以及与搜索词的第一相似度，又考虑了资源数据库 150 中每个应用软件与历史搜索中与该搜索词匹配的目标应用软件的第二相似度，通过这种综合、全面的搜索方式得到的搜索结果列表，可以提升用户
10 的搜索体验和后续的浏览或下载使用体验。

请参阅图 8，是本发明实施例提供的一种应用软件搜索装置 110 的功能模块框图。应用软件搜索装置 110 包括第一相似度计算模块 1102、质量数据计算模块 1104、第一搜索列表生成模块 1106、目标应用软件获取模块 1108、
15 第二相似度计算模块 1110、排序指标计算模块 1112、第二搜索列表生成模块 1114、搜索结果列表生成模块 1116。

所述第一相似度计算模块 1102，用于计算输入的搜索词与资源数据库 150 中每个应用软件的第一相似度。该第一相似度计算模块 1102 可用于执行图 3 中所示的步骤 S101 和图 7 中所示的步骤 S201，具体的操作方法可以参
20 照上述对步骤 S101 的详细阐述。

所述质量数据计算模块 1104，用于根据预先统计得到的所述资源数据库 150 中每个应用软件的用户行为数据，计算每个应用软件的质量数据。该质量数据计算模块 1104 可用于执行图 3 中所示的步骤 S103 和图 7 中所示的步

骤 S203，具体的操作方法可以参照上述对步骤 S103 的详细阐述。

所述第一搜索列表生成模块 1106，用于根据计算得到的所述资源数据库 150 中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库 150 中的至少一个应用软件。该第一搜索列表生成模块 1106 可用于执行图 3 中所示的步骤 S105 和图 7 中所示的步骤 S205，具体的操作方法可以参照上述对步骤 S105 的详细阐述。

所述目标应用软件获取模块 1108，用于根据所述搜索词的历史搜索数据，预先统计获得与该搜索词匹配的目标应用软件。该目标应用软件获取模块 1108 可用于执行图 7 中所示的步骤 S207，具体的操作方法可以参照上述对步骤 S207 的详细阐述。

所述第二相似度计算模块 1110，用于计算所述资源数据库 150 中每个应用软件与所述目标应用软件的第二相似度。该第二相似度计算模块 1110 可用于执行图 7 中所示的步骤 S209，具体的操作方法可以参照上述对步骤 S209 的详细阐述。

所述排序指标计算模块 1112，用于根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标。该排序指标计算模块 1112 可用于执行图 7 中所示的步骤 S211，具体的操作方法可以参照上述对步骤 S211 的详细阐述。

所述第二搜索列表生成模块 1114，用于根据所述排序指标对所述资源数据库 150 中的应用软件排序，生成第二搜索列表。该第二搜索列表生成模块 1114 可用于执行图 7 中所示的步骤 S213，具体的操作方法可以参照上述对步骤 S213 的详细阐述。

所述搜索结果列表生成模块 1116 ,用于结合所述第一搜索列表和所述第二搜索列表,生成所述搜索词当前的搜索结果列表。该搜索结果列表生成模块 1116 可用于执行图 7 中所示的步骤 S215,具体的操作方法可以参照上述对步骤 S215 的详细阐述。

5 以上各模块可以是由软件代码实现,此时,上述的各模块可存储于服务器的存储器内。以上各模块同样可以由硬件例如集成电路芯片实现。

需要说明的是,本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

10 本发明实施例所提供的应用软件搜索装置 110,其实现原理及产生的技术效果和前述方法实施例相同,为简要描述,装置实施例部分未提及之处,可参考前述方法实施例中相应内容。

在本申请所提供的几个实施例中,应该理解到,所揭露的装置和方法,也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本发明的多个实施例的装置、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分,所述模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现方式中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图

15

20

和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用

执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

另外,在本发明各个实施例中的各功能模块可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或两个以上模块集成形成一个独立的部分。

所述功能如果以软件功能模块的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器、随机存取存储器、磁碟或者光盘等各种可以存储程序代码的介质。需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到

变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

权利要求书

1.一种应用软件搜索方法，其特征在于，该方法包括：

5 计算输入的搜索词与资源数据库中每个应用软件的第一相似度；

根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，
计算每个应用软件的质量数据；

根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和
所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所
10 述资源数据库中的至少一个应用软件。

2.根据权利要求 1 所述的应用软件搜索方法，其特征在于，每个应用软件
的用户行为数据包括在预设时段内平均浏览该应用软件的用户数和平均
下载该应用软件的用户数，所述根据预先统计得到的所述资源数据库中每个
应用软件的用户行为数据，计算每个应用软件的质量数据的步骤，包括：

15 根据计算式： $Q_{ul} = \frac{num_{br} + num_{dl}}{Num_{sum}}$ 计算每个应用软件的质量数据，其中，
 Num_{sum} 表示在所述预设时段内平均曝光该应用软件的用户数， num_{br} 表示在所
述预设时段内平均浏览该应用软件的用户数， num_{dl} 表示在所述预设时段内平
均下载该应用软件的用户数。

3.根据权利要求 1 或 2 所述的应用软件搜索方法，其特征在于，所述根
20 据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质
量数据，生成该搜索词的第一搜索列表的步骤，包括：

初始化一目标列表；

根据所述第一相似度对所述资源数据库中的应用软件进行排序，得到候选搜索列表；

5 将所述候选搜索列表中的应用软件按序分为多组，每组包括第一预设数目个应用软件；

根据每组中每个应用软件的质量数据，对每组中的第一预设数目个应用软件进行排序；

在每组中根据上述排序的结果选择第二预设数目个应用软件加入所述目标列表中，得到所述第一搜索列表。

10 4.根据权利要求 1-3 任意一项所述的应用软件搜索方法，其特征在于，所述第一预设数目为一预设数值或为一预设的数值区间内的任意值，所述第二预设数目与所述第一预设数目之比为预定比例。

15 5.根据权利要求 1-4 任意一项所述的应用软件搜索方法，其特征在于，在根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成该搜索词的第一搜索列表的步骤之后，该方法还包括：

根据所述搜索词的历史搜索数据，预先统计获得与该搜索词匹配的目标应用软件；

计算所述资源数据库中每个应用软件与所述目标应用软件的第二相似度；

20 根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标；

根据所述排序指标对所述资源数据库中的应用软件排序，生成第二搜索列表；

结合所述第一搜索列表和所述第二搜索列表，生成所述搜索词当前的搜索结果列表。

- 6.根据权利要求 1-5 任意一项所述的应用软件搜索方法，其特征在于，所述计算所述资源数据库中每个应用软件与该目标应用软件的第二相似度的步骤，包括：

根据计算式：

$$Sim_{a,i} = \begin{cases} \frac{|UserInstalled_i \cap UserInstalled_a|}{|UserInstalled_i \cup UserInstalled_a|}, & T_i \cap T_a \neq \Phi \\ 0, & T_i \cap T_a = \Phi \end{cases}$$

- 计算所述资源数据库中每个应用软件与所述目标应用软件的相似度，其中， i 表示所述资源数据库中的任一应用软件， a 表示所述目标应用软件， $Sim_{a,i}$ 表示应用软件 i 与目标应用软件 a 的第二相似度， $UserInstalled_i$ 表示所述资源数据库的所有用户中安装了应用软件 i 的用户子集， $UserInstalled_a$ 表示所述资源数据库的所有用户中安装了目标应用软件 a 的用户子集， T_i 表示应用软件 i 的标签集合， T_a 表示目标应用软件 a 的标签集合。

- 7.根据权利要求 1-6 任意一项所述的应用软件搜索方法，其特征在于，所述根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标的步骤，包括：

根据计算式：

$$Or_i = \alpha \frac{Sim_{a,i}}{\text{avg}_{j \in I} (Sim_{a,j})} + \beta \frac{|NumUserInstalled_i|}{\text{avg}_{j \in I} (|NumUserInstalled_j|)}$$

- 计算每个应用软件的排序指标，其中， Or_i 表示资源数据库中任一应用软件 i 的排序指标， I 表示所述资源数据库中所有应用软件的集合， $\text{avg}_{j \in I} (Sim_{a,j})$ 表示所述资源数据库中所有应用软件对应的所述第二相似度的平均值，

$|NumUserInstalled_i|$ 表示应用软件 i 的安装用户的用户数量， $avg_{j \in I}(|NumUserInstalled_j|)$ 表示所述资源数据库中所有应用软件的安装用户的用户数量的平均值， α 、 β 是预设的权重参数。

8. 根据权利要求 1-7 任意一项所述的应用软件搜索方法，其特征在于，
5 所述结合所述第一搜索列表和所述第二搜索列表，生成该搜索词当前的搜索结果列表的步骤，包括：

初始化一目标列表；

向该目标列表中添加应用软件；其中，向该目标列表中添加应用软件的
步骤包括从所述第一搜索列表中按顺序获取第三预设数目个应用软件加入
10 所述目标列表中，以及从所述第二搜索列表中按顺序获取第四预设数目个应用
软件加入所述目标列表中，所述第三预设数目和所述第四预设数目根据用
于显示所述搜索结果列表的设备每屏能够显示的应用软件数目预先设定；

判断所述第一搜索列表和所述第二搜索列表是否为空；

若第一搜索列表、第二搜索列表都不为空，则返回上述向该目标列表中
15 添加应用软件的步骤；

若第一搜索列表、第二搜索列表中出现一个为空，则将另一个非空的搜索
列表中剩余的所有应用软件添加进上述目标列表中；

若第一搜索列表、第二搜索列表都为空，则将当前的所述目标列表作为
该搜索词当前的搜索结果列表并输出。

20 9. 根据权利要求 1-8 任意一项所述的应用软件搜索方法，其特征在于，
所述计算输入的搜索词与资源数据库中每个应用软件的第一相似度的步骤，
包括：

获取描述任一目标应用软件的文本信息；

对所述目标文本信息及所述搜索词进行分词，并计算出每个分词在所述目标文本信息、所述搜索词中出现的词频；

根据所述资源数据库中的所有应用软件对应的文本信息，计算所述每个分词的逆文档频率；

5 根据所述每个分词在所述目标文本信息中出现的词频以及所述每个分词的逆文档频率，计算得到所述目标应用软件的第一特征向量；

根据所述每个分词在所述搜索词中出现的词频以及所述每个分词的逆文档频率，计算得到所述搜索词的第二特征向量；

10 计算所述第一特征向量和所述第二特征向量的夹角的余弦值，该余弦值表示所述搜索词与所述目标应用软件的第一相似度。

10.一种应用软件搜索装置，其特征在于，该装置包括：

第一相似度计算模块，用于计算输入的搜索词与资源数据库中每个应用软件的第一相似度；

15 质量数据计算模块，用于根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据；

第一搜索列表生成模块，用于根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库中的至少一个应用软件。

20 11.根据权利要求 10 所述的应用软件搜索装置，其特征在于，每个应用软件的用户行为数据包括在预设时段内平均浏览该应用软件的用户数和平均下载该应用软件的用户数，所述质量数据计算模块根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据的方式，包括：

根据计算式： $Q_{ul} = \frac{num_{br} + num_{dl}}{Num_{sum}}$ 计算每个应用软件的质量数据，其中， Num_{sum} 表示在所述预设时段内平均曝光该应用软件的用户数， num_{br} 表示在所述预设时段内平均浏览该应用软件的用户数， num_{dl} 表示在所述预设时段内平均下载该应用软件的用户数。

5 12.根据权利要求 10 或 11 所述的应用软件搜索装置，其特征在于，所述第一搜索列表生成模块根据计算得到的所述资源数据库中每个应用软件对应的第一相似度和所述质量数据，生成所述搜索词的第一搜索列表的方式，包括：

初始化一目标列表；

10 根据所述第一相似度和所述质量数据对所述资源数据库中的应用软件进行排序，得到候选搜索列表；

将所述候选搜索列表中的应用软件按序分为多组，每组包括第一预设数目个应用软件；

15 根据每组中每个应用软件的质量数据，对每组中的第一预设数目个应用软件进行排序；

在每组中根据上述排序的结果选择第二预设数目个应用软件加入所述目标列表中，得到所述第一搜索列表。

13.根据权利要求 10-12 任意一项所述的应用软件搜索装置，其特征在于，所述第一预设数目为一预设数值或为一预设的数值区间内的任意值，所述第二预设数目与所述第一预设数目之比为预定比例。

14.根据权利要求 10-13 任意一项所述的应用软件搜索装置，其特征在于，该应用软件搜索装置还包括：

目标应用软件获取模块，用于根据所述搜索词的历史搜索数据，预先统计获得与该搜索词匹配的目标应用软件；

第二相似度计算模块，用于计算所述资源数据库中每个应用软件与所述目标应用软件的第二相似度；

5 排序指标计算模块，用于根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标；

第二搜索列表生成模块，用于根据所述排序指标对所述资源数据库中的应用软件排序，生成第二搜索列表；

10 搜索结果列表生成模块，用于结合所述第一搜索列表和所述第二搜索列表，生成所述搜索词当前的搜索结果列表。

15.根据权利要求 10-14 任意一项所述的应用软件搜索装置，其特征在于，所述第二相似度计算模块计算所述资源数据库中每个应用软件与所述目标应用软件的第二相似度的方式，包括：

根据计算式：

$$15 \quad Sim_{a,i} = \begin{cases} \frac{|UserInstalled_i \cap UserInstalled_a|}{|UserInstalled_i \cup UserInstalled_a|}, & T_i \cap T_a \neq \Phi \\ 0, & T_i \cap T_a = \Phi \end{cases}$$

10 计算所述资源数据库中每个应用软件与该目标应用软件的相似度，其中， i 表示所述资源数据库中的任一应用软件， a 表示所述目标应用软件， $Sim_{a,i}$ 表示应用软件 i 与目标应用软件 a 的第二相似度， $UserInstalled_i$ 表示所述资源数据库的所有用户中安装了应用软件 i 的用户子集， $UserInstalled_a$ 表示所述资源数据库的所有用户中安装了目标应用软件 a 的用户子集， T_i 表示应用软件 i 的标签集合， T_a 表示目标应用软件 a 的标签集合。

16.根据权利要求 10-15 任意一项所述的应用软件搜索装置，其特征在于，

所述排序指标计算模块根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标的方式，包括：

根据计算式：

$$Or_i = \alpha \frac{Sim_{a,i}}{avg_{j \in I}(Sim_{a,j})} + \beta \frac{|NumUserInstalled_i|}{avg_{j \in I}(|NumUserInstalled_j|)}$$

- 5 计算每个应用软件的排序指标，其中， Or_i 表示资源数据库中任一应用软件 i 的排序指标， I 表示所述资源数据库中所有应用软件的集合， $avg_{j \in I}(Sim_{a,j})$ 表示所述资源数据库中所有应用软件对应的所述第二相似度的平均值， $|NumUserInstalled_i|$ 表示应用软件 i 的安装用户的用户数量， $avg_{j \in I}(|NumUserInstalled_j|)$ 表示所述资源数据库中所有应用软件的安装用户
- 10 的用户数量的平均值， α 、 β 是预设的权重参数。

17.根据权利要求 11-16 任意一项所述的应用软件搜索装置 其特征在于，所述搜索结果列表生成模块结合所述第一搜索列表和所述第二搜索列表，生成该搜索词当前的搜索结果列表的方式，包括：

初始化一目标列表；

- 15 向该目标列表中添加应用软件；其中，向该目标列表中添加应用软件的步骤包括从所述第一搜索列表中按顺序获取第三预设数目个应用软件加入所述目标列表中，以及从所述第二搜索列表中按顺序获取第四预设数目个应用软件加入所述目标列表中，所述第三预设数目和所述第四预设数目根据用于显示所述搜索结果列表的设备每屏能够显示的应用软件数目预先设定；

- 20 判断所述第一搜索列表和所述第二搜索列表是否为空；

若第一搜索列表、第二搜索列表都不为空，则返回上述向该目标列表中添加应用软件的步骤；

若第一搜索列表、第二搜索列表中出现一个为空，则将另一个非空的搜索列表中剩余的所有应用软件添加进上述目标列表中；

若第一搜索列表、第二搜索列表都为空，则将当前的所述目标列表作为该搜索词当前的搜索结果列表并输出。

- 5 18.根据权利要求 10-17 任意一项所述的应用软件搜索装置，其特征在于，所述第一相似度计算模块计算输入的搜索词与资源数据库中每个应用软件的第一相似度的方式，包括：

获取描述任一目标应用软件的目标文本信息；

- 10 对所述目标文本信息及所述搜索词进行分词，并计算出每个分词在所述目标文本信息、所述搜索词中出现的词频；

根据所述资源数据库中的所有应用软件对应的文本信息，计算每个分词的逆文档频率；

根据每个分词在所述目标文本信息中出现的词频以及每个分词的逆文档频率，计算得到所述目标应用软件的第一特征向量；

- 15 根据每个分词在所述搜索词中出现的词频以及每个分词的逆文档频率，计算得到所述搜索词的第二特征向量；

计算所述第一特征向量和所述第二特征向量的夹角的余弦值，该余弦值表示所述搜索词与所述目标应用软件的第一相似度。

19.一种服务器，其特征在于，所述服务器包括：

- 20 存储器；

处理器；及

应用软件搜索装置，该应用软件搜索装置安装于所述存储器中并包括一个或多个由所述处理器执行的软件功能模块，该应用软件搜索装置包括：

第一相似度计算模块，用于计算输入搜索词与资源数据库中每个应用软件的所述第一相似度；

质量数据计算模块，用于根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据；

- 5 第一搜索列表生成模块，用于根据计算得到的所述资源数据库中每个应用软件对应的所述第一相似度和所述质量数据，生成所述搜索词的第一搜索列表，所述第一搜索列表包括所述资源数据库中的至少一个应用软件。

20. 根据权利要求 19 所述的服务器，其特征在于，每个应用软件的用户行为数据包括在预设时段内平均浏览该应用软件的用户数和平均下载该应用
10 软件的用户数，所述质量数据计算模块根据预先统计得到的所述资源数据库中每个应用软件的用户行为数据，计算每个应用软件的质量数据的方式，包括：

根据计算式： $Qul = \frac{num_{br} + num_{dl}}{Num_{sum}}$ 计算每个应用软件的质量数据，其中，

- 15 Num_{sum} 表示在所述预设时段内平均曝光该应用软件的用户数， num_{br} 表示在所述预设时段内平均浏览该应用软件的用户数， num_{dl} 表示在所述预设时段内平均下载该应用软件的用户数。

21. 根据权利要求 19 或 20 所述的服务器，其特征在于，所述第一搜索列表生成模块根据计算得到的所述资源数据库中每个应用软件对应的所述第一相似度和所述质量数据，生成所述搜索词的第一搜索列表的方式，包括：

- 20 初始化一目标列表；

根据所述第一相似度对所述资源数据库中的应用软件进行排序，得到候选搜索列表；

将所述候选搜索列表中的应用软件按序分为多组，每组包括第一预设数目个应用软件；

根据每组中每个应用软件的质量数据，对每组中的第一预设数目个应用软件进行排序；

5 在每组中根据上述排序的结果选择第二预设数目个应用软件加入所述目标列表中，得到所述第一搜索列表。

22.根据权利要求 19-21 任意一项所述的服务器，其特征在于，该应用软件搜索装置还包括：

10 目标应用软件获取模块，用于根据所述搜索词的历史搜索数据，预先统计获得与该搜索词匹配的目标应用软件；

第二相似度计算模块，用于计算所述资源数据库中每个应用软件与所述目标应用软件的第二相似度；

排序指标计算模块，用于根据计算得到的每个应用软件的第二相似度和安装每个应用软件的用户数量，计算每个应用软件的排序指标；

15 第二搜索列表生成模块，用于根据所述排序指标对所述资源数据库中的应用软件排序，生成第二搜索列表；

搜索结果列表生成模块，用于结合所述第一搜索列表和所述第二搜索列表，生成所述搜索词当前的搜索结果列表。

23.根据权利要求 19-22 任意一项所述的服务器，其特征在于，所述第二相似度计算模块计算所述资源数据库中每个应用软件与所述目标应用软件的第二相似度的方式，包括：

20

根据计算式：

$$Sim_{a,i} = \begin{cases} \frac{|UserInstalled_i \cap UserInstalled_a|}{|UserInstalled_i \cup UserInstalled_a|}, T_i \cap T_a \neq \Phi \\ 0, T_i \cap T_a = \Phi \end{cases}$$

计算所述资源数据库中每个应用软件与该目标应用软件的相似度，其中， i 表示所述资源数据库中的任一应用软件， a 表示所述目标应用软件， $Sim_{a,i}$ 表示应用软件 i 与目标应用软件 a 的第二相似度， $UserInstalled_i$ 表示所述资源数
 5 据库的所有用户中安装了应用软件 i 的用户子集， $UserInstalled_a$ 表示所述资源数据库的所有用户中安装了目标应用软件 a 的用户子集， T_i 表示应用软件 i 的标签集合， T_a 表示目标应用软件 a 的标签集合。

24. 根据权利要求 19-23 任意一项所述的服务器，其特征在于，所述排序指标计算模块根据计算得到的每个应用软件的第二相似度和安装每个应用
 10 软件的用户数量，计算每个应用软件的排序指标的方式，包括：

根据计算式：

$$Or_i = \alpha \frac{Sim_{a,i}}{avg_{j \in I} (Sim_{a,j})} + \beta \frac{|NumUserInstalled_i|}{avg_{j \in I} (|NumUserInstalled_j|)}$$

计算每个应用软件的排序指标，其中， Or_i 表示资源数据库中任一应用软件 i 的排序指标， I 表示所述资源数据库中所有应用软件的集合， $avg_{j \in I} (Sim_{a,j})$ 表
 15 示所述资源数据库中所有应用软件对应的所述第二相似度的平均值， $|NumUserInstalled_i|$ 表示应用软件 i 的安装用户的用户数量， $avg_{j \in I} (|NumUserInstalled_j|)$ 表示所述资源数据库中所有应用软件的安装用户的用户数量的平均值， α 、 β 是预设的权重参数。

25. 根据权利要求 19-24 任意一项所述的服务器，其特征在于，所述搜索结果列表生成模块结合所述第一搜索列表和所述第二搜索列表，生成该搜索
 20 词当前的搜索结果列表的方式，包括：

初始化一目标列表；

向该目标列表中添加应用软件；其中，向该目标列表中添加应用软件的步骤包括从所述第一搜索列表中按顺序获取第三预设数目个应用软件加入所述目标列表中，以及从所述第二搜索列表中按顺序获取第四预设数目个应用软件加入所述目标列表中，所述第三预设数目和所述第四预设数目根据用于显示所述搜索结果列表的设备每屏能够显示的应用软件数目预先设定；

判断所述第一搜索列表和所述第二搜索列表是否为空；

若第一搜索列表、第二搜索列表都不为空，则返回上述向该目标列表中添加应用软件的步骤；

10 若第一搜索列表、第二搜索列表中出现一个为空，则将另一个非空的搜索列表中剩余的所有应用软件添加进上述目标列表中；

若第一搜索列表、第二搜索列表都为空，则将当前的所述目标列表作为该搜索词当前的搜索结果列表并输出。

26.一种服务器，其中，包括：

15 存储器，用于存储可执行指令；

处理器，用于根据所述可执行指令的控制，运行所述服务器执行如权利要求 1-9 任意一项所述的应用软件搜索方法。

27.一种计算机可读取存储介质，其中，

所述计算机可读取存储介质中存储有计算机软件产品，所述计算机软件产品包括若干指令，用以使得一台计算机设备执行如权利要求 1-9 任意一项所述的应用软件搜索方法。

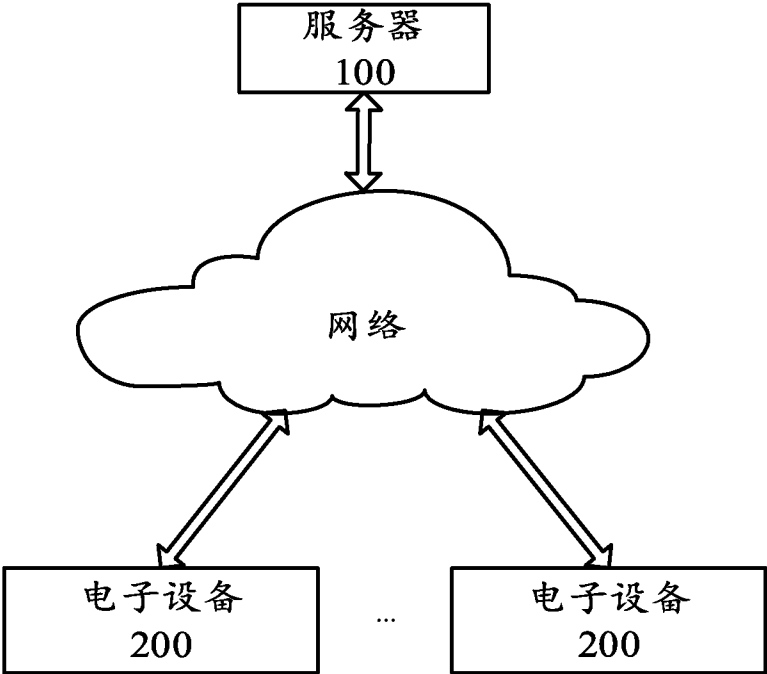


图 1

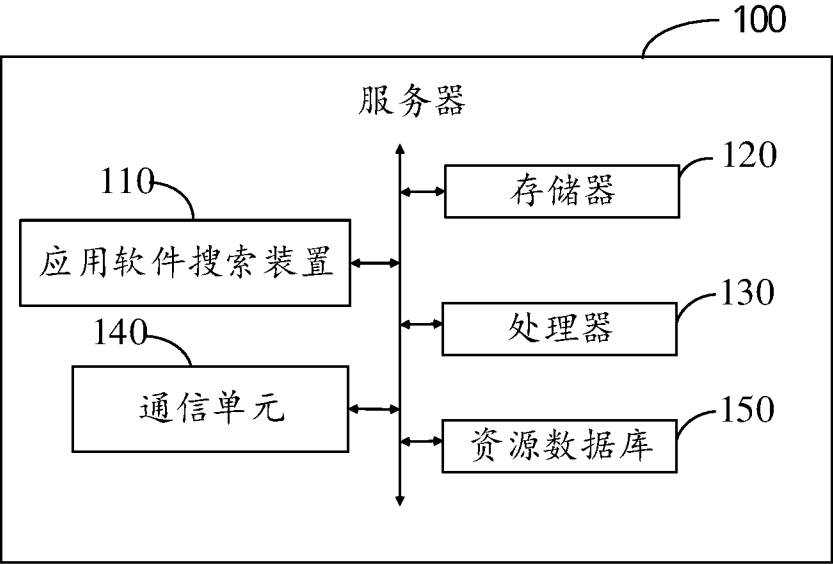


图 2

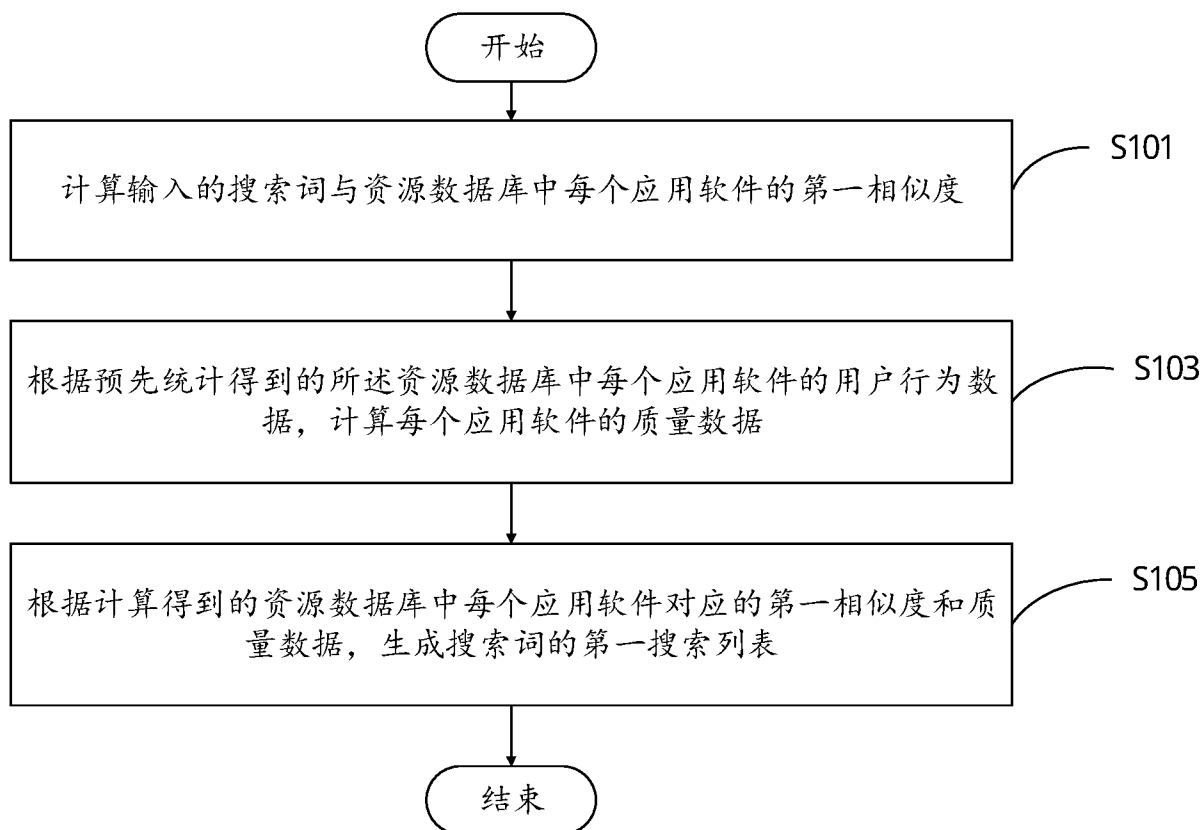


图 3

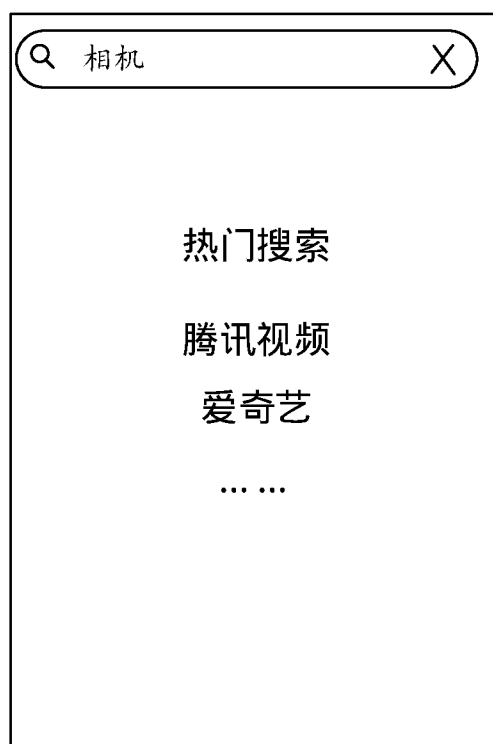


图 4

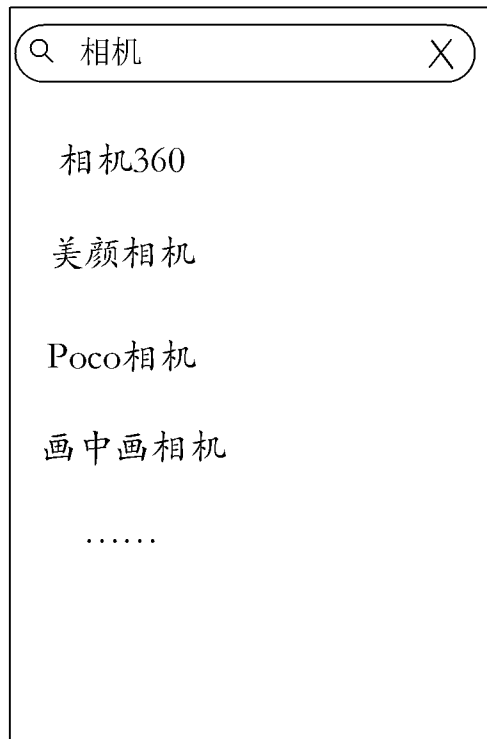


图 5

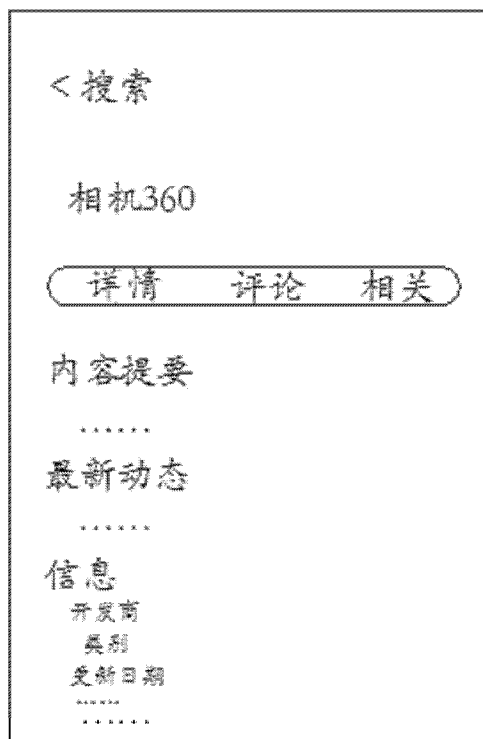


图 6

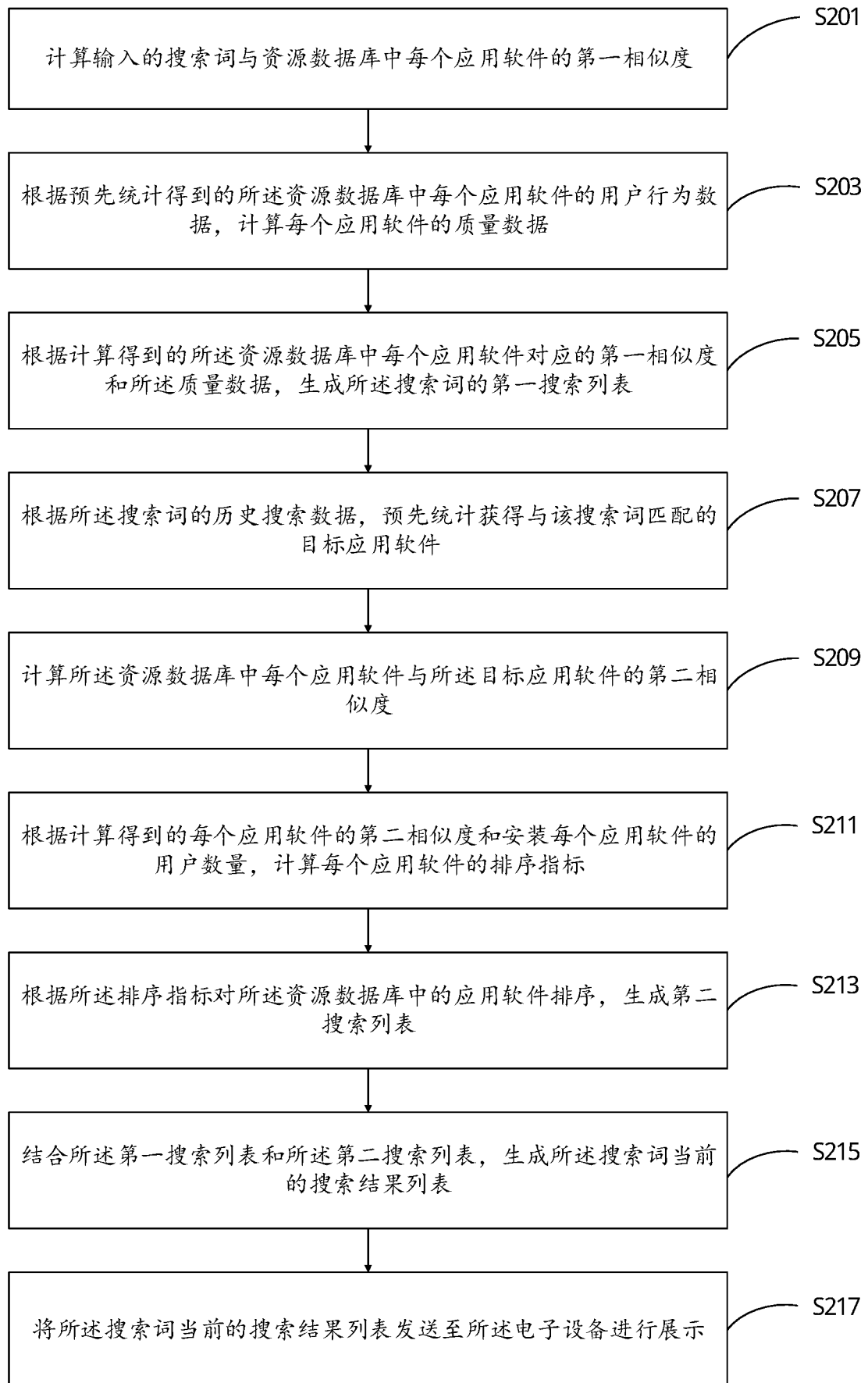


图 7

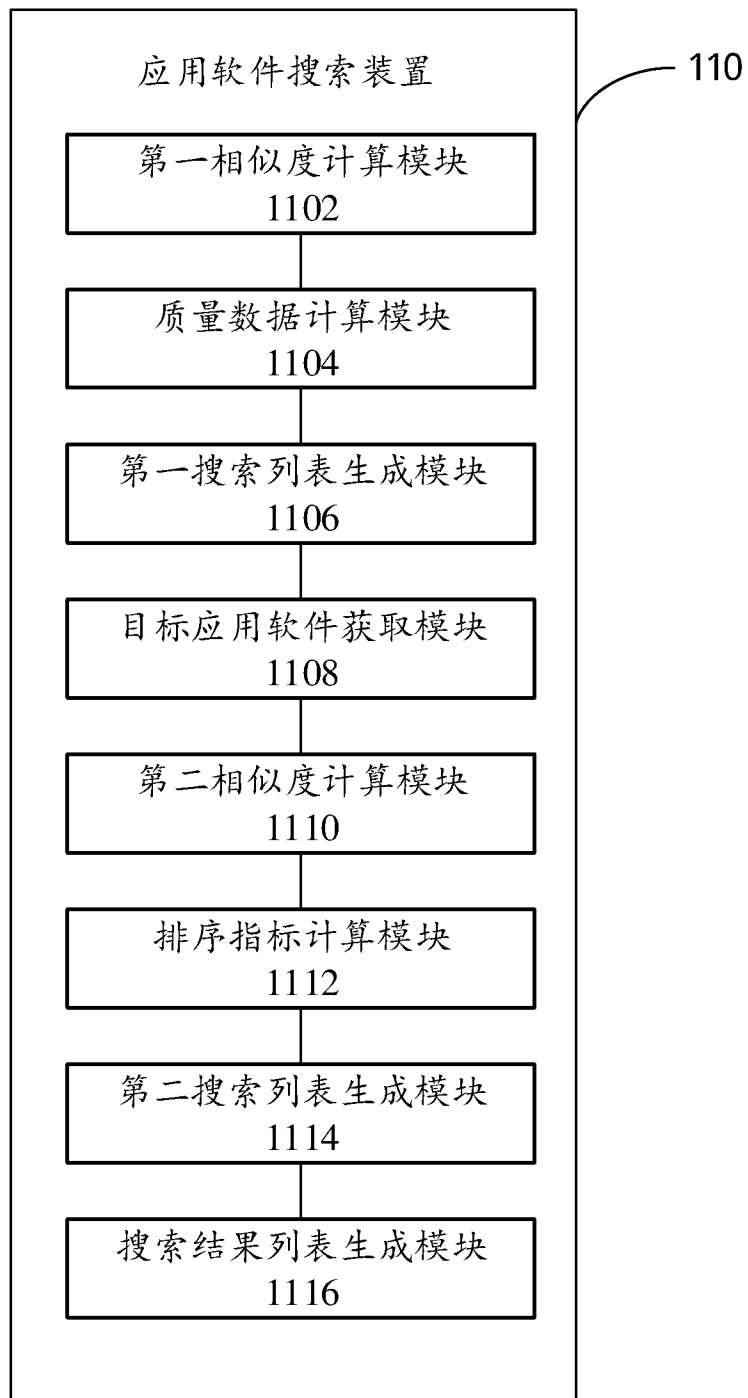


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/090150

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE: 软件, 应用, 检索, 搜索, 相似, 相关, 关联, 质量, 排序, 序列, 列表, 用户, 行为, 浏览, 点击, 下载, software, APP, application, search, correlation, related, quality, sort, sequence, list, user, behavior, browse, view, click, download

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107256251 A (GUANGZHOU UC NETWORK TECHNOLOGY CO., LTD.) 17 October 2017 (2017-10-17) claims 1-25, and description, paragraphs 122-124	1-27
X	CN 105893390 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 24 August 2016 (2016-08-24) description, paragraphs 51-126	1-4, 9-13, 18-21, 26-27
A	CN 106777282 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 31 May 2017 (2017-05-31) entire document	1-27
A	CN 105389328 A (NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY) 09 March 2016 (2016-03-09) entire document	1-27
A	CN 106503025 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 15 March 2017 (2017-03-15) entire document	1-27



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

24 August 2018

Date of mailing of the international search report

07 September 2018

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/090150

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009043732 A1 (NHN CORPORATION) 12 February 2009 (2009-02-12) entire document	1-27

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/090150

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107256251	A	17 October 2017	None			
CN	105893390	A	24 August 2016	None			
CN	106777282	A	31 May 2017	US	2018189291	A1	05 July 2018
CN	105389328	A	09 March 2016	None			
CN	106503025	A	15 March 2017	None			
US	2009043732	A1	12 February 2009	JP	2007524172	A	23 August 2007
				KR	100462292	B1	08 December 2004
				WO	2005083593	A1	09 September 2005

国际检索报告

国际申请号

PCT/CN2018/090150

A. 主题的分类 G06F 17/30 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																							
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) WPI, EPDOC, CNPAT, CNKI, IEEE: 软件, 应用, 检索, 搜索, 相似, 相关, 关联, 质量, 排序, 序列, 列表, 用户, 行为, 浏览, 点击, 下载, software, APP, application, search, correlation, related, quality, sort, sequence, list, user, behavior, browse, view, click, download																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 107256251 A (广州优视网络科技有限公司) 2017年 10月 17日 (2017 - 10 - 17) 权利要求1-25、说明书第122-124段</td> <td>1-27</td> </tr> <tr> <td>X</td> <td>CN 105893390 A (北京搜狗科技发展有限公司) 2016年 8月 24日 (2016 - 08 - 24) 说明书第51-126段</td> <td>1-4、9-13、 18-21、26-27</td> </tr> <tr> <td>A</td> <td>CN 106777282 A (百度在线网络技术北京有限公司) 2017年 5月 31日 (2017 - 05 - 31) 全文</td> <td>1-27</td> </tr> <tr> <td>A</td> <td>CN 105389328 A (中国人民解放军国防科学技术大学) 2016年 3月 9日 (2016 - 03 - 09) 全文</td> <td>1-27</td> </tr> <tr> <td>A</td> <td>CN 106503025 A (北京搜狗科技发展有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文</td> <td>1-27</td> </tr> <tr> <td>A</td> <td>US 2009043732 A1 (NHN CORPORATION) 2009年 2月 12日 (2009 - 02 - 12) 全文</td> <td>1-27</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 107256251 A (广州优视网络科技有限公司) 2017年 10月 17日 (2017 - 10 - 17) 权利要求1-25、说明书第122-124段	1-27	X	CN 105893390 A (北京搜狗科技发展有限公司) 2016年 8月 24日 (2016 - 08 - 24) 说明书第51-126段	1-4、9-13、 18-21、26-27	A	CN 106777282 A (百度在线网络技术北京有限公司) 2017年 5月 31日 (2017 - 05 - 31) 全文	1-27	A	CN 105389328 A (中国人民解放军国防科学技术大学) 2016年 3月 9日 (2016 - 03 - 09) 全文	1-27	A	CN 106503025 A (北京搜狗科技发展有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文	1-27	A	US 2009043732 A1 (NHN CORPORATION) 2009年 2月 12日 (2009 - 02 - 12) 全文	1-27
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 107256251 A (广州优视网络科技有限公司) 2017年 10月 17日 (2017 - 10 - 17) 权利要求1-25、说明书第122-124段	1-27																					
X	CN 105893390 A (北京搜狗科技发展有限公司) 2016年 8月 24日 (2016 - 08 - 24) 说明书第51-126段	1-4、9-13、 18-21、26-27																					
A	CN 106777282 A (百度在线网络技术北京有限公司) 2017年 5月 31日 (2017 - 05 - 31) 全文	1-27																					
A	CN 105389328 A (中国人民解放军国防科学技术大学) 2016年 3月 9日 (2016 - 03 - 09) 全文	1-27																					
A	CN 106503025 A (北京搜狗科技发展有限公司) 2017年 3月 15日 (2017 - 03 - 15) 全文	1-27																					
A	US 2009043732 A1 (NHN CORPORATION) 2009年 2月 12日 (2009 - 02 - 12) 全文	1-27																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																							
国际检索实际完成的日期 2018年 8月 24日		国际检索报告邮寄日期 2018年 9月 7日																					
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10) 62019451		受权官员 杨庆丽 电话号码 86-(10)-53961433																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/090150

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	107256251	A	2017年 10月 17日	无			
CN	105893390	A	2016年 8月 24日	无			
CN	106777282	A	2017年 5月 31日	US	2018189291	A1	2018年 7月 5日
CN	105389328	A	2016年 3月 9日	无			
CN	106503025	A	2017年 3月 15日	无			
US	2009043732	A1	2009年 2月 12日	JP	2007524172	A	2007年 8月 23日
				KR	100462292	B1	2004年 12月 8日
				WO	2005083593	A1	2005年 9月 9日