

(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 2019/141109 A1

(43) 国际公布日
2019 年 7 月 25 日 (25.07.2019)

(51) 国际专利分类号:
G06F 16/00 (2019.01)

(21) 国际申请号: PCT/CN2019/070830

(22) 国际申请日: 2019 年 1 月 8 日 (08.01.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810043114.2 2018 年 1 月 17 日 (17.01.2018) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 刘阳阳 (LIU, Yangyang); 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: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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, 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))。

(54) Title: METHOD AND DEVICE FOR CONTENT RECOMMENDATION

(54) 发明名称: 内容推荐方法及装置

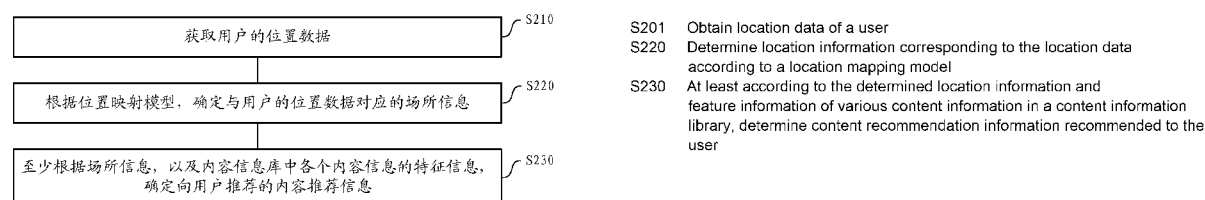


图 2

(57) Abstract: Provided by an embodiment disclosed in the description is a content recommendation method, the method comprising: obtaining location data of a user, and determining location information corresponding to the location data according to a location mapping model; then, at least according to the determined location information and feature information of various content information in a content information library, determining content recommendation information recommended to the user.

(57) 摘要: 本说明书披露的实施例提供一种内容推荐方法, 该方法包括: 获取用户的位置数据, 并根据位置映射模型, 确定与位置数据对应的场所信息。然后, 至少根据确定的场所信息, 以及内容信息库中各个内容信息的特征信息, 确定向所述用户推荐的内容推荐信息。



WO 2019/141109 A1

内容推荐方法及装置

技术领域

[01] 本说明书披露的多个实施例涉及互联网技术领域，尤其涉及一种内容推荐方法及装置。

5 背景技术

[02] 随着互联网技术的发展，人们越来越频繁地浏览网络平台提供的内容信息。例如，在网络购物平台中浏览商品信息，或者在新闻平台浏览热点信息，或者在理财平台浏览理财资讯等。

[03] 不同用户在使用同一网络平台时，对其提供的内容信息的需求有着或多或少的差异。

10 另一方面，网络平台中海量增长也常常让用户难以选择。目前，向用户推荐的内容信息由于存在不够精准、不够及时等不足，难以满足用户的要求。因此，需要提供一种合理的方法，以满足用户浏览网络平台中提供的内容信息的多种需求。

发明内容

15 [04] 本说明书描述了一种内容推荐方法及装置，根据用户的场所信息和内容信息库中各个内容信息的特征信息确定向用户推荐的内容推荐信息，从而及时地向用户推荐更加精准的内容信息。

[05] 第一方面，提供了一种内容推荐方法。该方法包括：

[06] 获取用户的位置数据；

20 [07] 根据位置映射模型，确定与所述位置数据对应的场所信息，所述位置映射模型基于预先获取的服务信息而训练获得，所述服务信息中包含服务位置数据和对应的服务场所数据；

[08] 至少根据所述场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息。

[09] 在一种可能的实施方式中，所述位置映射模型通过以下步骤训练获得：

25 [10] 对所述服务信息中包含的服务位置数据进行聚类，获得多个类簇；

[11]根据所述服务信息中与所述服务位置数据对应的服务场所数据，确定所述多个类簇对应的场所标签。

[12]在一种可能的实施方式中，所述确定与所述位置数据对应的场所信息包括：

[13]从所述多个类簇中确定与所述用户的位置数据对应的类簇；

5 [14]根据确定出的类簇所对应的场所标签，确定所述场所信息。

[15]在一种可能的实施方式中，所述确定向所述用户推荐的内容推荐信息，包括：

[16]确定与所述场所信息对应的服务类别信息；

[17]根据所述各个内容信息的特征信息，从所述内容信息库中确定与所述服务类别信息相关的内容信息，并将所述内容信息作为所述内容推荐信息。

10 [18]在一种可能的实施方式中，所述特征信息包括所述内容信息的所述服务类别信息，所述从所述内容信息库中确定与所述服务类别信息相关的内容信息，包括：

[19]从所述内容信息库中确定与所述服务类别信息对应的所述内容信息。

[20]在一种可能的实施方式中，所述各个内容信息的特征信息包括预先确定的与该内容信息对应的场所信息，所述确定向所述用户推荐的内容推荐信息，包括：

15 [21]当与所述各个内容信息对应的场所信息和与所述位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为所述内容推荐信息。

[22]在一种可能的实施方式中，所述特征信息基于从所述内容信息中提取出的关键词信息和场所信息而确定。

20 [23]在一种可能的实施方式中，所述位置数据包括无线保真 WiFi 指纹数据和经纬度数据中的至少一种。

[24]在一种可能的实施方式中，所述场所信息包括建筑名称、商户名称和商圈信息中的至少一种。

[25]第二方面，提供了一种内容推荐装置。该装置包括：

[26]获取单元，用于获取用户的位置数据；

25 [27]确定单元，用于根据位置映射模型，确定与所述位置数据对应的场所信息，所述位置映射模型基于预先获取的服务信息而训练获得，所述服务信息中包含服务位置数据和对应的服务场所数据；

[28]处理单元，用于至少根据所述场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息。

[29]在一种可能的实施方式中，所述位置映射模型由所述确定单元通过以下步骤训练获得：

5 [30]对所述服务信息中包含的服务位置数据进行聚类，获得多个类簇；

[31]根据所述服务信息中与所述服务位置数据对应的服务场所数据，确定所述多个类簇对应的场所标签。

[32]在一种可能的实施方式中，所述确定单元具体包括：

[33]第一确定子单元，用于从所述多个类簇中确定与所述用户的位置数据对应的类簇；

10 [34]第二确定子单元，用于根据确定出的类簇所对应的场所标签，确定所述场所信息。

[35]在一种可能的实施方式中，所述处理单元具体包括：

[36]第一处理子单元，用于确定与所述场所信息对应的服务类别信息；

[37]第二处理子单元，用于根据所述各个内容信息的特征信息，从所述内容信息库中确定与所述服务类别信息相关的内容信息，并将所述内容信息作为所述内容推荐信息。

15 [38]在一种可能的实施方式中，所述第二处理子单元中的特征信息包括所述内容信息的所述服务类别信息，所述第二处理子单元具体用于：

[39]从所述内容信息库中确定与所述服务类别信息对应的所述内容信息。

[40]在一种可能的实施方式中，所述处理单元中包括的各个内容信息的特征信息包括预先确定的与该内容信息对应的场所信息，所述处理单元具体用于：

20 [41]当与所述各个内容信息对应的场所信息和与所述位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为所述内容推荐信息。

[42]在一种可能的实施方式中，所述处理单元中包括的特征信息基于从所述内容信息中提取出的关键词信息和场所信息而确定。

[43]在一种可能的实施方式中，所述获取单元获取的位置数据包括无线保真 WiFi 指纹
25 数据和经纬度数据中的至少一种。

[44]在一种可能的实施方式中，所述确定单元确定的场所信息包括建筑名称、商户名称和商圈信息中的至少一种。

[45]第三方面，提供了一种计算机可读存储介质，其上存储有计算机程序。当所述计算机程序在计算机中执行时，令计算机执行上述第一方面中任一种实施方式提供的方法。

[46]第四方面，提供了一种计算设备，包括存储器和处理器。所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现上述第一方面中任一种实施方式提供的方法。

[47]本说明书提供的一种内容推荐方法及装置，通过获取用户的位置数据，并根据位置映射模型，确定与位置数据对应的场所信息。然后，至少根据该场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息，从而及时地向用户推荐更加精准的内容信息。

10 附图说明

[48]为了更清楚地说明本说明书披露的多个实施例的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书披露的多个实施例，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它的附图。

15 [49]图 1 为本说明书披露的一个实施例提供的一种内容推荐方法的应用场景示意图；

[50]图 2 为本说明书披露的一个实施例提供的一种内容推荐方法的流程图；

[51]图 3 为本说明书披露的一个实施例提供的一种将服务位置数据聚类成多个类簇的示意图；

[52]图 4 为本说明书披露的一个实施例提供的一种特征信息确定方法的流程图；

20 [53]图 5 为本说明书披露的一个实施例提供的一种内容推荐装置的示意图。

具体实施方式

[54]下面结合附图，对本说明书披露的多个实施例进行描述。

[55]图 1 为本说明书披露的一个实施例提供的一种内容推荐方法的应用场景示意图。所述推荐方法的执行主体可以为服务器。当用户通过终端（如，终端可以为手机、平板电脑、可穿戴智能设备等）登录内容推荐平台（如，支付宝应用平台）时，可以采用本说明书披露的多个实施例提供的业务功能的推荐方法，服务器获取用户的位置数据（如，

经纬度数据)，并根据位置映射模型（如，位置映射模型可以包括位置数据与场所信息的映射关系），确定与该位置数据对应的场所信息（如，汽车 4S 店）。然后，根据该场所信息，以及内容信息库中各个内容信息的特征信息（如，特征信息可以包括服务类别信息），确定向用户推荐的内容推荐信息（如，汽车保养知识等）。

- 5 [56]本说明书披露的多个实施例提供的业务功能的推荐方法,通过获取用户的位置数据,并根据位置映射模型,确定与位置数据对应的场所信息。然后,至少根据该场所信息,以及内容信息库中各个内容信息的特征信息,确定向所述用户推荐的内容推荐信息,从而及时地向用户推荐更加精准的内容信息。

- 10 [57]图 2 为本说明书披露的一个实施例提供的一种内容推荐方法的流程图。所述方法的执行主体可以为具有处理能力的设备: 服务器或者系统或者装置, 例如, 如图 1 所示的服务器。如图 2 所示, 所述方法具体包括:

[58]步骤 S210, 获取用户的位置数据。

- 15 [59]具体地, 位置数据可以为通过基于位置的服务 (Location Based Service, 简称 LBS) 从用户的终端采集的数据。LBS 中包括多种定位方式, 如全球定位系统 (Global Positioning System, 简称 GPS) 定位、基站定位、无线保真 (Wireless Fidelity, 简称 WiFi) 定位等。

[60]在一个实施例中, 位置数据可以包括通过 GPS 定位或基站定位得到的经纬度数据。例如, 位置数据中包括的经纬度数据为: 北纬 39° 54' 25.70" 和东经 116° 23' 28.49"。

- 20 [61]在另一个实施例中, 位置数据可以包括通过 WiFi 定位得到的 WiFi 指纹数据。在一个例子中, WiFi 指纹数据可以包括 WiFi 接入点 (Access Point, 简称 AP) 的地址 (如, 媒体访问控制 (Media Access Control, 简称 MAC) 地址) 和相应的信号强度等。

[62]步骤 S220, 根据位置映射模型, 确定与用户的位置数据对应的场所信息。

- 25 [63]具体地, 场所信息与位置数据的不同之处在于, 场所信息是用户可以感知到的、具有明确语义的位置信息。在一个实施例中, 场所信息可以包括建筑名称 (如, 新中关大厦)、商户名称 (如, 哈姆玩具店) 和商圈信息 (如, 三里屯太古里) 等。在另一个实施例中, 场所信息还可以包括与场所 (如, 商户) 相关的服务类别信息 (如, 美食、服装、美容、电影等)。

[64]从位置数据到场所信息的转换可以根据预先获得的位置映射模型来执行。在一个实施例中, 上述位置映射模型包括 GPS 坐标数据与场所标签之间的映射关系, 该映射关

系预先通过人工采集而获得。

[65]在另一实施例中，上述位置映射模型可以基于预先获取的服务信息中包含的服务位置数据和对应的服务场所数据而训练获得。与服务信息对应的服务可以包括支付服务（如，使用支付应用对订单进行支付）和定位服务（如，在社交平台发送状态信息时

5 当前位置进行定位）等。

[66]其中，服务信息可以通过多种方式获取。在一个实施例中，与服务信息对应的服务可以包括支付服务，相应的服务信息可以包括支付信息。获取支付信息，可以包括：服务器在检测到用户通过终端提交订单或对订单进行支付时，可以采集支付信息，具体包括通过采集终端的位置数据作为支付信息中的服务位置数据，以及从提交的订单中获取

10 与服务位置数据相应的服务场所信息。在一个例子中，用户使用已连接 WiFi 的终端上的支付应用（如，支付宝）进行支付操作时，该支付应用的服务器可以通过 WiFi 定位采集此时终端的位置数据，（如，WiFi 指纹数据）和支付信息中包括的服务场所数据（如，麦当劳学院路店）。

[67]在另一个实施例中，与服务信息对应的服务可以包括定位服务，相应的服务信息可以包括定位服务信息。获取定位服务信息，可以包括：服务器在检测到用户通过终端使用定位服务时，可以获取定位服务信息，具体包括采集终端的位置数据作为定位服务信息中的服务位置数据，以及获取用户在使用定位服务时选定或创建的位置信息。在一个例子中，用户使用支持 GPS 定位的终端登录工作应用（如，钉钉）进行上班打卡操作，此时，可以通过 GPS 定位采集终端的位置数据（如，经纬度数据）和打卡信息中包括

15 的服务场所数据（如，加州美食餐厅）。在另一个例子中，用户使用连接有蜂窝网的终端在社交平台上发布包括场所信息的状态，此时，可以通过基站定位采集终端的位置数据（如，经纬度数据）和用户选定（或创建）的服务场所数据（如，一二咖啡馆）。

[68]进一步地，基于预先获取的服务信息，位置映射模型可以通过以下步骤训练获得：对服务信息中包含的服务位置数据进行聚类，获得多个类簇。然后，根据服务信息中与

20 服务位置数据对应的服务场所数据，确定多个类簇对应的场所标签。如此，可以建立多个类簇与多个场所标签之间的映射关系。

[69]在一个实施例中，可以采用聚类算法对服务位置数据进行聚类，并获得多个类簇。其中聚类算法可以为 GEOHASH 算法或 DBSCAN 算法，在此不作限定。

[70]在一个例子中，服务位置数据包括多个经纬度数据，此时可以采用 GEOHASH 算法

将该数据转换成 GEOHASH 网格，并确定各个经纬度数据的网格编号。例如，服务位置数据包括 100 个经纬度坐标，采用 GEOHASH 算法将这 100 经纬度坐标转换成 GEOHASH 网格后，如图 3 所示，每 20 个经纬度坐标具有相同的网格编号，得到的 5 个网格编号分别为：WX4G01、WX4H02、WX4I03、WX4J04、WX4K05。如此，以上

5 100 个经纬度坐标被聚类为 5 个类簇。

[71]在另一个例子中，服务位置数据包括多个 WiFi 指纹数据，此时可以采用 DBSCAN 算法对该数据进行聚类，获取多个具有不同编号的类簇。DBSCAN 算法是一种基于密度的聚类算法。与划分和层次聚类方法不同，它将簇定义为密度相连的点的最大集合，能够把具有足够高密度的区域划分为簇，并可在噪声的空间数据库中发现任意形状的聚

10 类。具体而言，在 DBSCAN 算法中，首先将所有位置点标记为核心点、边界点或噪声点，删除其中的噪声点。然后为距离在预设参数之内的所有核心点之间赋予一条边，每组连通的核心点形成一个簇，将每个边界点指派到一个与之关联的核心点的簇中，由此完成位置点的聚类。

[72]可以理解的是，可以通过调整聚类算法的参数，控制聚类得到的类簇的精度（如，

15 类簇的范围大小）。例如，可以通过控制 GEOHASH 算法中对服务位置数据进行编码的编码位数，控制类簇的精度。更具体地，服务位置数据的编码位数越多，得到的类簇范围越精确。又例如，可以通过控制 DBSCAN 算法中邻域半径 ϵ 的大小，控制类簇的精度。更具体地，输入的领域半径 ϵ 的值越小，得到的类簇范围越精确。

[73]在一个实施例，对聚类获得的多个类簇，可以根据服务信息中与服务位置数据对

20 应的服务场所数据，确定多个类簇中各个类簇所对应的场所标签。

[74]在一个例子中，与服务位置数据对应的服务场所数据中包括服务场所的名称（如，幸福超市），则可以将此服务场所的名称作为该服务位置数据所在类簇对应的场所标签。

[75]在另一个例子中，单个类簇中可以对应多个服务位置数据，而与多个服务位置数据对应的多个服务场所数据所标识的范围（如，商户或商圈）可能不同。相应地，单个类

25 簇可以具有多个场所标签，如，商户标签和商圈标签等。例如，某个类簇的场所标签可以包括商户标签（如，西部马华）和商圈标签（如，大钟寺）等。

[76]此步骤中，确定与用户的位置数据对应的场所信息，可以包括：从聚类获得的多个类簇中确定与用户的位置数据对应的类簇。接着，根据确定出的类簇所对应的场所标签，确定场所信息。

[77]在一个实施例中，位置映射模型中包括多个类簇中各个类簇的中心点的位置数据，以及多个类簇与多个场所标签之间的映射关系。如此，从聚类获得的多个类簇中确定与用户的位置数据对应的类簇，可以包括：计算用户的位置数据与各中心点的位置数据之间距离，将计算出的距离的最小值所对应的类簇作为与用户的位置数据对应的类簇。

- 5 [78]进一步地，根据确定出的类簇所对应的场所标签，确定场所信息，可以包括：根据多个类簇与多个场所标签之间的映射关系，以及确定出的与用户的位置数据对应类簇，确定与该类簇对应的场所标签，进而确定场所信息。

[79]在一个实施例中，确定出的场所标签包括商户标签（如，西部马华），可以据此确定出包括商户名称（如，西部马华）的场所信息。

- 10 [80]在步骤 S220 中确定出与用户的位置数据对应的场所信息后，接着，在步骤 S230，至少根据场所信息，以及内容信息库中各个内容信息的特征信息，确定向用户推荐的内容推荐信息。

[81]具体地，确定向用户推荐的内容推荐信息，可以包括：确定与场所信息对应的服务类别信息，以及根据各个内容信息的特征信息，从内容信息库中确定与服务类别信息相关的
15 内容信息，并将该内容信息作为内容推荐信息。

[82]在一个实施例中，从内容信息库中确定与服务类别信息相关的内容信息，可以包括：根据各个内容信息的特征信息，计算各个内容信息与服务类别信息的相关度，并根据计算出的相关度确定与服务类别信息相关的内容信息。

- [83]在一个例子中，可以将相关度在预定范围内（如，预定范围可以为大于 0.6）的内容
20 信息确定为与服务类别信息相关的内容信息。在另一个例子中，可以根据相关度对内容信息进行排名，并将名次在预定范围（如，预定范围可以为前五名）内的内容信息确定为与服务类别信息相关的内容信息。

- [84]在另一个实施例中，特征信息可以包括内容信息的服务类别信息。从内容信息库中确定与服务类别信息相关的内容信息，可以包括：从内容信息库中确定与服务类别信息
25 对应的内容信息。

[85]在一个例子中，步骤 S220 中确定的场所信息为“汽车 4S 店”，相应地，可以确定出与该场所信息对应的服务类别信息为“汽车”。相应地，可以从内容信息库中确定服务类别信息为“汽车”的内容信息，并将此内容信息作为向用户推荐的内容推荐信息。

[86]或者，各个内容信息的特征信息中可以包括预先确定的与该内容信息对应的场所信

息。在这样的情况下，确定向用户推荐的内容推荐信息，可以包括：当与各个内容信息对应的场所信息和与位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为内容推荐信息。

5 [87]在一个实施例中，预设规则可以由业务方根据业务的属性（如，该业务对场所范围的精度要求）确定。例如，当业务为向用户推送商家优惠券时，预设规则可以包括要求与内容推荐信息对应的场所信息与位置数据对应的场所信息完全相同。当业务为向用户推荐附近的类似的场所时，预设规则可以包括要求与内容推荐信息对应的场所信息与位置数据对应的场所信息部分相同。

10 [88]在一个例子中，预设规则包括要求与内容推荐信息对应的场所信息与位置数据对应的场所信息完全相同，与位置数据对应的场所信息包括“老好吃香锅”、“三里屯”。相应地，可以将与场所信息“老好吃香锅”（商户）、“三里屯”（商圈）对应的内容信息（如，商户优惠活动、商户公众号信息、商户评价信息）作为内容推荐信息。

15 [89]在另一个例子中，预设规则可以包括要求与内容推荐信息对应的场所信息与位置数据对应的场所信息部分相同，与位置数据对应的场所信息包括“优衣库”（商户）、“三里屯”（商圈）。相应地，可以将与场所信息“三里屯”对应的内容信息（如，三里屯商圈内的服装店、餐厅等）作为内容推荐信息。

20 [90]在通过上述步骤确定向用户推荐的内容推荐信息后，还可以包括：向用户发送内容推荐信息。例如，当用户打开相应的应用 App，或打开推荐频道时，或当检测到用户的位置发生变化时，获取用户最新的位置信息，据此确定向用户推荐的内容推荐信息，并向用户发送该内容推荐信息。

[91]需要说明的是，在步骤 S220 中，还可以包括：确定与场所信息对应的服务类别信息。在一个实施例中，场所信息中除了包括场所的名称外，还可以包括与场所对应的服务类别信息，相应地，可以据此直接确定出与场所信息对应的服务类别信息。

25 [92]在另一个实施例中，可以在确定场所信息后，根据预存的多个场所信息和多个服务类别信息的对应关系，确定该场所信息对应的服务类别信息。在一个例子中，确定出的场所标签包括商户标签（如，西部马华），可以据此确定出包括商户名称（如，西部马华）、以及与商户名称对应的服务类别信息（如，美食）。

[93]在步骤 S230 中提及的特征信息可以基于从内容信息中提取的关键词信息和/或位置信息而确定，具体可以采用如图 4 所示的方法进行确定：

[94] 步骤 S410, 对内容信息进行预处理。

[95] 具体地, 预处理可以包括结构化分析、分词处理、去停用词处理、词性标注 (postag) 等。其中, 结构化分析可以包括对内容信息中段落结构的分析, 例如, 判断出内容信息中的标题和正文; 分词处理可以包括一元分词 (unigram)、二元分词 (bigram)、三元分词 (trigram) 等; 去停用词可以包括根据预设的停用词表去除内容信息中的停用词 (如, 无实际意义的功能词: 这、那、的); 词性标注是指对内容信息中的词语的词性 (如, 名词、副词、形容词等) 进行标注。

[96] 步骤 S420, 根据预处理后的内容信息, 提取关键词信息。

[97] 具体地, 可以根据预处理得到的词语在内容信息中的位置 (如, 位于标题中或位于正文中)、标注的词性, 采用 TextRank 算法或 TF-IDF (Term Frequency - inverse Document Frequency) 算法加权识别出关键词信息。

[98] 步骤 S430, 根据预处理后的内容信息, 提取位置信息。

[99] 具体地, 可以根据内容信息中包括的位置标签提取位置信息, 位置信息中可以包括场所信息。在一个实施例中, 该位置标签可以为内容信息的生产者在发布该内容信息时, 为其所贴的位置标签, 如, 北京、杭州等。

[100] 或者, 可以采用命名实体识别 (Named Entity Recognition, 简称 NER) 的方法, 识别内容信息中的地名、机构名等。例如, 可以识别出内容信息中的地名为五台山, 机构名为海淀区民政局 (属于场所信息)。

[101] 又或者, 服务器中可以预先存储有位置信息库, 可以据此从预处理后的内容信息中提取出与所述位置信息库中的信息匹配的位置信息。

[102] 步骤 S440, 将初步提取的关键词信息和位置信息输入预先训练的特征提取模型中, 确定出特征信息。

[103] 具体地, 特征提取模型可以为对大规模内容语料数据进行离线训练, 而产生的词嵌入 (Word Embedding) 模型或基于双向循环神经网络 (Recurrent Neural Networks, 简称 RNN) 的 NER 模型。

[104] 此外, 上述特征信息的确定可以在内容信息产生以后的预定时间 (如, 5min 或 10min) 内进行。

[105] 由上可知, 在本说明书披露的多个实施例提供的内容推荐方法中, 通过获取用

户的位置数据,并根据位置映射模型,确定与位置数据对应的场所信息。然后,至少根据该场所信息,以及内容信息库中各个内容信息的特征信息,确定向所述用户推荐的内容推荐信息,从而及时地向用户推荐更加精准的内容信息。

[106] 与内容推荐方法对应地,本说明书披露的多个实施例还提供一种内容推荐装置,如图 5 所示,该装置 500 包括:

[107] 获取单元 510,用于获取用户的位置数据;

[108] 确定单元 520,用于根据位置映射模型,确定与位置数据对应的场所信息,该位置映射模型基于预先获取的服务信息而训练获得,该服务信息中包含服务位置数据和对应的服务场所数据;

10 [109] 处理单元 530,用于至少根据场所信息,以及内容信息库中各个内容信息的特征信息,确定向用户推荐的内容推荐信息。

[110] 在一种可能的实施方式中,位置映射模型由确定单元 520 通过以下步骤训练获得:

[111] 对服务信息中包含的服务位置数据进行聚类,获得多个类簇;

15 [112] 根据服务信息中与服务位置数据对应的服务场所数据,确定多个类簇对应的场所标签。

[113] 在一种可能的实施方式中,确定单元 520 具体包括:

[114] 第一确定子单元 521,用于从多个类簇中确定与用户的位置数据对应的类簇;

[115] 第二确定子单元 522,用于根据确定出的类簇所对应的场所标签,确定场所信息。

20 [116] 在一种可能的实施方式中,处理单元 530 具体包括:

[117] 第一处理子单元 531,用于确定与场所信息对应的服务类别信息;

[118] 第二处理子单元 532,用于根据各个内容信息的特征信息,从内容信息库中确定与服务类别信息相关的内容信息,并将内容信息作为内容推荐信息。

25 [119] 在一种可能的实施方式中,第二处理子单元 532 中的特征信息包括内容信息的服务类别信息,第二处理子单元 532 具体用于:

[120] 从内容信息库中确定与服务类别信息对应的内容信息。

[121] 在一种可能的实施方式中,处理单元 530 中包括的各个内容信息的特征信息包

括预先确定的与该内容信息对应的场所信息，处理单元 530 具体用于：

[122] 当与各个内容信息对应的场所信息和与位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为内容推荐信息。

5 [123] 在一种可能的实施方式中，处理单元 530 中包括的特征信息基于从内容信息中提取出的关键词信息和场所信息而确定。

[124] 在一种可能的实施方式中，其特征信息在于，获取单元 510 获取的位置数据包括无线保真 WiFi 指纹数据和经纬度数据中的至少一种。

[125] 在一种可能的实施方式中，确定单元 520 确定的场所信息包括建筑名称、商户名称和商圈信息中的至少一种。

10 [126] 由上可知，在本说明书披露的多个实施例提供的内容推荐装置中，获取单元 510 获取用户的位置数据，确定单元 520 根据位置映射模型，确定与位置数据对应的场所信息。处理单元 530 至少根据该场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息，从而及时地向用户推荐更加精准的内容信息。

15 [127] 本领域技术人员应该可以意识到，在上述一个或多个示例中，本说明书披露的多个实施例所描述的功能可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。

20 [128] 以上所述的具体实施方式，对本说明书披露的多个实施例的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本说明书披露的多个实施例的具体实施方式而已，并不用于限定本说明书披露的多个实施例的保护范围，凡在本说明书披露的多个实施例的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本说明书披露的多个实施例的保护范围之内。

权利要求书

1、一种内容推荐方法，其特征在于，包括：

获取用户的位置数据；

5 根据位置映射模型,确定与所述位置数据对应的场所信息,所述位置映射模型基于预先获取的服务信息而训练获得,所述服务信息中包含服务位置数据和对应的服务场所数据；

至少根据所述场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息。

10 2、根据权利要求 1 所述的方法，其特征在于，所述位置映射模型通过以下步骤训练获得：

对所述服务信息中包含的服务位置数据进行聚类，获得多个类簇；

根据所述服务信息中与所述服务位置数据对应的服务场所数据，确定所述多个类簇对应的场所标签。

15 3、根据权利要求 2 所述的方法，其特征在于，所述确定与所述位置数据对应的场所信息包括：

从所述多个类簇中确定与所述用户的位置数据对应的类簇；

根据确定出的类簇所对应的场所标签，确定所述场所信息。

4、根据权利要求 1 所述的方法，其特征在于，所述确定向所述用户推荐的内容推荐信息，包括：

20 确定与所述场所信息对应的服务类别信息；

根据所述各个内容信息的特征信息，从所述内容信息库中确定与所述服务类别信息相关的内容信息，并将所述内容信息作为所述内容推荐信息。

25 5、根据权利要求 4 所述的方法，其特征在于，所述特征信息包括所述内容信息的所述服务类别信息，所述从所述内容信息库中确定与所述服务类别信息相关的内容信息，包括：

从所述内容信息库中确定与所述服务类别信息对应的所述内容信息。

6、根据权利要求 1 所述的方法，其特征在于，所述各个内容信息的特征信息包括预先确定的与该内容信息对应的场所信息，所述确定向所述用户推荐的内容推荐信息，包括：

30 当与所述各个内容信息对应的场所信息和与所述位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为所述内容推荐信息。

7、根据权利要求 1-6 任一项所述的方法，其特征在于，所述特征信息基于从所述内容信息中提取出的关键词信息和场所信息而确定。

8、根据权利要求 1-6 中任一项所述的方法，其特征在于，所述位置数据包括无线保真 WiFi 指纹数据和经纬度数据中的至少一种。

5 9、根据权利要求 1-6 任一项所述的方法，其特征在于，所述场所信息包括建筑名称、商户名称和商圈信息中的至少一种。

10、一种内容推荐装置，其特征在于，包括：

获取单元，用于获取用户的位置数据；

10 确定单元，用于根据位置映射模型，确定与所述位置数据对应的场所信息，所述位置映射模型基于预先获取的服务信息而训练获得，所述服务信息中包含服务位置数据和对应的服务场所数据；

处理单元，用于至少根据所述场所信息，以及内容信息库中各个内容信息的特征信息，确定向所述用户推荐的内容推荐信息。

15 11、根据权利要求 10 所述的装置，其特征在于，所述位置映射模型由所述确定单元通过以下步骤训练获得：

对所述服务信息中包含的服务位置数据进行聚类，获得多个类簇；

根据所述服务信息中与所述服务位置数据对应的服务场所数据，确定所述多个类簇对应的场所标签。

12、根据权利要求 11 所述的装置，其特征在于，所述确定单元具体包括：

20 第一确定子单元，用于从所述多个类簇中确定与所述用户的位置数据对应的类簇；

第二确定子单元，用于根据确定出的类簇所对应的场所标签，确定所述场所信息。

13、根据权利要求 10 所述的装置，其特征在于，所述处理单元具体包括：

第一处理子单元，用于确定与所述场所信息对应的服务类别信息；

25 第二处理子单元，用于根据所述各个内容信息的特征信息，从所述内容信息库中确定与所述服务类别信息相关的内容信息，并将所述内容信息作为所述内容推荐信息。

14、根据权利要求 13 所述的装置，其特征在于，所述第二处理子单元中的特征信息包括所述内容信息的所述服务类别信息，所述第二处理子单元具体用于：

从所述内容信息库中确定与所述服务类别信息对应的所述内容信息。

30 15、根据权利要求 10 所述的装置，其特征在于，所述处理单元中包括的各个内容信息的特征信息包括预先确定的与该内容信息对应的场所信息，所述处理单元具体用于：

当与所述各个内容信息对应的场所信息和与所述位置数据对应的场所信息之间的从属关系符合预设规则时，将该内容信息作为所述内容推荐信息。

16、根据权利要求 10-15 任一项所述的装置，其特征在于，所述处理单元中包括的特征信息基于从所述内容信息中提取出的关键词信息和场所信息而确定。

5 17、根据权利要求 10-15 中任一项所述的装置，其特征在于，所述获取单元获取的位置数据包括无线保真 WiFi 指纹数据和经纬度数据中的至少一种。

18、根据权利要求 10-15 任一项所述的装置，其特征在于，所述确定单元确定的场所信息包括建筑名称、商户名称和商圈信息中的至少一种。

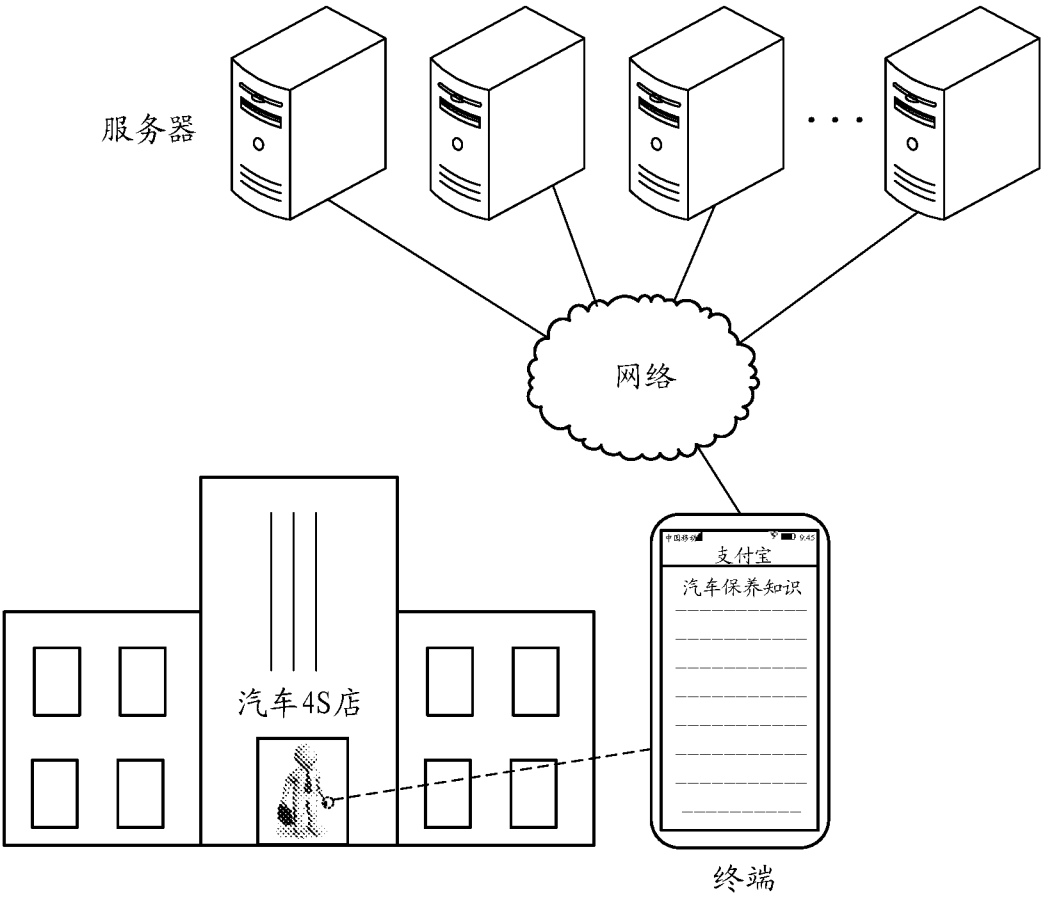


图 1

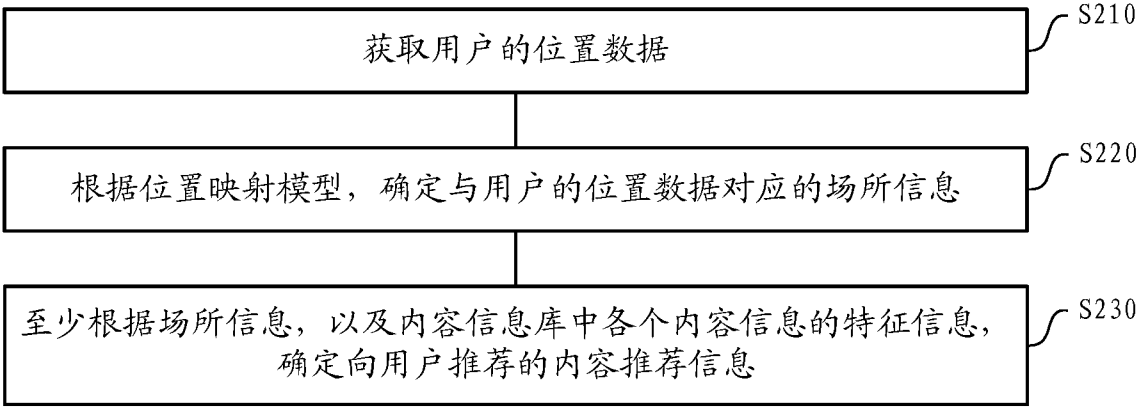


图 2



图 3

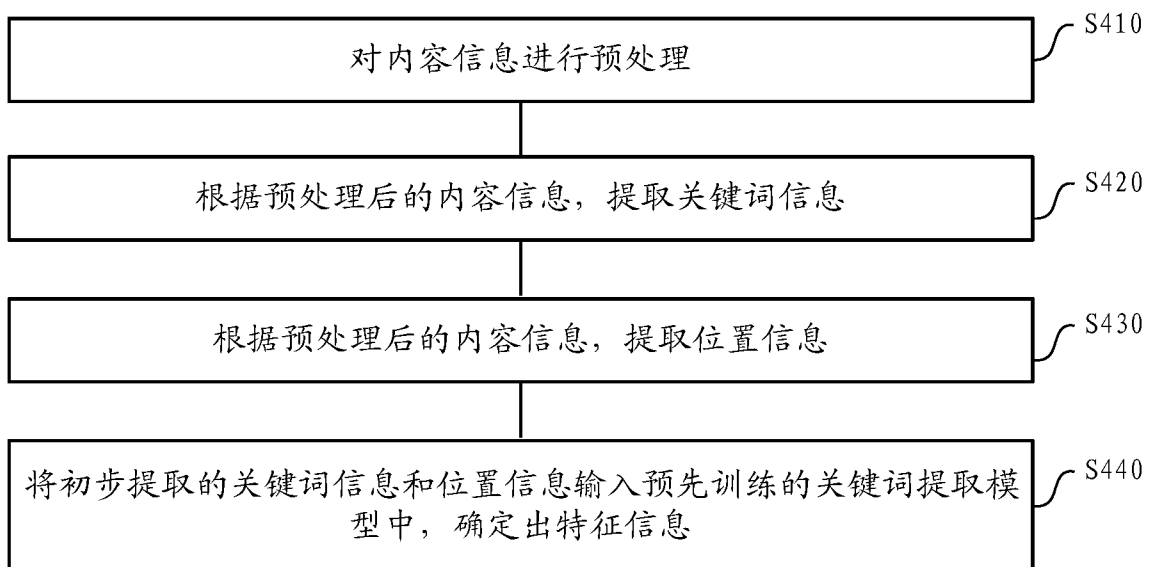


图 4

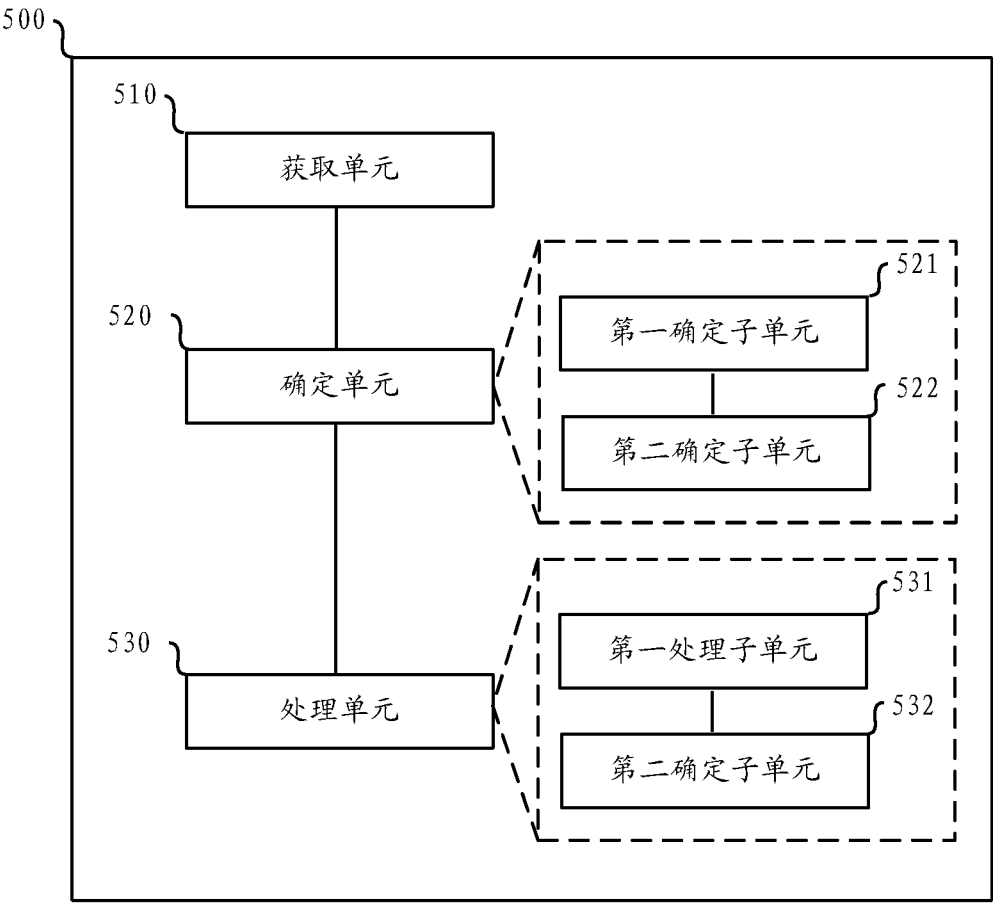


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/070830

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/00(2019.01)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)

CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 内容, 推荐, 训练, 位置, 模型; content, recommending, train, location, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 108363733 A (ALIBABA GROUP HOLDING LIMITED) 03 August 2018 (2018-08-03) entire document	1-18
X	CN 106294489 A (BEIJING SAMSUNG TELECOMMUNICATION R&D CENTER ET AL.) 04 January 2017 (2017-01-04) description, paragraphs [0031]-[0033], [0046]-[0059], [0065], [0066] and [0078]-[0080], and figure 1	1-18
A	CN 103023977 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 03 April 2013 (2013-04-03) entire document	1-18
A	US 2009094189 A1 (AT & T BLS INTELLECTUAL PROPERTY, INC.) 09 April 2009 (2009-04-09) entire document	1-18

☐ 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

25 March 2019

Date of mailing of the international search report

11 April 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/070830

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108363733	A	03 August 2018	None			
CN	106294489	A	04 January 2017	KR	20160144305	A	16 December 2016
				WO	2016200150	A1	15 December 2016
				US	2016358225	A1	08 December 2016
CN	103023977	A	03 April 2013	CN	103023977	B	01 July 2015
US	2009094189	A1	09 April 2009	None			

国际检索报告

国际申请号

PCT/CN2019/070830

A. 主题的分类

G06F 16/00 (2019.01) 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))

CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 内容, 推荐, 训练, 位置, 模型; content, recommending, train, location, model

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 108363733 A (阿里巴巴集团控股有限公司) 2018年 8月 3日 (2018 - 08 - 03) 全文	1-18
X	CN 106294489 A (北京三星通信技术研究有限公司等) 2017年 1月 4日 (2017 - 01 - 04) 说明书第31-33、46-59、65、66、78-80段, 图1	1-18
A	CN 103023977 A (华南理工大学) 2013年 4月 3日 (2013 - 04 - 03) 全文	1-18
A	US 2009094189 A1 (AT & T BLS INTELLECTUAL PROPER, INC.) 2009年 4月 9日 (2009 - 04 - 09) 全文	1-18

☐ 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

2019年 3月 25日

国际检索报告邮寄日期

2019年 4月 11日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (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

授权官员

古志春

电话号码 86-(20)-28950386

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/070830

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	108363733	A	2018年 8月 3日	无			
CN	106294489	A	2017年 1月 4日	KR	20160144305	A	2016年 12月 16日
				WO	2016200150	A1	2016年 12月 15日
				US	2016358225	A1	2016年 12月 8日
CN	103023977	A	2013年 4月 3日	CN	103023977	B	2015年 7月 1日
US	2009094189	A1	2009年 4月 9日	无			

表 PCT/ISA/210 (同族专利附件) (2015年1月)