

(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

(43) 国际公布日
2017 年 10 月 5 日 (05.10.2017)



(10) 国际公布号
WO 2017/167043 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2017/077245
- (22) 国际申请日: 2017 年 3 月 20 日 (20.03.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610203900.5 2016 年 4 月 1 日 (01.04.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属开曼群岛大开曼资本大厦一座四层 847 号邮箱, Cayman (KY)。
- (72) 发明人: 李晓菲 (LI, Xiaofei); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。
- (74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY

AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, 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, 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)。

[见续页]

(54) Title: USER-BASED PERSONALIZED DATA SEARCH METHOD AND APPARATUS

(54) 发明名称: 一种基于用户的个性化数据搜索方法和装置

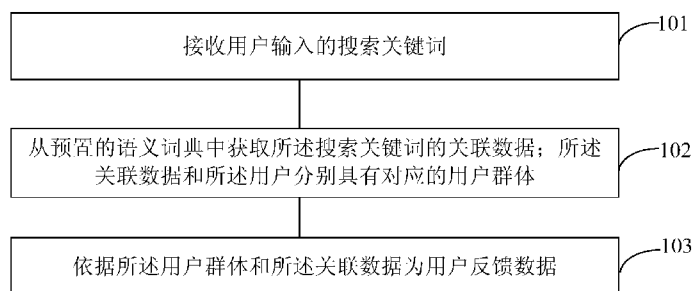


图 1

- 101 Receive a search keyword entered by a user
102 Obtain associated data of the search keyword from a preset semantic dictionary, the associated data and the user separately having a corresponding user group
103 Feed back data to the user according to the user groups and the associated data

(57) Abstract: Embodiments of the present application provide a user-based personalized data search method and apparatus. The method comprises: receiving a search keyword entered by a user; obtaining associated data of the search keyword from a preset semantic dictionary, the associated data and the user separately having a corresponding user group; and feeding back data to the user according to the user groups and the associated data. In the embodiments of the present application, by using a semantic dictionary, the problem of differences between search keywords entered by a user and words stored in a system is resolved, and the distance between data itself and value generated from the use of the data is shortened, thereby making it possible that required data is rapidly found from massive data.

(57) 摘要: 本申请实施例提供了一种基于用户的个性化数据搜索方法和装置, 所述的方法包括: 接收用户输入的搜索关键词; 从预置的语义词典中获取所述搜索关键词的关联数据; 所述关联数据和所述用户分别具有对应的用户群体; 依据所述用户群体和所述关联数据为用户反馈数据。本申请实施例通过利用语义词典, 解决用户输入的搜索关键词与系统存储词之间的差异问题, 缩短从数据本身到使用数据产生价值的距离, 使得从海量数据从快速搜索到所需数据成为可能。

WO 2017/167043 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种基于用户的个性化数据搜索方法和装置

本申请要求 2016 年 04 月 01 日递交的申请号为 201610203900.5、发明名称为“一种基于用户的个性化数据搜索方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及数据处理技术领域，特别是涉及一种基于用户的个性化数据搜索方法和一种基于用户的个性化数据搜索装置。

10 背景技术

随着大数据时代的到来，企业收集到的数据越来越多，与此同时，要在海量数据中找到所需数据往往如大海捞针一般，费时费力。目前，大部分企业存储的数据表基本都是以英文名称、英文缩略词进行存储，而用户在搜索数据时，是根据自己的语言习惯，使用中文全称或英文全称进行搜索。

15 用户在大数据平台上使用搜索关键词进行搜索时，最快得到搜索结果的方式是，按照数据表名称全匹配方式进行搜索，当然，所有用户在使用相同的搜索关键词时，搜索到数据的排序结果都是相同的。然而，在实际业务场景中，用户往往不知道大数据平台中，所想要数据表的具体名称，因此，只能使用自己所理解的中文来描述需求，由于大数据平台中的数据表通常是使用英文名称或者英文缩略词作为数据表名称，因此若使用
20 中文进行搜索，很难在海量数据中精准找到所需数据。另一方面，使用同一个搜索关键词，对于业务不同的用户，所期望的搜索结果通常有所不同，如果大数据平台对所有用户给出一样的搜索结果，只会浪费用户搜索数据的时间。

综上所述，目前在大数据平台上搜索数据的明显缺点是：

1、用户使用中文搜索，再将中文与大数据平台上存储的数据表的中文名称、中文
25 描述信息的匹配，但是大数据平台有上百万张数据表，数据负责人很难维护这些数据表的全部中文信息；

2、用户即使使用英文搜索，比如使用 seller，虽不再是使用中文，但是大数据平台上该数据命名名称有可能不是 seller，而是缩写的 srl，在这种情况下，也是无法搜索到所需的结果；

30 3、用户需要知道所有数据表的命名名称，才能快速搜索到所需数据，否则将无法

搜索到所需的结果，而这在数据为海量的大数据时代几乎是不可能做到的；

4、在上述 3 的情况下，用户只能通过咨询有经验的人，咨询数据开发接口人，才能快速搜索到所需数据，无形中增大了多方的时间成本；

5、使用同一个搜索关键词（key）进行搜索时，对做安全业务的用户和做售后服务的用户给出的结果相同，但是两者的需求是不同的，故而降低大数据平台的服务能力，用户体验效果较差。

发明内容

鉴于上述问题，提出了本申请实施例以便提供一种克服上述问题或者至少部分地解决上述问题的一种基于用户的个性化数据搜索方法和相应的一种基于用户的个性化数据搜索装置。

为了解决上述问题，本申请实施例公开了一种基于用户的个性化数据搜索方法，包括：

接收用户输入的搜索关键词；

15 从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和所述用户分别具有对应的用户群体；

依据所述用户群体和所述关联数据为用户反馈数据。

优选地，所述语义词典通过如下方式生成：

获取一个或多个用户群体的源数据文档；

20 从所述源数据文档中提取出所述一个或多个用户群体对应的关联数据；

按照用户群体将所述关联数据组织为的语义词典。

优选地，所述依据用户群体和所述关联数据为用户反馈数据的步骤包括：

依据所述用户群体确定所述关联数据的权重数值；

采用所述关联数据进行搜索得到搜索结果；

25 按照所述权重数值将所述关联数据对应的搜索结果反馈给用户。

优选地，所述依据用户群体确定所述关联数据的权重数值的步骤包括：

判断所述用户下的所述关联数据是否已记录有对应的权重数值；

若是，则采用所述已记录的权重数值作为所述关联数据的权重数值；

30 若否，则采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

优选地，所述采用用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值的步骤包括：

分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致；

若是，则为所述关联数据分配第一权重数值；

5 若否，则为所述关联数据分配第二权重数值；

其中，所述第一权重数值大于所述第二权重数值。

优选地，所述搜索结果具有对应的用户群体，所述搜索结果对应的关联数据具有对应的用户群体，在所述按照权重数值将所述关联数据对应的搜索结果反馈给用户的步骤之后，还包括：

10 判断所述用户点击的搜索结果对应的用户群体是否与所述用户的用户群体一致；

若否，则修改所述关联数据的权重数值。

优选地，所述修改所述关联数据的权重数值的步骤包括：

将所述关联数据的第一权重数值修改为第三权重数值，以及将关联数据的第二权重数值修改为第四权重数值；其中，所述第三权重数值等于所述第四权重数值。

15 优选地，所述关联数据包括中文名称，英文名称，英文缩略词，中文缩略词，相似词，近义词，和/或同义词。

本申请实施例还公开了一种基于用户的个性化数据搜索装置，包括：

搜索关键词接收模块，用于接收用户输入的搜索关键词；

关联数据获取模块，用于从预置的语义词典中获取所述搜索关键词的关联数据；所

20 述关联数据和所述用户分别具有对应的用户群体；

用户数据反馈模块，用于依据所述用户群体和所述关联数据为用户反馈数据。

优选地，所述装置还包括：

源数据文档获取模块，用于获取一个或多个用户群体的源数据文档；

关联数据提取模块，用于从所述源数据文档中提取出所述一个或多个用户群体对应

25 的关联数据；

语义词典组织模块，用于按照用户群体将所述关联数据组织为的语义词典。

优选地，所述用户数据反馈模块包括：

权重数值确定子模块，用于依据所述用户群体确定所述关联数据的权重数值；

结果数据搜索子模块，用于采用所述关联数据进行搜索得到搜索结果；

30 搜索结果反馈子模块，用于按照所述权重数值将所述关联数据对应的搜索结果反馈

给用户。

优选地，所述权重数值确定子模块包括：

权重数值判断单元，用于判断所述用户下的所述关联数据是否已记录有对应的权重数值；若是，则调用第一权重数值赋值单元，若否，则调用二权重数值赋值单元；

5 第一权重数值赋值单元，用于采用所述已记录的权重数值作为所述关联数据的权重数值；

第二权重数值赋值单元，用于采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

优选地，所述第二权重数值赋值单元包括：

10 用户群体判断子单元，用于分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致；若是，则调用第一权重数值分配子单元，若否，则调用第二权重数值分配子单元；

第一权重数值分配子单元，用于为所述关联数据分配第一权重数值；

第二权重数值分配子单元，用于为所述关联数据分配第二权重数值；

15 其中，所述第一权重数值大于所述第二权重数值。

优选地，所述搜索结果具有对应的用户群体，所述的装置还包括：

用户群体一致性判断模块，用于判断所述用户点击的搜索结果对应的用户群体是否与所述用户的用户群体一致；若否，则调用权重数值修改模块；

权重数值修改模块，用于修改所述关联数据的权重数值。

20 优选地，所述权重数值修改模块包括：

第三权重数值赋值子模块，用于将所述关联数据的第一权重数值修改为第三权重数值，以及将关联数据的第二权重数值修改为第四权重数值；其中，所述第三权重数值等于所述第四权重数值。

本申请实施例还公开了一种基于用户的个性化数据搜索方法，包括：

25 获取用户输入的搜索关键词；

将所述搜索关键词发送至服务器；所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据，所述关联数据和所述用户分别具有对应的用户群体；

接收所述服务器依据所述用户群体和关联数据为用户反馈的数据；

30 展现所述反馈的数据。

本申请实施例还公开了一种基于用户的个性化数据搜索装置，包括：

搜索关键词获取模块，用于获取用户输入的搜索关键词；

搜索关键词发送模块，用于将所述搜索关键词发送至服务器；所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据，所述关联数据和
5 所述用户分别具有对应的用户群体；

反馈数据接收模块，用于接收所述服务器依据所述用户群体和关联数据为用户反馈的数据；

反馈数据展现模块，用于展现所述反馈的数据。

10 本申请实施例包括以下优点：

本申请实施例针对大数据场景，考虑到数据在大数据平台上存储时使用的系统存储词，即数据存储时所使用的名称与用户输入的搜索关键词不一致，以及，归属于不同用户群体的用户所需数据有所不同的情况下，建立语义词典。语义词典中建立了用户输入的搜索关键词与系统存储词之间的关联，当用户输入搜索关键词进行搜索时，从语义词典中获取该搜索关键词的关联数据，也即是用户所需数据潜在的系统存储词，关联数据
15 和用户都分别属于某个或多个的用户群体，而不同用户群体的用户，其所需数据通常有所不同，故将按照关联数据和用户群体为用户反馈数据，以能够用户提供符合其需求的数据。本申请实施例通过利用语义词典，解决用户输入的搜索关键词与系统存储词之间的差异问题，缩短从数据本身到使用数据产生价值的距离，使得从海量数据从快速搜索
20 到所需数据成为可能。

本申请实施例为用户反馈数据时，先按照关联数据对应的用户群体和用户对应的用户群体分配权重数值，最后再使用关联数据进行搜索得到搜索结果后，按照权重数值进行排序后展现给用户。由于搜索结果是按照权重数值，也即是按照用户对于搜索结果的需要程度来按序反馈给用户，在排除其他影响因素的情况下，反馈的数据更加符合用户
25 的需求，提高了用户的搜索体验。

本申请实施例获取用户对于搜索结果的点击信息，再去进一步调整关联数据的权重数值，若用户点击的搜索结果，与其所属的用户群体不一致，将修改关联数据的权重数值为一致，那么当再次使用该关联数据进行搜索时，在排除其他影响因素的情况下，它们对应的搜索结果应当同等展现给用户；反之，若用户点击的搜索结果，与其所属的用户群体一致，则无需对应关联数据的权重数值进行修改，那么当再次使用该关联数据进
30

行搜索时，在排除其他影响因素的情况下，它们对应的搜索结果应当按照权重数值展现给用户。

附图说明

- 5 图 1 是本申请的一种基于用户的个性化数据搜索方法实施例 1 的步骤流程图；
图 2 是本申请的一种基于用户的个性化数据搜索方法实施例 2 的步骤流程图；
图 3 是本申请的一种基于用户的个性化数据搜索方法实施例 3 的步骤流程图；
图 4 是本申请的一种基于用户的个性化数据搜索方法实施例 3 的步骤流程图；
图 5 是本申请的一种基于语义词典的个性化大数据搜索流程图；
10 图 6 是本申请的一种用户搜索场景的示意图一；
图 7 是本申请的一种用户搜索场景的示意图二；
图 8 是本申请的一种用户搜索场景的示意图三；
图 9 是本申请的一种基于用户的个性化数据搜索装置实施例 1 的结构框图；
图 10 是本申请的一种基于用户的个性化数据搜索装置实施例 2 的结构框图。

15

具体实施方式

为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本申请作进一步详细的说明。

- 参照图 1，示出了本申请的一种个性化数据的搜索方法实施例 1 的步骤流程图，具体可以包括如下步骤：

步骤 101，接收用户输入的搜索关键词；

步骤 102，从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和所述用户分别具有对应的用户群体；

步骤 103，依据所述用户群体和所述关联数据为用户反馈数据。

- 25 在本申请的一种优选实施例中，所述语义词典可以通过如下方式生成：

获取一个或多个用户群体的源数据文档；

从所述源数据文档中提取出所述一个或多个用户群体对应的关联数据；

按照用户群体将所述关联数据组织为的语义词典。

- 在具体实现中，用户具有其所归属的用户群体，可以仅归属于一个用户群体，当然
30 也同时归属于多个用户群体。以一个公司内部的人员配置为例，某个用户所归属用户群

体可以部门 A，部门 B。可以理解，不同用户群体的用户，其所需数据有所不同，比如，部门 A 的用户更加关注交易相关的信息，部门 B 更加关注商品库存信息。

本申请实施例在考虑到数据存储时使用的系统存储词与用户输入的搜索关键词不一致，以及，归属于不同用户群体的用户所需数据有所不同的情况下，建立语义词典。

5 具体来说，首先通过分析不同业务的数据建模规范文档、数据设计说明文档、表使用说明文档等相关文档，抽取其中的文本，建立一个“语义词典”，由于在大数据场景下，基于数据通常以英文缩略词的命名方式存储的情况，故本申请实施例的语义词典具体包含如下内容，其中业务域是指代用户归属的用户群体：

1、业务域：不同业务域下的用户进行搜索将匹配不同业务域的语义词典；

10 2、英文名称：如 seller、computer、item、product；

3、中文名称：如卖家、电脑、商品、产品；

4、英文缩略词：如 slr、comp、itm、prod；

5、相似词、近义词、同义词；

以下举例说明具体的业务场景：

15 通过收集、分析各数据文档，假设得到如下语义词典：

业务域	缩略词	英文名称	中文名称	近义词
部门 A	trd	trade	交易	成交
部门 A	comp	computer	电脑	台式机，笔记本
部门 A	itm	item	商品	宝贝
部门 B	trd	trade	成交	交易
部门 B	prod	product	商品	产品，宝贝
部门 B	comp	compensation	赔偿金	

同时，语义词典还对应有用户信息表：

用户名	所属部门
U1	部门 A
	部门 B

本申请实施例在用户输入搜索关键词进行搜索时，将按照语义词典和对应的用户信息表，来为用户提供搜索服务。需要说明的是，本申请实施例除了适用于公司内部外，也适用于更加大的数据搜索场景，同样可通过事先收集用户相关信息再分类出用户群体，然后建立语义词典，并在此基础上进行搜索，本申请实施例对此不加以限制。

20

本申请实施例针对大数据场景，考虑到数据在大数据平台上存储时使用的系统存储

词，即数据存储时所使用的名称与用户输入的搜索关键词不一致，以及，归属于不同用户群体的用户所需数据有所不同的情况下，建立语义词典。语义词典中建立了用户输入的搜索关键词与系统存储词之间的关联，当用户输入搜索关键词进行搜索时，从语义词典中获取该搜索关键词的关联数据，也即是用户所需数据潜在的系统存储词，关联数据

5 和用户都分别属于某个或多个的用户群体，而不同用户群体的用户，其所需数据通常有所不同，故按照关联数据和用户群体为用户反馈数据，以给用户符合其需求的数据。本申请实施例通过利用语义词典，解决用户输入的搜索关键词与系统存储词之间的差异问题，缩短从数据本身到使用数据产生价值的距离，使得从海量数据从快速搜索到所需数据成为可能。

10

参照图 2，示出了本申请的一种个性化数据的搜索方法实施例 2 的步骤流程图，具体可以包括如下步骤：

步骤 201，接收用户输入的搜索关键词；

步骤 202，从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和

15 所述用户分别具有对应的用户群体；

在本申请实施例中，当用户输入的搜索关键词后，调用预先创立的语义词典，通过使用语义词典查找多用户输入的搜索关键词的关联数据，其中，关联数据也具有其归属的用户群体。

参考上述的语义词典和用户信息表，假设用户输入的关键词为“成交”，那么将得

20 到部门 A 下的“trade”，以及部门 B 下“trade”这两个关联数据，它们分属于的用户群体分别是部门 A 和部门 B。同时，用户具有归属的用户群体，调用用户信息表，假设用户的用户标识为 U1，那么可以从用户信息表查找到，U1 所属的用户群体为部门 A。

步骤 203，依据所述用户群体确定所述关联数据的权重数值；

在本申请的一种优选实施例中，所述步骤 203 可以包括如下子步骤：

子步骤 S11，判断所述用户下的所述关联数据是否已记录有对应的权重数值；若是，

25 则执行子步骤 S12，若否，则执行子步骤 S13；

子步骤 S12，采用所述已记录的权重数值作为所述关联数据的权重数值；

子步骤 S13，采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

30 在本申请的一种示例中，若用户在先已经使用过关联数据进行搜索，那么在系统中

已存储有该关联数据的权重数值，此时无需再次对于该关联数据分配权重数值，直接使用即可。

如果用户在先没有使用过关联数据进行搜索，那么就需要为该关联数据的权重数值，则将根据进行搜索的用户所属的用户群体，以及，该关联数据的用户群体，来进行权重数值的分配。

在本申请的一种优选实施例中，采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值，也即是所述子步骤 S13 可以包括如下子步骤：

子步骤 a1，分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致；若是，则执行子步骤 a2，若否，则执行子步骤 a3。

10 子步骤 a2，为所述关联数据分配第一权重数值；

子步骤 a3，若否，则为所述关联数据分配第二权重数值。其中，所述第一权重数值大于所述第二权重数值。

可以理解，若关联数据的用户群体与用户的用户群体一致，说明该关联数据与用户的关联度较高，那么关联数据对应的搜索结果应该更加接近用户的需求，故可以为该关联数据分配较大的权重数值，反之，若关联数据的用户群体与用户的用户群体不一致，说明该关联数据与用户的关联度较低，那么关联数据对应的搜索结果应该与用户需求有一定距离，故可以为该关联数据分配较小的权重数值。

步骤 204，采用所述关联数据进行搜索得到搜索结果；

步骤 205，按照所述权重数值将所述关联数据对应的搜索结果反馈给用户。

20 在本申请的一种示例中，采用关联数据进行搜索时将得到相应的搜索结果，在不考虑其他影响因素的情况下，按照已分配的关联数据的权重数值，按序将搜索结果展现给用户。

在实际搜索场景中，用户通常是根据自己的语言习惯，使用中文全称或英文全称进行搜索，而非数据通常所使用的英文缩略词。故而本申请实施例通过语义词典，使得用户可以使用自然语言，就能快速查找到对应的英文缩略词，而无需要了解底层表的表命名规范，就可以进行快速搜索得到所需结果，用户体验效果极佳。

25 本申请实施例为用户反馈数据时，先按照关联数据对应的用户群体和用户对应的用户群体分配权重数值，最后再使用关联数据进行搜索得到搜索结果后，按照权重数值进行排序后展现给用户。由于搜索结果是按照权重数值，也即是按照用户对于搜索结果的需要程度来反馈，在排除其他影响因素的情况下，反馈的数据更加符合用户的需求，提

高了用户的搜索体验。

参照图 3，示出了本申请的一种个性化数据的搜索方法实施例 3 的步骤流程图，具体可以包括如下步骤：

5 步骤 301，接收用户输入的搜索关键词；

步骤 302，从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和所述用户分别具有对应的用户群体；

步骤 303，依据所述用户群体确定所述关联数据的权重数值；

步骤 304，采用所述关联数据进行搜索得到搜索结果；

10 步骤 305，按照所述权重数值将所述关联数据对应的搜索结果反馈给用户；

步骤 306，判断所述用户选择的搜索结果对应的用户群体是否与所述用户的用户群体一致；若否，则执行步骤 307；

步骤 307，修改所述关联数据的权重数值。

在本申请的一种优选实施例中，所述步骤 307 可以包括如下子步骤：

15 子步骤 S21，将所述关联数据的第一权重数值修改为第三权重数值，以及将关联数据的第二权重数值修改为第四权重数值；其中，所述第三权重数值等于所述第四权重数值。

在本申请实施例中，当反馈给用户搜索结果时，将收集用户对搜索结果的点击信息，如果用户点击的搜索结果所对应的用户群体与用户对应的用户群体不一致，则需要重新
20 改变关联数据的权重数值，反之，如果用户点击的搜索结果所对应的用户群体与用户对应的用户群体一致，则无需重新改变关联数据的权重数值。

具体来说，如果用户点击的搜索结果所对应的用户群体与用户对应的用户群体不一致，则将关联数据的权重数值修改为一致，使得两者的优先级相同。例如，假设在先两个关联数据的权重数值分别为 0.9 和 0.1，修改后的关联数据的权重数值则为 0.5 和 0.5。
25 当然，上述权重数值仅仅是作为示例，在实际中也可以采用其他数值，本申请实施例对此不加以限制。

本申请的一种示例中，将对于修改后的关联数据的权重数值进行存储，当下次用户使用相同的关联数据进行搜索时，可直接使用该权重数值，当然，如果没有修改，也可以将关联数据的权重数值进行存储，本申请实施例对此并不加以限制。

30 本申请实施例获取用户对于搜索结果的点击信息，再去进一步调整关联数据的权重

数值，若用户点击的搜索结果，与其所属的用户群体不一致，将修改关联数据的权重数值为一致，那么当再次使用该关联数据进行搜索时，在排除其他影响因素的情况下，它们对应的搜索结果应当同等展现给用户；反之，若用户点击的搜索结果，与其所属的用户群体一致，则无需对应关联数据的权重数值进行修改，那么当再次使用该关联数据进行搜索时，在排除其他影响因素的情况下，它们对应的搜索结果应当按照权重数值展现给用户。

参照图 4，示出了本申请的一种个性化数据的搜索方法实施例 4 的步骤流程图，具体可以包括如下步骤：

10 步骤 401，获取用户输入的搜索关键词；

步骤 402，将所述搜索关键词发送至服务器；所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据，所述关联数据和所述用户分别具有对应的用户群体；

步骤 403，接收所述服务器依据所述用户群体和关联数据为用户反馈的数据；

15 步骤 404，展现所述反馈的数据。

在本申请实施例中，当用户在客户端中输入搜索关键词时，客户端获取到该搜索关键词，然后将该搜索关键词发送到服务器，服务器则将基于该搜索关键词从预置的语义词典中，查找到该搜索关键词的关联数据，再基于该关联数据，关联数据的用户群体以及用户的用户群体，来为用户反馈数据，客户端接收到服务器的反馈数据后，将在客户端上展现给用户。

应用本申请实施例，即使用户所输入的搜索关键词与数据在大数据平台上存储时所使用的命名名称有所不同，服务器也可以从预置的语义词典中获取该搜索关键词的关联数据，即所需数据在大数据平台上有可能使用的命名名称，公知的是，直接采用数据的命名名称来查找数据，能够快速并准确地查找到对应的数据，基于这个特点，用户能够快速准确地查找到所需数据，用户体验效果极佳。

为了使本领域技术人员更好地理解本申请实施例，以下对于本申请的基于语义词典的个性化大数据搜索方案进行说明。

参照图 5 所示的本申请的一种基于语义词典的个性化大数据搜索流程图，具体的实现过程可以分为如下两个部分：

30 1、建立语义词典。

收集分析不同业务域的数据建模规范文档、数据设计说明文档、表使用说明文档等，抽取其中的文本，建立一个“语义词典”；其中，语义词典中包括有用户群体（以下称为业务域），以及互相关联的中文名称，英文名称，英文缩略词，中文缩略词，相似词，近义词，和/或同义词等等。

5 2、用户搜索场景。

当用户在任意用户界面上进行数据搜索时，通过匹配用户信息表确定用户的所属业务域，并通过匹配语义词典确定用户输入搜索关键词（输入搜索 key）的关联数据，关联数据包括与搜索关键词关联的英文缩略词、中文名称，英文名称等信息。当获取关联数据后，按照用户所属业务域和关联数据所属的业务域，来为关联数据分配权重数值。

10 将关联数据导入查询改写模块进行查询条件改写，在查询改写模块中同时记录有用户、关联数据、权重数值，并将这些信息提交给搜索引擎进行搜索查询，最终反馈给用户搜索结果。其中，搜索结果按照权重数值在用户界面上按序展现给用户。

3、权重数值改写。

本示例中收集用户对搜索结果的点击数据，将重新改写关联数据的权重数值，并反馈给查询改写模块，用以优化该用户下关联数据的权重数值。

以下采用几个具体的用户搜索场景来进一步说明，需参考前文中的语义词典和用户信息表，三个搜索场景分别如下所示。

一，参照图 6 所示的场景 1：

当部门 A 的用户 U1，在用户界面上输入“商品”时，将得到部门 A 下的“itm”和
20 部门 B 下的“prod”这两个关联数据，由于部门 A 下的“itm”与用户 U1 同属于部门 A，故可以为部门 A 下的“itm”分配较大的权重数值 0.9，并且为部门 B 下的“prod”分配较小的权重数值 0.1；

当采用上述两个关联数据得到搜索结果时，在用户界面上按照权重数值的大小，按序将关联数据的搜索结果展现给用户。即将返回含“itm”的表排名靠前，而将含“prod”
25 的表排名靠后；

若用户最终点击了含“prod”的表，则需要返回查询改写模块，将该关联数据的权重数值进行修改，具体来说，可以将部门 A 下的“itm”的权重数值和部门 B 下的“prod”的权重数值修改为 0.5。

二，参照图 7 所示的场景 2：

30 当部门 A 的用户 U1，在用户界面上输入“赔偿金”时，将得到部门 A 下的“comp”

和部门 B 下的“comp”这两个关联数据，由于部门 A 下的“comp”与用户 U1 同属于部门 A，故可以为部门 A 下的“comp”分配较大的权重数值 0.9，并且为部门 B 下的“comp”分配较小的权重数值 0.1；

当采用上述两个关联数据得到搜索结果时，在用户界面上按照权重数值的大小，按序将关联数据的搜索结果展现给用户。即将返回部门 A 下的“comp”的表排名靠前，而将部门 B 下的“comp”的表排名靠后；

若用户最终点击了部门 B 下的“comp”的表，则需要返回查询改写模块，将该关联数据的权重数值进行修改，具体来说，可以将部门 A 下的“comp”的权重数值和部门 B 下的“comp”的权重数值修改为 0.5。

三，参照图 8 所示的场景 2：

当部门 A 的用户 U1，在用户界面上输入“交易”时，将得到部门 A 下的“trd”和部门 B 下的“trd”这两个关联数据，由于部门 A 下的“trd”与用户 U1 同属于部门 A，故可以为部门 A 下的“trd”分配较大的权重数值 0.9，并且为部门 B 下的“trd”分配较小的权重数值 0.1；

当采用上述两个关联数据得到搜索结果时，在用户界面上按照权重数值的大小，按序将关联数据的搜索结果展现给用户。即将返回部门 A 下的“trd”的表排名靠前，而将部门 B 下的“trd”的表排名靠后；

若用户最终点击了部门 B 下的“trd”的表，则需要返回查询改写模块，将该关联数据的权重数值进行修改，具体来说，可以将部门 A 下的“trd”的权重数值和部门 B 下的“trd”的权重数值修改为 0.5。

当然，以上的几种搜索场景仅仅是作为一个例子，具体应用上，还可以根据需要进行扩充，本申请实施例对此不加以限制。

本申请实施例主要针对大数据场景，基于在大数据平台中的数据通常以英文缩略词作为命名存储的情况，并且考虑用户关注/所属领域的不同，而使得英文缩略词语义有所不同的场景，从数据建模规范等数据文档中抽取语义词典，应用于用户使用中文、或英文全称搜索时的智能匹配，使得从海量数据从快速搜索到所需数据成为可能。

需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是

本申请实施例所必须的。

参照图 9，示出了本申请的一种基于用户的个性化数据搜索装置实施例 1 的结构框图，具体可以包括如下模块：

5 搜索关键词接收模块 501，用于接收用户输入的搜索关键词；

关联数据获取模块 502，用于从预置的语义词典中获取所述搜索关键词的关联数据；
所述关联数据和所述用户分别具有对应的用户群体；

用户数据反馈模块 503，用于依据所述用户群体和所述关联数据为用户反馈数据。

在本申请的一种优选实施例中，所述用户数据反馈模块 503 可以包括如下子模块：

10 权重数值确定子模块，用于依据所述用户群体确定所述关联数据的权重数值；

在本申请的一种优选实施例中，所述权重数值确定模块可以包括如下单元：

权重数值判断单元，用于判断所述用户下的所述关联数据是否已记录有对应的权重数值；若是，则调用第一权重数值赋值单元，若否，则调用二权重数值赋值单元；

15 第一权重数值赋值单元，用于采用所述已记录的权重数值作为所述关联数据的权重数值；

第二权重数值赋值单元，用于采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

在本申请的一种优选实施例中，所述第二权重数值赋值单元可以包括如下单元：

20 用户群体判断子单元，用于分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致；若是，则调用第一权重数值分配子单元，若否，则调用第二权重数值分配子单元；

第一权重数值分配子单元，用于为所述关联数据分配第一权重数值；

第二权重数值分配子单元，用于为所述关联数据分配第二权重数值；

其中，所述第一权重数值大于所述第二权重数值。

25 结果数据搜索模块，用于采用所述关联数据进行搜索得到搜索结果；

搜索结果反馈模块，用于按照所述权重数值将所述关联数据对应的搜索结果反馈给用户。

在本申请的一种优选实施例中，所述装置还可以包括如下模块：

源数据文档获取模块，用于获取一个或多个用户群体的源数据文档；

30 关联数据提取模块，用于从所述源数据文档中提取出所述一个或多个用户群体对应

的关联数据；

语义词典组织模块，用于按照用户群体将所述关联数据组织为的语义词典。

在本申请的一种优选实施例中，所述搜索结果可以具有对应的用户群体，所述的装置还可以包括如下模块：

5 用户群体一致性判断模块，用于判断所述用户点击的搜索结果对应的用户群体是否与所述用户的用户群体一致；若否，则调用权重数值修改模块；

权重数值修改模块，用于修改所述关联数据的权重数值。

在本申请的一种优选实施例中，所述权重数值修改模块可以包括如下子模块：

10 第三权重数值赋值子模块，用于将所述关联数据的第一权重数值修改为第三权重数值，以及将关联数据的第二权重数值修改为第四权重数值；其中，所述第三权重数值等于所述第四权重数值。

在本申请的一种优选实施例中，所述关联数据可以包括中文名称，英文名称，英文缩略词，中文缩略词，相似词，近义词，和/或同义词。

15 参照图 10，示出了本申请的一种基于用户的个性化数据搜索装置实施例 2 的结构框图，具体可以包括如下模块：

搜索关键词获取模块 601，用于获取用户输入的搜索关键词；

20 搜索关键词发送模块 602，用于将所述搜索关键词发送至服务器；所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据，所述关联数据和所述用户分别具有对应的用户群体；

反馈数据接收模块 603，用于接收所述服务器依据所述用户群体和关联数据为用户反馈的数据；

反馈数据展现模块 604，用于展现所述反馈的数据。

25 对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

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。

30 本领域内的技术人员应明白，本申请实施例的实施例可提供为方法、装置、或计算机程序产品。因此，本申请实施例可采用完全硬件实施例、完全软件实施例、或结合软

件和硬件方面的实施例的形式。而且，本申请实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

在一个典型的配置中，所述计算机设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非持续性的电脑可读媒体(transitory media)，如调制的数据信号和载波。

本申请实施例是参照根据本申请实施例的方法、终端设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上，使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本申请实施例的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例做出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本申请实施例范围的所有变更和修改。

最后，还需要说明的是，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、物品或者终端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、物品或者终端设备所固有的要素。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、物品或者终端设备中还存在另外的相同要素。

以上对本申请所提供的一种基于用户的个性化数据搜索方法和一种基于用户的个性化数据搜索装置，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权利要求书

1. 一种基于用户的个性化数据搜索方法，其特征在于，包括：

接收用户输入的搜索关键词；

从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和所述用户分

5 别具有对应的用户群体；

依据所述用户群体和所述关联数据为用户反馈数据。

2. 根据权利要求 1 所述的方法，其特征在于，所述语义词典通过如下方式生成：

获取一个或多个用户群体的源数据文档；

从所述源数据文档中提取出所述一个或多个用户群体对应的关联数据；

10 按照用户群体将所述关联数据组织为的语义词典。

3. 根据权利要求 1 或 2 所述的方法，其特征在于，所述依据用户群体和所述关联数据为用户反馈数据的步骤包括：

依据所述用户群体确定所述关联数据的权重数值；

采用所述关联数据进行搜索得到搜索结果；

15 按照所述权重数值将所述关联数据对应的搜索结果反馈给用户。

4. 根据权利要求 3 所述的方法，其特征在于，所述依据用户群体确定所述关联数据的权重数值的步骤包括：

判断所述用户下的所述关联数据是否已记录有对应的权重数值；

若是，则采用所述已记录的权重数值作为所述关联数据的权重数值；

20 若否，则采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

5. 根据权利要求 4 所述的方法，其特征在于，所述采用用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值的步骤包括：

分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致；

25 若是，则为所述关联数据分配第一权重数值；

若否，则为所述关联数据分配第二权重数值；

其中，所述第一权重数值大于所述第二权重数值。

6. 根据权利要求 3 所述的方法，其特征在于，所述搜索结果具有对应的用户群体，所述搜索结果对应的关联数据具有对应的用户群体，在所述按照权重数值将所述关联数据对应的搜索结果反馈给用户的步骤之后，还包括：

30

判断所述用户点击的搜索结果对应的用户群体是否与所述用户的用户群体一致；
若否，则修改所述关联数据的权重数值。

7. 根据权利要求 6 所述的方法，其特征在于，所述修改所述关联数据的权重数值的步骤包括：

5 将所述关联数据的第一权重数值修改为第三权重数值，以及将关联数据的第二权重数值修改为第四权重数值；其中，所述第三权重数值等于所述第四权重数值。

8. 根据权利要求 1 所述的方法，其特征在于，所述关联数据包括中文名称，英文名称，英文缩略词，中文缩略词，相似词，近义词，和/或同义词。

9. 一种基于用户的个性化数据搜索装置，其特征在于，包括：

10 搜索关键词接收模块，用于接收用户输入的搜索关键词；

关联数据获取模块，用于从预置的语义词典中获取所述搜索关键词的关联数据；所述关联数据和所述用户分别具有对应的用户群体；

用户数据反馈模块，用于依据所述用户群体和所述关联数据为用户反馈数据。

10. 根据权利要求 9 所述的装置，其特征在于，所述装置还包括：

15 源数据文档获取模块，用于获取一个或多个用户群体的源数据文档；

关联数据提取模块，用于从所述源数据文档中提取出所述一个或多个用户群体对应的关联数据；

语义词典组织模块，用于按照用户群体将所述关联数据组织为的语义词典。

11. 根据权利要求 9 或 10 所述的装置，其特征在于，所述用户数据反馈模块包括：

20 权重数值确定子模块，用于依据所述用户群体确定所述关联数据的权重数值；

结果数据搜索子模块，用于采用所述关联数据进行搜索得到搜索结果；

搜索结果反馈子模块，用于按照所述权重数值将所述关联数据对应的搜索结果反馈给用户。

12. 根据权利要求 11 所述的装置，其特征在于，所述权重数值确定子模块包括：

25 权重数值判断单元，用于判断所述用户下的所述关联数据是否已记录有对应的权重数值；若是，则调用第一权重数值赋值单元，若否，则调用二权重数值赋值单元；

第一权重数值赋值单元，用于采用所述已记录的权重数值作为所述关联数据的权重数值；

30 第二权重数值赋值单元，用于采用所述用户对应的用户群体和所述关联数据的用户群体确定所述关联数据的权重数值。

13. 根据权利要求 12 所述的装置, 其特征在于, 所述第二权重数值赋值单元包括:
用户群体判断子单元, 用于分别判断所述关联数据的用户群体与所述用户对应的用户群体是否一致; 若是, 则调用第一权重数值分配子单元, 若否, 则调用第二权重数值分配子单元;

5 第一权重数值分配子单元, 用于为所述关联数据分配第一权重数值;
第二权重数值分配子单元, 用于为所述关联数据分配第二权重数值;
其中, 所述第一权重数值大于所述第二权重数值。

14. 根据权利要求 11 所述的装置, 其特征在于, 所述搜索结果具有对应的用户群体, 所述的装置还包括:

10 用户群体一致性判断模块, 用于判断所述用户点击的搜索结果对应的用户群体是否与所述用户的用户群体一致; 若否, 则调用权重数值修改模块;
权重数值修改模块, 用于修改所述关联数据的权重数值。

15 15. 根据权利要求 14 所述的装置, 其特征在于, 所述权重数值修改模块包括:
第三权重数值赋值子模块, 用于将所述关联数据的第一权重数值修改为第三权重数值, 以及将关联数据的第二权重数值修改为第四权重数值; 其中, 所述第三权重数值等于所述第四权重数值。

16. 一种基于用户的个性化数据搜索方法, 其特征在于, 包括:
获取用户输入的搜索关键词;
将所述搜索关键词发送至服务器; 所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据, 所述关联数据和所述用户分别具有对应的用户群体;

接收所述服务器依据所述用户群体和关联数据为用户反馈的数据;
展现所述反馈的数据。

17. 一种基于用户的个性化数据搜索装置, 其特征在于, 包括:
25 搜索关键词获取模块, 用于获取用户输入的搜索关键词;
搜索关键词发送模块, 用于将所述搜索关键词发送至服务器; 所述服务器用于采用所述搜索关键词从预置的语义词典中获取所述搜索关键词的关联数据, 所述关联数据和所述用户分别具有对应的用户群体;

反馈数据接收模块, 用于接收所述服务器依据所述用户群体和关联数据为用户反馈的数据;

反馈数据展现模块，用于展现所述反馈的数据。

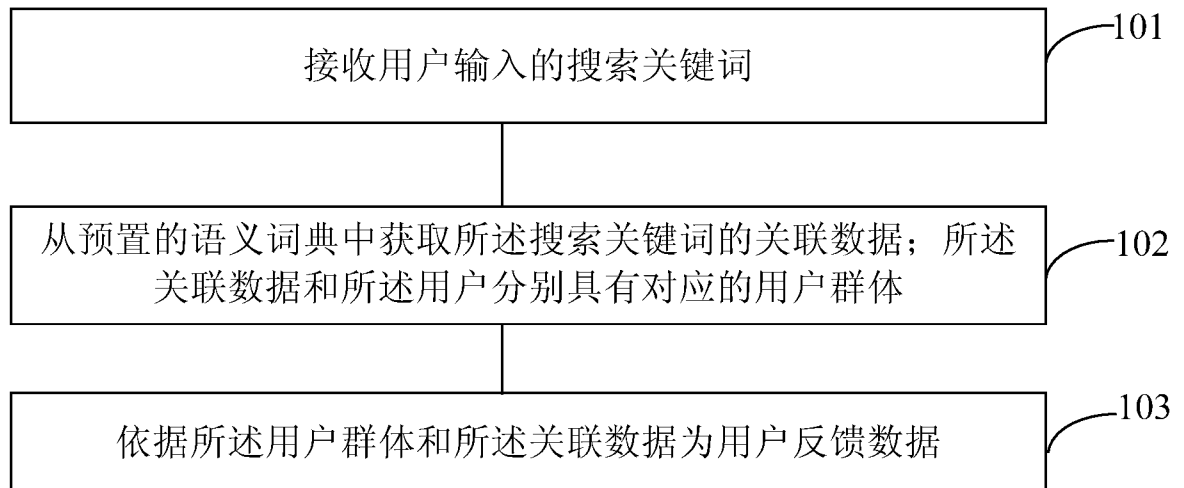


图 1

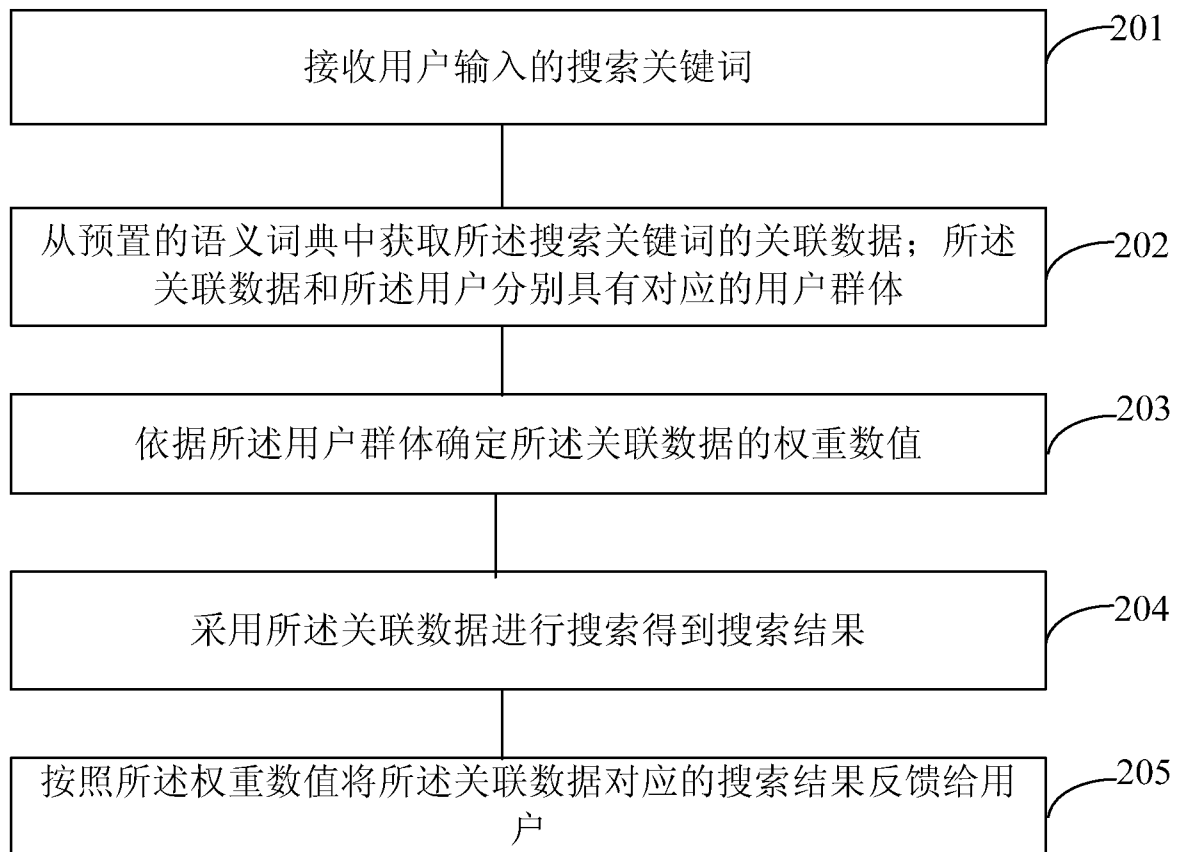


图 2

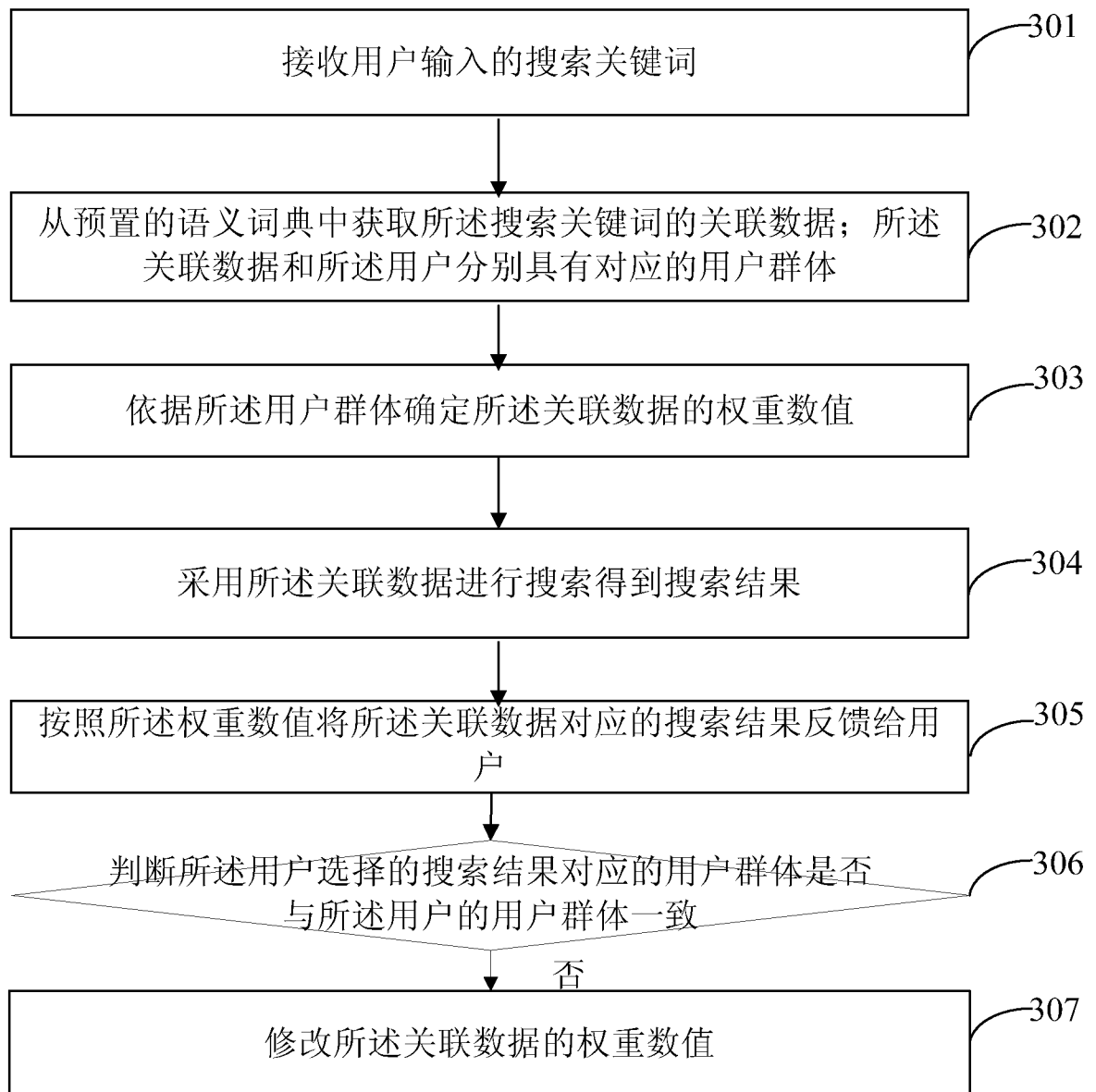


图 3

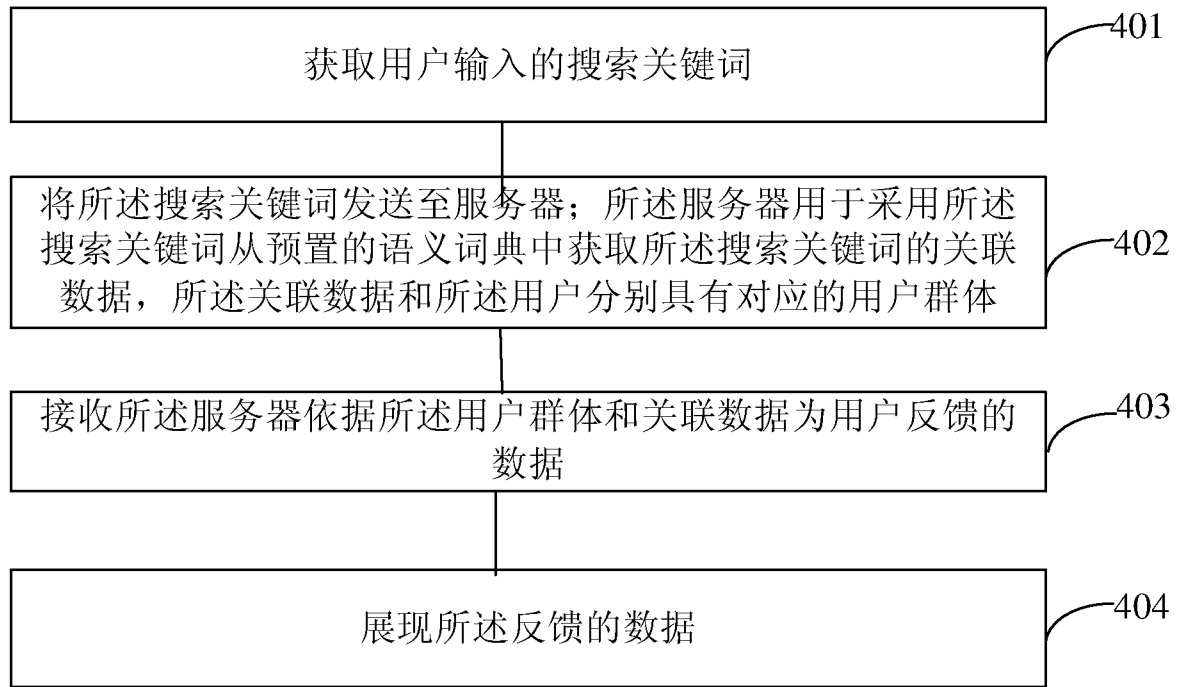


图 4

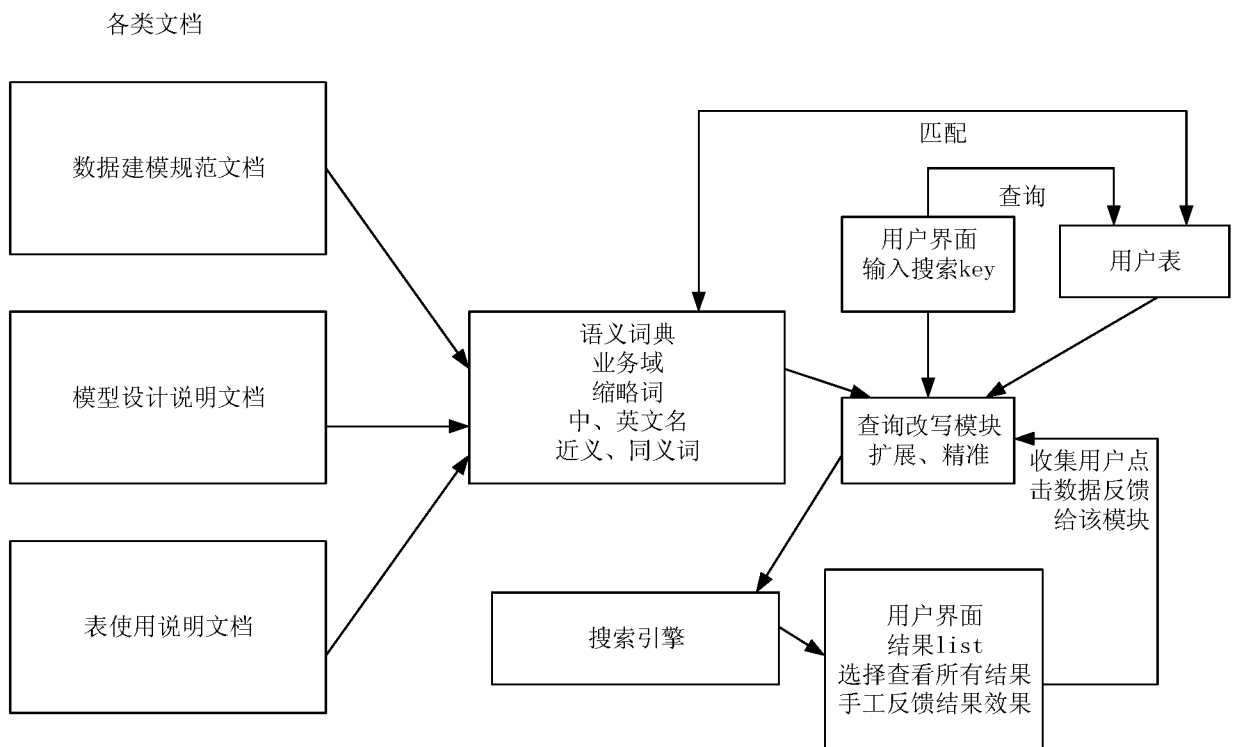


图 5

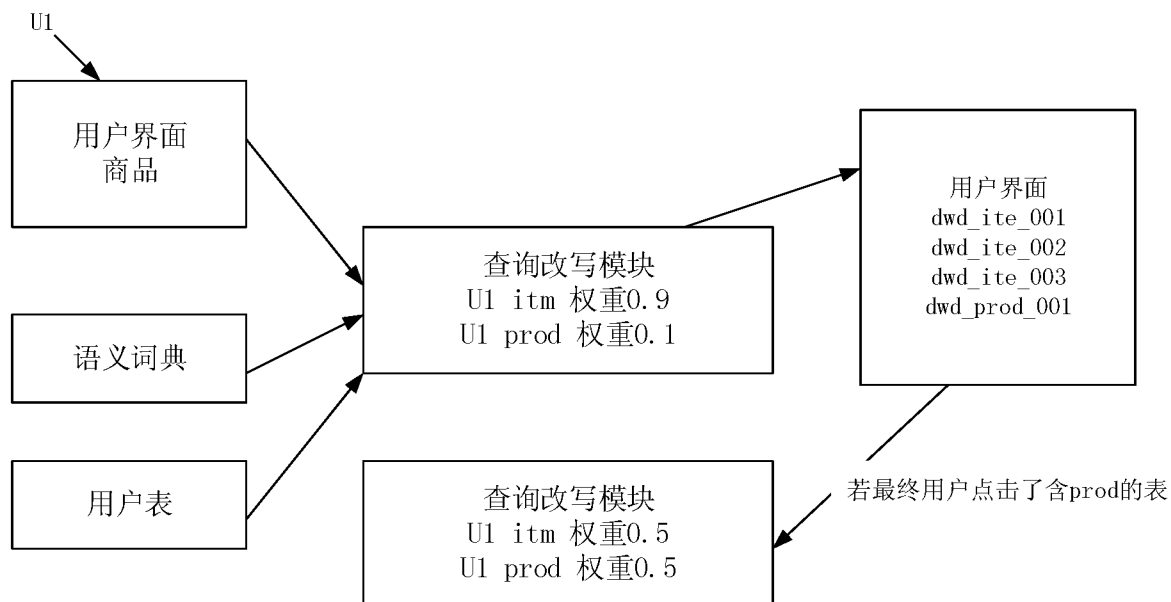


图 6

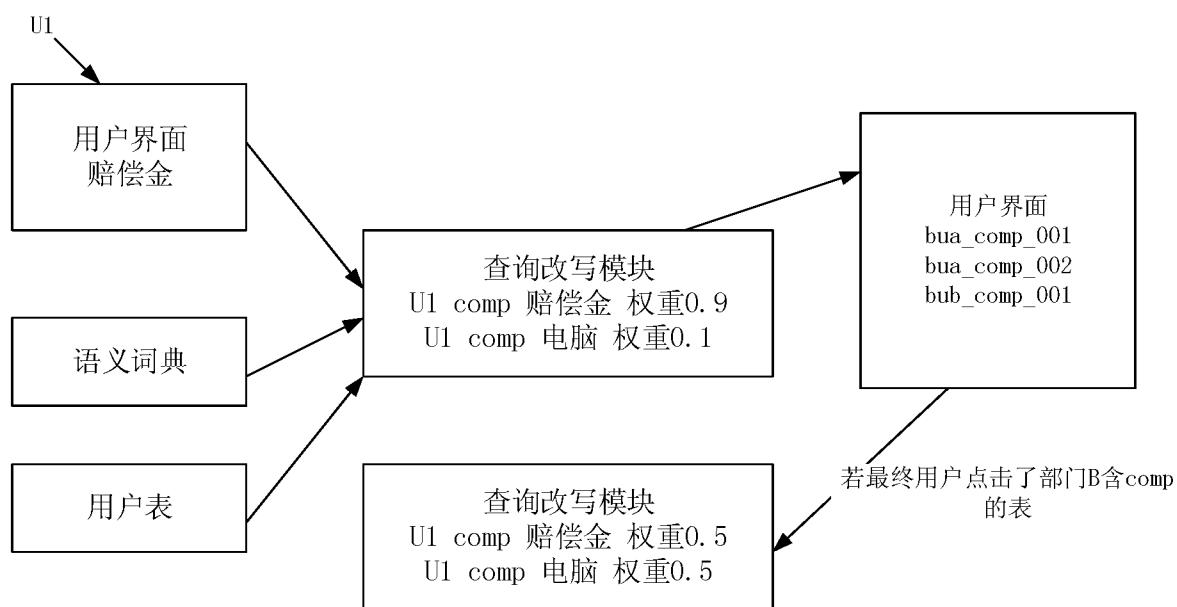


图 7

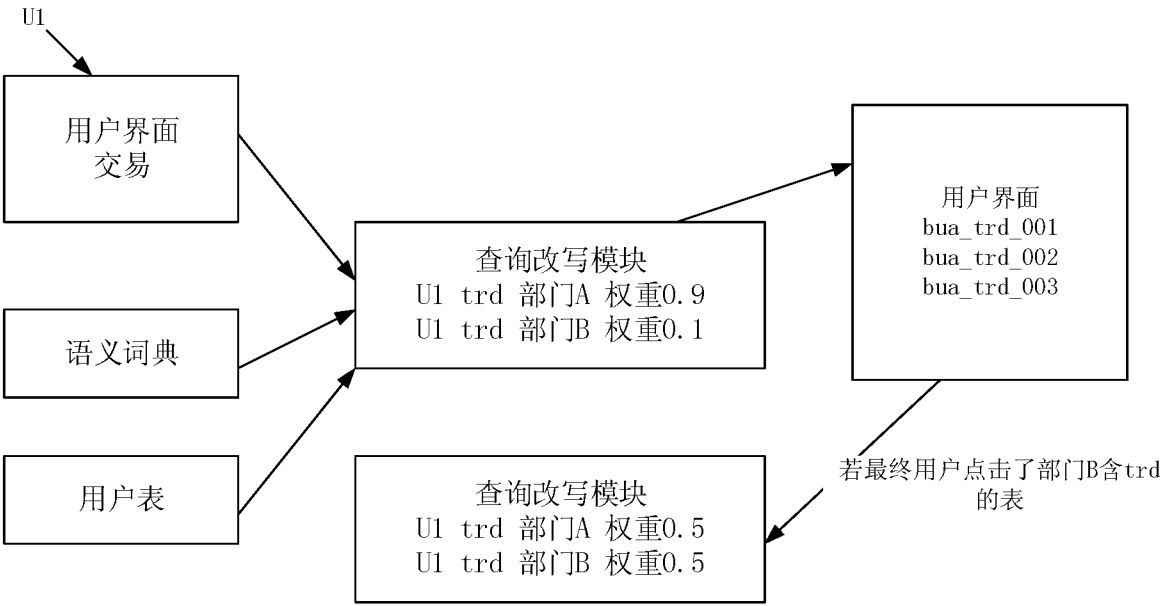


图 8

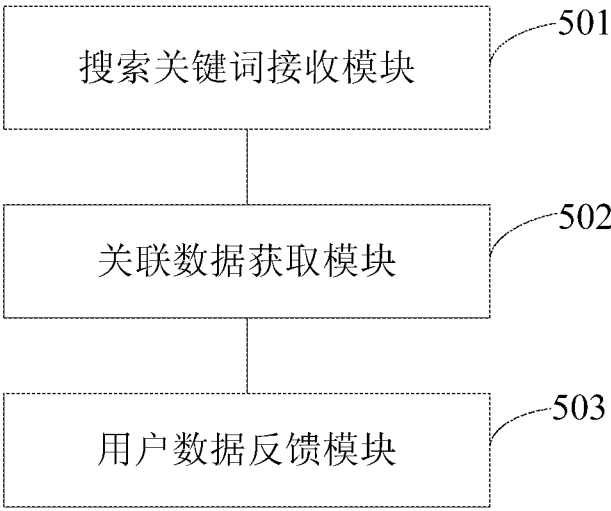


图 9

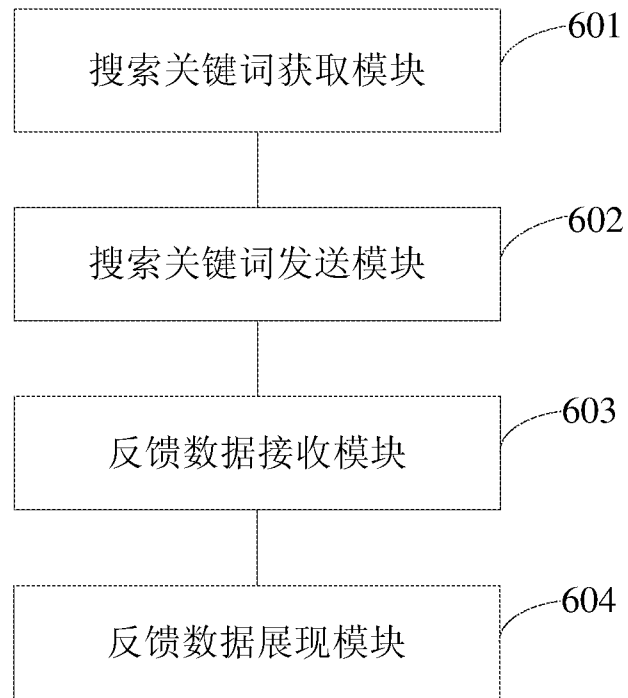


图 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/077245

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)

CPRS, CNKI, EPODOC, WPI: search, synonym, abbreviation, semantics, exten+, weight, community, group, associate+, relevant, key word, thesaurus, query, retrieval

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103020049 A (CHINA TELECOMMUNICATIONS CO., LTD.) 03 April 2013 (03.04.2013) description, paragraphs [0009]-[0041]	1, 2, 8, 9, 10, 16, 17
A	CN 103955480 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 30 July 2014 (30.07.2014) the whole document	1-17
A	CN 105302810 A (BEIJING SOGOU TECHNOLOGY DEVELOPMENT CO., LTD.) 03 February 2016 (03.02.2016) the whole document	1-17
A	US 2009083029 A1 (DOI et al.) 26 March 2009 (26.03.2009) the whole document	1-17

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 24 April 2017	Date of mailing of the international search report 28 April 2017
Name and mailing address of the ISA 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 BAI, Xuetao Telephone No. (86-10) 62412069

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/077245

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103020049 A	03 April 2013	None	
CN 103955480 A	30 July 2014	None	
CN 105302810 A	03 April 2013	None	
US 2009083029 A1	26 March 2009	JP 2009080579 A	16 April 2009
		JP 5178109 B2	10 April 2013
		US 8374845 B2	12 February 2013

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)) CPRS, CNKI, EPDOC, WPI 搜索, 检索, 查询, 同义词, 近义词, 缩略词, 关键词, 群, 权重, 语义, search, synonym, abbreviation, semantics, exten+, weight, community, group, associat+, relevant																	
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>X</td> <td>CN 103020049 A (中国电信股份有限公司) 2013年 4月 3日 (2013 - 04 - 03) 说明书【0009】-【0041】段</td> <td>1, 2, 8, 9, 10, 16, 17</td> </tr> <tr> <td>A</td> <td>CN 103955480 A (百度在线网络技术北京有限公司) 2014年 7月 30日 (2014 - 07 - 30) 全文</td> <td>1-17</td> </tr> <tr> <td>A</td> <td>CN 105302810 A (北京搜狗科技发展有限公司) 2016年 2月 3日 (2016 - 02 - 03) 全文</td> <td>1-17</td> </tr> <tr> <td>A</td> <td>US 2009083029 A1 (DOI 等) 2009年 3月 26日 (2009 - 03 - 26) 全文</td> <td>1-17</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	X	CN 103020049 A (中国电信股份有限公司) 2013年 4月 3日 (2013 - 04 - 03) 说明书【0009】-【0041】段	1, 2, 8, 9, 10, 16, 17	A	CN 103955480 A (百度在线网络技术北京有限公司) 2014年 7月 30日 (2014 - 07 - 30) 全文	1-17	A	CN 105302810 A (北京搜狗科技发展有限公司) 2016年 2月 3日 (2016 - 02 - 03) 全文	1-17	A	US 2009083029 A1 (DOI 等) 2009年 3月 26日 (2009 - 03 - 26) 全文	1-17
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 103020049 A (中国电信股份有限公司) 2013年 4月 3日 (2013 - 04 - 03) 说明书【0009】-【0041】段	1, 2, 8, 9, 10, 16, 17															
A	CN 103955480 A (百度在线网络技术北京有限公司) 2014年 7月 30日 (2014 - 07 - 30) 全文	1-17															
A	CN 105302810 A (北京搜狗科技发展有限公司) 2016年 2月 3日 (2016 - 02 - 03) 全文	1-17															
A	US 2009083029 A1 (DOI 等) 2009年 3月 26日 (2009 - 03 - 26) 全文	1-17															
☐ 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 2017年 4月 24日		国际检索报告邮寄日期 2017年 4月 28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 白雪涛 电话号码 (86-10)62412069															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/077245

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	103020049	A	2013年 4月 3日	无			
CN	103955480	A	2014年 7月 30日	无			
CN	105302810	A	2013年 4月 3日	无			
US	2009083029	A1	2009年 3月 26日	JP	2009080579	A	2009年 4月 16日
				JP	5178109	B2	2013年 4月 10日
				US	8374845	B2	2013年 2月 12日