

(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
2019 年 10 月 31 日 (31.10.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/205373 A1

(51) 国际专利分类号:
G06F 16/9536 (2019.01)

(21) 国际申请号: PCT/CN2018/102112

(22) 国际申请日: 2018 年 8 月 24 日 (24.08.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810382302.8 2018 年 4 月 26 日 (26.04.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 王健宗(WANG, JianZong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000

(CN)。 吴天博(WU, TianBo); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 黄章成(HUANG, ZhangCheng); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, Guangdong 518000 (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,

(54) Title: SIMILAR USER SEARCH APPARATUS AND METHOD, AND COMPUTER READABLE STORAGE MEDIUM

(54) 发明名称: 相似用户查找装置、方法及计算机可读存储介质

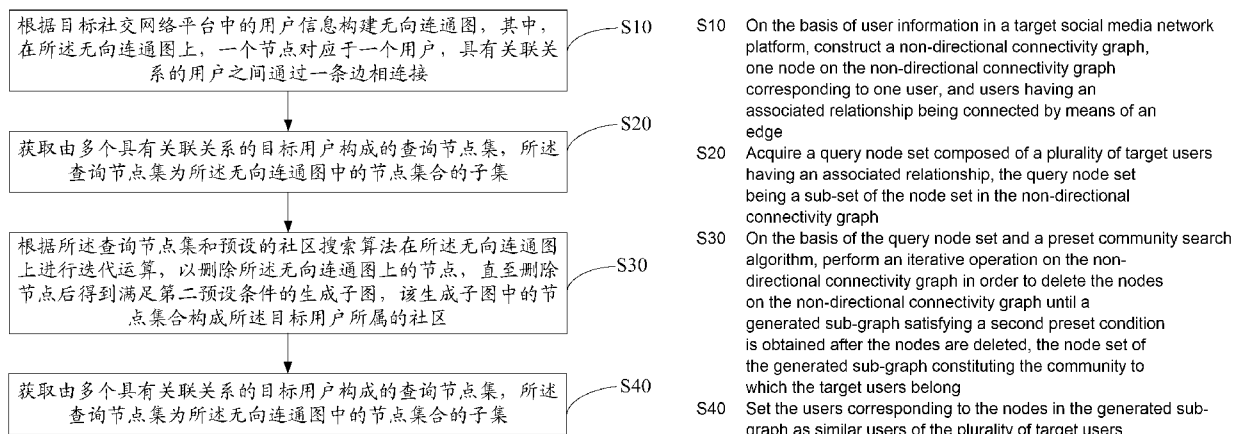


图 3

(57) Abstract: A similar user search apparatus based on community search, comprising a memory and a processor, a user search program capable of running on the processor being stored on the memory, and the program implementing the following steps when executed by the processor: on the basis of user information in a target social media network platform, constructing a non-directional connectivity graph; acquiring a query node set composed of a plurality of target users having an associated relationship, the query node set being a sub-set of the node set in the non-directional connectivity graph; on the basis of the query node set and a preset community search algorithm, performing an iterative operation on the non-directional connectivity graph in order to delete the nodes on the non-directional connectivity graph until a generated sub-graph satisfying a second preset condition is obtained after the nodes are deleted; and setting the users corresponding to the nodes in the generated sub-graph as similar users of the plurality of target users. The present apparatus improves the accuracy of the search for similar users.

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))。

(57) 摘要: 一种基于社区搜索的相似用户查找装置, 包括存储器和处理器, 存储器上存储有可在处理器上运行的用户查找程序, 该程序被处理器执行时实现如下步骤: 根据目标社交网络平台中的用户信息构建无向连通图, 其中; 获取由多个具有关联关系的目标用户构成的查询节点集, 查询节点集为无向连通图中的节点集合的子集; 根据查询节点集和预设的社区搜索算法在无向连通图上进行迭代运算, 以删除无向连通图上的节点, 直至删除节点后得到满足第二预设条件的生成子图; 将生成子图中的节点对应的用户作为多个目标用户的相似用户。所述装置提高了查找相似用户的准确度。

相似用户查找装置、方法及计算机可读存储介质

本申请基于巴黎公约申明享有2018年04月26日递交的申请号为2018103823028、名称为“相似用户查找装置、方法及计算机可读存储介质”的中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

技术领域

本申请涉及计算机技术领域，尤其涉及一种相似用户查找装置、方法及计算机可读存储介质。

背景技术

随着信息技术的发展，互联网、移动互联网、物联网能够收集到越来越多的用户信息，如何将这些信息采集、存储并分析，进而发现不同用户之间的相似性正在成为研究热点。

目前，现有的用户相似度的计算方法主要依赖于用户资料。这种方式存在以下缺陷：一方面，用户资料很难反映一个人的全部特征，而另一方面，如果用户资料填写不完整，那么相似度的计算是不理想的，因此，通过这种方法计算得到的相似用户的准确度较低。

发明内容

本申请提供一种相似用户查找装置、方法及计算机可读存储介质，其主要目的在于提高查找相似用户的准确度。

为实现上述目的，本申请提供一种基于社区搜索的相似用户查找装置，该装置包括存储器和处理器，所述存储器中存储有可在所述处理器上运行的用户查找程序，所述用户查找程序被所述处理器执行时实现如下步骤：

根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接；

获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点

集为所述无向连通图中的节点集合的子集;

根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算, 以删除所述无向连通图上的节点, 直至删除节点后得到满足第二预设条件的生成子图, 该生成子图中的节点集合构成所述目标用户所属的社区;

5 将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

此外, 为实现上述目的, 本申请还提供一种基于社区搜索的相似用户查找方法, 该方法包括:

10 根据目标社交网络平台中的用户信息构建无向连通图, 其中, 在所述无向连通图上, 一个节点对应于一个用户, 具有关联关系的用户之间通过一条边相连接;

获取由多个具有关联关系的目标用户构成的查询节点集, 所述查询节点集为所述无向连通图中的节点集合的子集;

15 根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算, 以删除所述无向连通图上的节点, 直至删除节点后得到满足第二预设条件的生成子图, 该生成子图中的节点集合构成所述目标用户所属的社区;

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

此外, 为实现上述目的, 本申请还提供一种计算机可读存储介质, 所述计算机可读存储介质上存储有用户查找程序, 所述用户查找程序可被一个或者多个处理器执行, 以实现如下步骤:

20 根据目标社交网络平台中的用户信息构建无向连通图, 其中, 在所述无向连通图上, 一个节点对应于一个用户, 具有关联关系的用户之间通过一条边相连接;

获取由多个具有关联关系的目标用户构成的查询节点集, 所述查询节点集为所述无向连通图中的节点集合的子集;

25 根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算, 以删除所述无向连通图上的节点, 直至删除节点后得到满足第二预设条件的生成子图, 该生成子图中的节点集合构成所述目标用户所属的社区;

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

30 本申请提出的相似用户查找装置、方法及计算机可读存储介质, 根据目标社交网络平台中的用户信息构建无向连通图, 其中, 在无向连通图上一个

节点对应于一个用户，具有关联关系的用户之间通过一条边相连接，获取由多个具有关联关系的目标用户构成的查询节点集，同时，这个查询节点集是五香连通图中的节点集合的子集，根据查询节点集合预设的社区搜索算法在无向连通图上进行迭代运算，以删除无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成目标用户所属的社区，而在无向连通图上，属于一个社区的用户具有相似的属性，因此，该生成子图上的节点对应的用户可以作为上述目标用户的相似用户。本申请通过社交网络平台进行社区搜索，查找目标用户所述的社区，进而获取目标用户的相似用户，提高了查找相似用户的准确度。

附图说明

图1为本申请基于社区搜索的相似用户查找装置较佳实施例的示意图；

图2为本申请基于社区搜索的相似用户查找装置一实施例中用户查找程序的程序模块示意图；

图3为本申请基于社区搜索的相似用户查找方法第一实施例的流程图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请提供一种基于社区搜索的相似用户查找装置。参照图 1 所示，为本申请基于社区搜索的相似用户查找装置较佳实施例的示意图。

在本实施例中，基于社区搜索的相似用户查找装置 1 可以是 PC (Personal Computer, 个人电脑)，也可以是智能手机、平板电脑、便携计算机等终端设备。该基于社区搜索的相似用户查找装置 1 至少包括存储器 11、处理器 12，网络接口 13，以及通信总线 14。

其中，存储器 11 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD 或 DX 存储器等）、磁性存储器、磁盘、光盘等。存储器 11 在一些实施例中可以是基于社区搜索的相似用户查找装置 1 的内部存储单元，例如该基于社区搜索的相似用户查找装置 1 的硬盘。存储器 11 在另一些实施例中也可以是基于社区搜索的相似用

户查找装置 1 的外部存储设备, 例如基于社区搜索的相似用户查找装置 1 上配备的插接式硬盘, 智能存储卡 (Smart Media Card, SMC), 安全数字 (Secure Digital, SD) 卡, 闪存卡 (Flash Card) 等。进一步地, 存储器 11 还可以既包括基于社区搜索的相似用户查找装置 1 的内部存储单元也包括外部存储设备。

5 存储器 11 不仅可以用于存储安装于基于社区搜索的相似用户查找装置 1 的应用软件及各类数据, 例如用户查找程序 01 的代码等, 还可以用于暂时地存储已经输出或者将要输出的数据。

处理器 12 在一些实施例中可以是一中央处理器 (Central Processing Unit, CPU)、控制器、微控制器、微处理器或其他数据处理芯片, 用于运行存储器
10 11 中存储的程序代码或处理数据, 例如执行用户查找程序 01 等。

网络接口 13 可选的可以包括标准的有线接口、无线接口(如 WI-FI 接口), 通常用于在该装置 1 与其他电子设备之间建立通信连接。

通信总线 14 用于实现这些组件之间的连接通信。

图 1 仅示出了具有组件 11-14 以及用户查找程序 01 的基于社区搜索的相
15 似用户查找装置 1, 但是应理解的是, 并不要求实施所有示出的组件, 可以替代的实施更多或者更少的组件。

可选地, 该装置 1 还可以包括用户接口, 用户接口可以包括显示器 (Display)、输入单元比如键盘 (Keyboard), 可选的用户接口还可以包括标准的有线接口、无线接口。可选地, 在一些实施例中, 显示器可以是 LED 显
20 示器、液晶显示器、触控式液晶显示器以及 OLED (Organic Light-Emitting Diode, 有机发光二极管) 触摸器等。其中, 显示器也可以适当的称为显示屏或显示单元, 用于显示在基于社区搜索的相似用户查找装置 1 中处理的信息以及用于显示可视化的用户界面。

在图 1 所示的装置 1 实施例中, 存储器 11 中存储有用户查找程序 01; 处
25 理器 12 执行存储器 11 中存储的用户查找程序 01 时实现如下步骤:

根据目标社交网络平台中的用户信息构建无向连通图, 其中, 在所述无向连通图上, 一个节点对应于一个用户, 具有关联关系的用户之间通过一条边相连接。

获取一个社交网络平台上预设数量用户的用户信息, 用户信息中除了包
30 含用户的个人资料之外, 还包含用户的关注者信息和被关注者信息, 这些信

息体现了用户之间的关联关系，其中，关注关系可以是用户在上述平台上的关注关系、好友关系、共同投资等。其中，关注关系可以单向关注关系或者双向关注关系。基于用户之间的关联关系构建一个复杂网络，在该复杂网络上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接。

- 5 该复杂网络可以抽象表示成一个无向连通图 $G=(V, E)$ ，其中， V 是用户节点的集合， E 是边的集合，体现用户之间的关联关系，例如边 $e=(v_1, v_2)$ 代表用户 v_1 和用户 v_2 之间的关系。本实施例中，社交网络平台包括微博平台、Twitter 平台或者金融论坛等。

10 获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点集为所述无向连通图中的节点集合的子集。

根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成所述目标用户所属的社区。

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

- 15 获取多个目标用户，这些目标用户构成的查询节点集，查询节点集是图 G 的用户节点集 V 的一个子集。接下来根据社区搜索算法在无向连通图 G 上进行迭代计算，以找到与这些目标用户关系密切或者具有相似偏好的用户群。这些用户群以及他们之间的关联关系形成无向连通图 G 的一个生成子图。

20 具体地，根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图的步骤包括：

- 删除无向连通图中具有最小度的节点，以及与该节点相连的边；判断经节点删除操作后的无向连通图是否满足第二预设条件，其中第二预设条件为：所述查询节点集中的节点在节点删除操作后的图无向连通图中具有最小度，
25 或者，经节点删除操作后，查询节点集中的节点在该无向连通图上不再连通；若是，则终止节点删除操作，将当前的无向连通图作为所述生成子图；若否，则基于当前的无向连通图继续执行所述删除无向连通图中具有最小度的节点，以及与该节点相连的边的步骤。

- 该社区搜索算法主要包括：将 G_0 作为算法的初始输入图，其中， $G_0=G$ ；
30 每一步删除该步骤中输入的图上的一个节点以及与该节点关联的边，该节点

是当前的图中具有最小度的节点，其中，在无向连通图中，节点的度等于连接该节点的边的数量。即在第 t 步中，计算当前的图 G_{t-1} 上具有最小度的节点，删除该节点以及所有与该节点连接的边，得到图 G_t ，作为下一步的输入。

如此循环往复的删除每一步骤的图中具有最小度的节点，直至执行了节点删除操作后得到的图满足以下条件之一：查询节点集中的节点在图中具有最小度，或者，删除有最小度的节点后，查询节点集中的点不再连通。这时，终止节点删除操作，获取计算终止后得到的生成子图，该生成子图是一个连通图，并且查询节点集是该生成子图中节点集的一个子集。

按照上述算法获取到的生成子图中的节点对应的用户即为要搜索的社区，这些用户具有与目标用户相同的特征，例如，具有相同的感兴趣产品或者具有相同的特征标签等。

进一步地，在其他的实施例中，为了避免计算得到的社区范围过大，可以进一步地增加预设条件：生成子图中的节点与查询节点集 Q 中的节点的最大距离不超过预设距离。在判定经节点删除操作后的无向连通图满足第二预设条件之后，终止节点删除操作之前，计算当前的无向连通图中的节点到所述查询节点集之间的距离，判断是否有节点到查询节点的距离大于预设距离；若是，则删除到查询节点的距离大于预设距离的节点，并将删除节点后的无向连通图作为所述生成子图；若否，则执行将终止节点删除操作，将当前的无向连通图作为所述生成子图的步骤。

关于节点之间的距离的计算可以采用如下方式中的任意一种：方式一，将两个节点之间的最短路径中的边的数量作为节点之间的距离，由于在一个复杂的无向连通图中，可以由多种路径实现两节点之间的连通，此处可以根据多条路径中的最短路径来计算节点之间的距离；方式二，根据两个节点之间的最短路径中的边的权重计算节点之间的距离，具体地，从目标社交网络平台上获取到的用户信息中包含有标签信息，根据两个用户拥有的相同标签的数量设置他们之间的权重，相同标签的数量越多，则权重越大，否则权重越小，将两个节点之间的最短路径中的边的权重之和作为两个节点之间的距离。其中，标签信息可以是用户在社交网络平台上可以为自己添加的标签，这些标签体现用户的感兴趣标签信息，或者，也可以将从用户在该社交网络平台上发布的内容提取的关键字作为用户的标签信息。此外，在计算一个节

点到查询节点集的距离时，先计算该节点到查询节点集中每一个节点的距离，然后将上述多个距离中的最小距离作为该节点到查询节点集的距离。

此外，可以理解的是，在上述删除节点的过程中，如果删除了某个节点之后，有节点在图上变成孤立的点，与其他节点之间没有任何连接关系时，

5 要把这些孤立的节点删除。

通过上述计算获取到目标用户所在的社区，即由生成子图中的节点对应的用户构成的社区，这些用户是目标用户的相似用户，例如，具体在一个金融论坛中，可以认为在同一个社区中的用户有相同或者相似的投资偏好、属性、标签、感兴趣事件等。在得到相似用户后，根据它们之间相同或者相似

10 的投资偏好、属性、标签、感兴趣事件等信息做精准推荐或定向投放。

以上实施例提出的相似用户查找装置，根据目标社交网络平台中的用户信息构建无向连通图，其中，在无向连通图上一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接，获取由多个具有关联关系的目标用户构成的查询节点集，同时，这个查询节点集是五香连通图中的节点集合

15 的子集，根据查询节点集合预设的社区搜索算法在无向连通图上进行迭代运算，以删除无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成目标用户所属的社区，而在无向连通图上，属于一个社区的用户具有相似的属性，因此，该生成子图上的节点对应的用户可以作为上述目标用户的相似用户。本申请通过社交网络平台进

20 行社区搜索，查找目标用户所述的社区，进而获取目标用户的相似用户，提高了查找相似用户的准确度。

可选地，在其他的实施例中，用户查找程序还可以被分割为一个或者多个模块，一个或者多个模块被存储于存储器 11 中，并由一个或多个处理器（本实施例为处理器 12）所执行以完成本申请，本申请所称的模块是指能够完成

25 特定功能的一系列计算机程序指令段，用于描述用户查找程序在基于社区搜索的相似用户查找装置中的执行过程。

例如，参照图 2 所示，为本申请基于社区搜索的相似用户查找装置一实施例中的用户查找程序的程序模块示意图，该实施例中，用户查找程序可以被分割为信息生成模块 10、信息获取模块 20、社区搜索模块 30、用户获取模

30 块 40，示例性地：

信息生成模块 10 用于：根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接；

5 信息获取模块 20 用于：获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点集为所述无向连通图中的节点集合的子集；

社区搜索模块 30 用于：根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成所述目标用户所属的社区；

10 用户获取模块 40 用于：将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

上述信息生成模块 10、信息获取模块 20、社区搜索模块 30、用户获取模块 40 等程序模块被执行时所实现的功能或操作步骤与上述实施例大体相同，在此不再赘述。

15 此外，本申请还提供一种基于社区搜索的相似用户查找方法。参照图 3 所示，为本申请基于社区搜索的相似用户查找方法第一实施例的流程图。该方法可以由一个装置执行，该装置可以由软件和/或硬件实现。

在本实施例中，基于社区搜索的相似用户查找方法包括：

20 步骤 S10，根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接。

25 获取一个社交网络上预设数量用户的用户信息，用户信息中除了包含用户的个人资料之外，还包含用户的关注者信息和被关注者信息，这些信息体现了用户之间的关联关系，其中，关注关系可以是用户在上述平台上的关注关系、好友关系、共同投资等。其中，关注关系可以单向关注关系或者双向关注关系。基于用户之间的关联关系构建一个复杂网络，在该复杂网络上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接。该复杂网络可以抽象表示成一个无向连通图 $G=(V, E)$ ，其中， V 是用户节点的集合， E 是边的集合，体现用户之间的关联关系，例如边 $e=(v_1, v_2)$ 代表用户 v_1 和用户 v_2 之间的关系。本实施例中，社交网络平台包括微博平台、Twitter

30

平台或者金融论坛等。

步骤 S20, 获取由多个具有关联关系的目标用户构成的查询节点集, 所述查询节点集为所述无向连通图中的节点集合的子集。

5 步骤 S30, 根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算, 以删除所述无向连通图上的节点, 直至删除节点后得到满足第二预设条件的生成子图, 该生成子图中的节点集合构成所述目标用户所属的社区。

步骤 S40, 将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

10 获取多个目标用户, 这些目标用户构成的查询节点集, 查询节点集是图 G 的用户节点集 V 的一个子集。接下来根据社区搜索算法在无向连通图 G 上进行迭代计算, 以找到与这些目标用户关系密切或者具有相似偏好的用户群。这些用户群以及他们之间的关联关系形成无向连通图 G 的一个生成子图。

具体地, 步骤 S30 包括:

15 删除无向连通图中具有最小度的节点, 以及与该节点相连的边; 判断经节点删除操作后的无向连通图是否满足第二预设条件, 其中第二预设条件为: 所述查询节点集中的节点在节点删除操作后的图无向连通图中具有最小度, 或者, 经节点删除操作后, 查询节点集中的节点在该无向连通图上不再连通; 若是, 则终止节点删除操作, 将当前的无向连通图作为所述生成子图; 若否,
20 则基于当前的无向连通图继续执行所述删除无向连通图中具有最小度的节点, 以及与该节点相连的边的步骤。

该社区搜索算法主要包括: 将 G_0 作为算法的初始输入图, 其中, $G_0=G$; 每一步删除该步骤中输入的图上的一个节点以及与该节点关联的边, 该节点是当前的图中具有最小度的节点, 其中, 在无向连通图中, 节点的度等于连接该节点的边的数量。即在第 t 步中, 计算当前的图 G_{t-1} 上具有最小度的节点, 删除该节点以及所有与该节点连接的边, 得到图 G_t , 作为下一步的输入。
25 如此循环往复的删除每一步骤的图中具有最小度的节点, 直至执行了节点删除操作后得到的图满足以下条件之一: 查询节点集中的节点在图中具有最小度, 或者, 删除有最小度的节点后, 查询节点集中的点不再连通。这时, 终止节点删除操作, 获取计算终止后得到的生成子图, 该生成子图是一个连通
30

图，并且查询节点集是该生成子图中节点集的一个子集。

按照上述算法获取到的生成子图中的节点对应的用户即为要搜索的社区，这些用户具有与目标用户相同的特征，例如，具有相同的感兴趣产品或者具有相同的特征标签等。

5 进一步地，在其他的实施例中，为了避免计算得到的社区范围过大，可以进一步地增加预设条件：生成子图中的节点与查询节点集 Q 中的节点的最大距离不超过预设距离。在判定经节点删除操作后的无向连通图满足第二预设条件之后，终止节点删除操作之前，计算当前的无向连通图中的节点到所述查询节点集之间的距离，判断是否有节点到查询节点的距离大于预设距离；
10 若是，则删除到查询节点的距离大于预设距离的节点，并将删除节点后的无向连通图作为所述生成子图；若否，则执行将终止节点删除操作，将当前的无向连通图作为所述生成子图的步骤。

 关于节点之间的距离的计算可以采用如下方式中的任意一种：方式一，将两个节点之间的最短路径中的边的数量作为节点之间的距离，由于在一个
15 复杂的无向连通图中，可以由多种路径实现两节点之间的连通，此处可以根据多条路径中的最短路径来计算节点之间的距离；方式二，根据两个节点之间的最短路径中的边的权重计算节点之间的距离，具体地，从目标社交网络平台上获取到的用户信息中包含有标签信息，根据两个用户拥有的相同标签的数量设置他们之间的权重，相同标签的数量越多，则权重越大，否则权重
20 越小，将两个节点之间的最短路径中的边的权重之和作为两个节点之间的距离。其中，标签信息可以是用户在社交网络平台上可以为自己添加的标签，这些标签体现用户的感兴趣标签信息，或者，也可以将从用户在该社交网路平台上发布的内容提取的关键字作为用户的标签信息。此外，在计算一个节点到查询节点集的距离时，先计算该节点到查询节点集中每一个节点的距离，
25 然后将上述多个距离中的最小距离作为该节点到查询节点集的距离。

 此外，可以理解的是，在上述删除节点的过程中，如果删除了某个节点之后，有节点在图上变成孤立的点，与其他节点之间没有任何连接关系时，要把这些孤立的节点删除。

 通过上述计算获取到目标用户所在的社区，即由生成子图中的节点对应的
30 的用户构成的社区，这些用户是目标用户的相似用户，例如，具体在一个金

融论坛中，可以认为在同一个社区中的用户有相同或者相似的投资偏好、属性、标签、感兴趣事件等。在得到相似用户后，根据它们之间相同或者相似的投资偏好、属性、标签、感兴趣事件等信息做精准推荐或定向投放。

以上实施例提出的相似用户查找方法，根据目标社交网络平台中的用户信息构建无向连通图，其中，在无向连通图上一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接，获取由多个具有关联关系的目标用户构成的查询节点集，同时，这个查询节点集是五香连通图中的节点集合的子集，根据查询节点集合预设的社区搜索算法在无向连通图上进行迭代运算，以删除无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成目标用户所属的社区，而在无向连通图上，属于一个社区的用户具有相似的属性，因此，该生成子图上的节点对应的用户可以作为上述目标用户的相似用户。本申请通过社交网络平台进行社区搜索，查找目标用户所述的社区，进而获取目标用户的相似用户，提高了查找相似用户的准确度。

此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质上存储有用户查找程序，所述用户查找程序可被一个或多个处理器执行，以实现如下操作：

根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接；

获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点集为所述无向连通图中的节点集合的子集；

根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成所述目标用户所属的社区；

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

本申请计算机可读存储介质具体实施方式与上述基于社区搜索的相似用户查找装置和方法各实施例基本相同，在此不作累述。

需要说明的是，上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。并且本文中的术语“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、装置、物品或者方法不仅包括那些要素，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、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献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如上所述的一个存储介质(如 ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

权利要求书

1、一种基于社区搜索的相似用户查找装置，其特征在于，所述装置包括存储器和处理器，所述存储器上存储有可在所述处理器上运行的用户查找程序，所述用户查找程序被所述处理器执行时实现如下步骤：

5 根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接；

获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点集为所述无向连通图中的节点集合的子集；

10 根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成所述目标用户所属的社区；

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

2、如权利要求1所述的基于社区搜索的相似用户查找装置，其特征在于，
15 所述根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图的步骤包括：

删除无向连通图中具有最小度的节点，以及与该节点相连的边；

判断经节点删除操作后的无向连通图是否满足第二预设条件，其中第二
20 预设条件为：所述查询节点集中的节点在节点删除操作后的图无向连通图中具有最小度，或者，经节点删除操作后，查询节点集中的节点在该无向连通图上不再连通；

若是，则终止节点删除操作，将当前的无向连通图作为所述生成子图；

25 若否，则基于当前的无向连通图继续执行所述删除无向连通图中具有最小度的节点，以及与该节点相连的边的步骤。

3、如权利要求2所述的基于社区搜索的相似用户查找装置，其特征在于，所述用户查找程序还可被所述处理器执行，以在所述终止节点删除操作，将当前的无向连通图作为所述生成子图的步骤之前，还实现如下步骤：

30 若经节点删除操作后的无向连通图满足第二预设条件，则计算当前的无向连通图中的节点到所述查询节点集之间的距离，判断是否有节点到查询节

点的距离大于预设距离;

若是,则删除到查询节点的距离大于预设距离的节点,并将删除节点后的无向连通图作为所述生成子图;

5 若否,则执行将终止节点删除操作,将当前的无向连通图作为所述生成子图的步骤。

4、如权利要求3所述的基于社区搜索的相似用户查找装置,其特征在于,所述计算当前的无向连通图中的节点到所述查询节点集之间的距离的步骤包括:

10 根据两个节点之间的最短路径中的边的数量或者权重,计算所述无向连通图中的节点到所述查询节点集之间的距离。

5、如权利要求1所述的基于社区搜索的相似用户查找装置,其特征在于,所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括:

获取目标社交网络平台中的用户信息,根据获取的用户信息确定用户之间的关联关系;

15 根据获取到的用户信息和用户之间的关联关系构建无向连通图。

6、如权利要求2所述的基于社区搜索的相似用户查找装置,其特征在于,所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括:

获取目标社交网络平台中的用户信息,根据获取的用户信息确定用户之间的关联关系;

20 根据获取到的用户信息和用户之间的关联关系构建无向连通图。

7、如权利要求3所述的基于社区搜索的相似用户查找装置,其特征在于,所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括:

获取目标社交网络平台中的用户信息,根据获取的用户信息确定用户之间的关联关系;

25 根据获取到的用户信息和用户之间的关联关系构建无向连通图。

8、一种基于社区搜索的相似用户查找方法,其特征在于,所述方法包括:

根据目标社交网络平台中的用户信息构建无向连通图,其中,在所述无向连通图上,一个节点对应于一个用户,具有关联关系的用户之间通过一条边相连接;

30 获取由多个具有关联关系的目标用户构成的查询节点集,所述查询节点

集为所述无向连通图中的节点集合的子集;

根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算,以删除所述无向连通图上的节点,直至删除节点后得到满足第二预设条件的生成子图,该生成子图中的节点集合构成所述目标用户所属的社区;

5 将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

9、如权利要求 8 所述的基于社区搜索的相似用户查找方法,其特征在于,所述根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算,以删除所述无向连通图上的节点,直至删除节点后得到满足第二预设条件的生成子图的步骤包括:

10 删除无向连通图中具有最小度的节点,以及与该节点相连的边;

判断经节点删除操作后的无向连通图是否满足第二预设条件,其中第二预设条件为:所述查询节点集中的节点在节点删除操作后的图无向连通图中具有最小度,或者,经节点删除操作后,查询节点集中的节点在该无向连通图上不再连通;

15 若是,则终止节点删除操作,将当前的无向连通图作为所述生成子图;

若否,则基于当前的无向连通图继续执行所述删除无向连通图中具有最小度的节点,以及与该节点相连的边的步骤。

10、如权利要求 9 所述的基于社区搜索的相似用户查找方法,其特征在于,所述终止节点删除操作,将当前的无向连通图作为所述生成子图的步骤之前,所述方法还包括如下步骤:

若经节点删除操作后的无向连通图满足第二预设条件,则计算当前的无向连通图中的节点到所述查询节点集之间的距离,判断是否有节点到查询节点的距离大于预设距离;

25 若是,则删除到查询节点的距离大于预设距离的节点,并将删除节点后的无向连通图作为所述生成子图;

若否,则执行将终止节点删除操作,将当前的无向连通图作为所述生成子图的步骤。

11、如权利要求 10 所述的基于社区搜索的相似用户查找方法,其特征在于,所述计算当前的无向连通图中的节点到所述查询节点集之间的距离的步骤包括:

30

根据两个节点之间的最短路径中的边的数量或者权重，计算所述无向连通图中的节点到所述查询节点集之间的距离。

12、如权利要求 8 所述的基于社区搜索的相似用户查找方法，其特征在于，所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括：

5 获取目标社交网络平台中的用户信息，根据获取的用户信息确定用户之间的关联关系；

根据获取到的用户信息和用户之间的关联关系构建无向连通图。

13、如权利要求 9 所述的基于社区搜索的相似用户查找方法，其特征在于，所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括：

10 获取目标社交网络平台中的用户信息，根据获取的用户信息确定用户之间的关联关系；

根据获取到的用户信息和用户之间的关联关系构建无向连通图。

14、如权利要求 10 所述的基于社区搜索的相似用户查找方法，其特征在于，所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括：

15 获取目标社交网络平台中的用户信息，根据获取的用户信息确定用户之间的关联关系；

根据获取到的用户信息和用户之间的关联关系构建无向连通图。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有用户查找程序，所述用户查找程序可被一个或者多个处理器执行，

20 以实现如下步骤：

根据目标社交网络平台中的用户信息构建无向连通图，其中，在所述无向连通图上，一个节点对应于一个用户，具有关联关系的用户之间通过一条边相连接；

25 获取由多个具有关联关系的目标用户构成的查询节点集，所述查询节点集为所述无向连通图中的节点集合的子集；

根据所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图，该生成子图中的节点集合构成所述目标用户所属的社区；

将所述生成子图中的节点对应的用户作为所述多个目标用户的相似用户。

30 16、如权利要求 15 所述的计算机可读存储介质，其特征在于，所述根据

所述查询节点集和预设的社区搜索算法在所述无向连通图上进行迭代运算，以删除所述无向连通图上的节点，直至删除节点后得到满足第二预设条件的生成子图的步骤包括：

删除无向连通图中具有最小度的节点，以及与该节点相连的边；

- 5 判断经节点删除操作后的无向连通图是否满足第二预设条件，其中第二预设条件为：所述查询节点集中的节点在节点删除操作后的图无向连通图中具有最小度，或者，经节点删除操作后，查询节点集中的节点在该无向连通图上不再连通；

若是，则终止节点删除操作，将当前的无向连通图作为所述生成子图；

- 10 若否，则基于当前的无向连通图继续执行所述删除无向连通图中具有最小度的节点，以及与该节点相连的边的步骤。

17、如权利要求 16 所述的计算机可读存储介质，其特征在于，所述用户查找程序可被一个或者多个处理器执行，以在所述终止节点删除操作，将当前的无向连通图作为所述生成子图的步骤之前，还实现如下步骤：

- 15 若经节点删除操作后的无向连通图满足第二预设条件，则计算当前的无向连通图中的节点到所述查询节点集之间的距离，判断是否有节点到查询节点的距离大于预设距离；

若是，则删除到查询节点的距离大于预设距离的节点，并将删除节点后的无向连通图作为所述生成子图；

- 20 若否，则执行将终止节点删除操作，将当前的无向连通图作为所述生成子图的步骤。

18、如权利要求 17 所述的计算机可读存储介质，其特征在于，所述计算当前的无向连通图中的节点到所述查询节点集之间的距离的步骤包括：

- 25 根据两个节点之间的最短路径中的边的数量或者权重，计算所述无向连通图中的节点到所述查询节点集之间的距离。

19、如权利要求 15 所述的计算机可读存储介质，其特征在于，所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括：

获取目标社交网络平台中的用户信息，根据获取的用户信息确定用户之间的关联关系；

- 30 根据获取到的用户信息和用户之间的关联关系构建无向连通图。

20、如权利要求 16 所述的计算机可读存储介质，其特征在于，所述根据目标社交网络平台中的用户信息构建无向连通图的步骤包括：

获取目标社交网络平台中的用户信息，根据获取的用户信息确定用户之间的关联关系；

5 根据获取到的用户信息和用户之间的关联关系构建无向连通图。

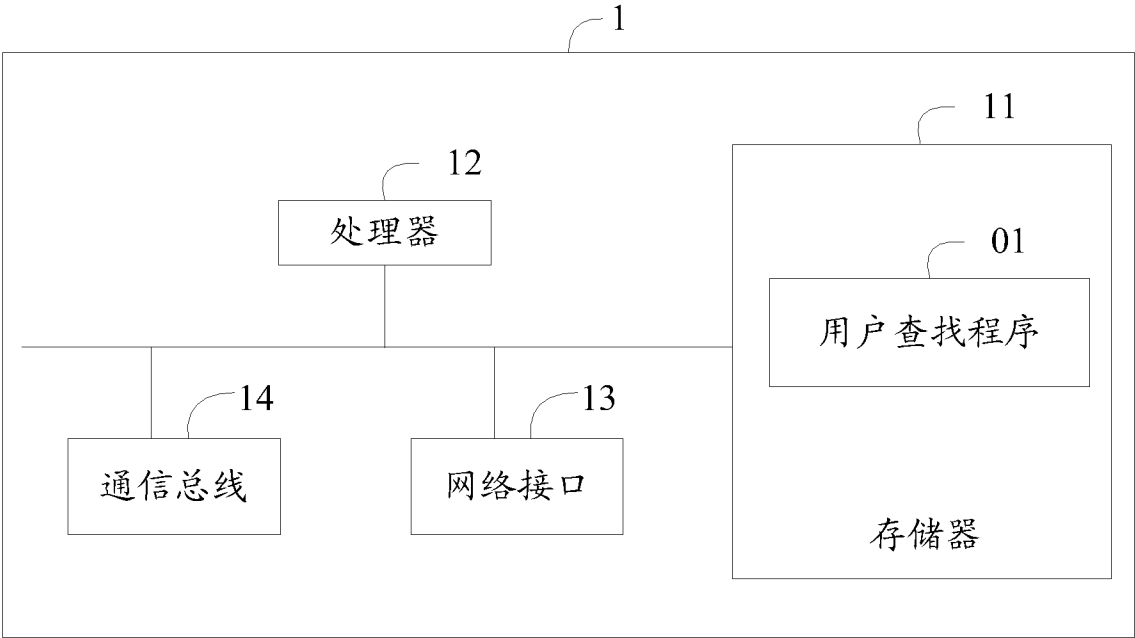


图 1

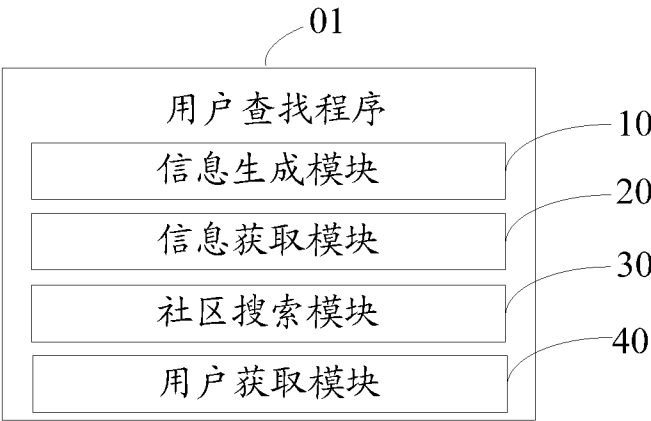


图 2

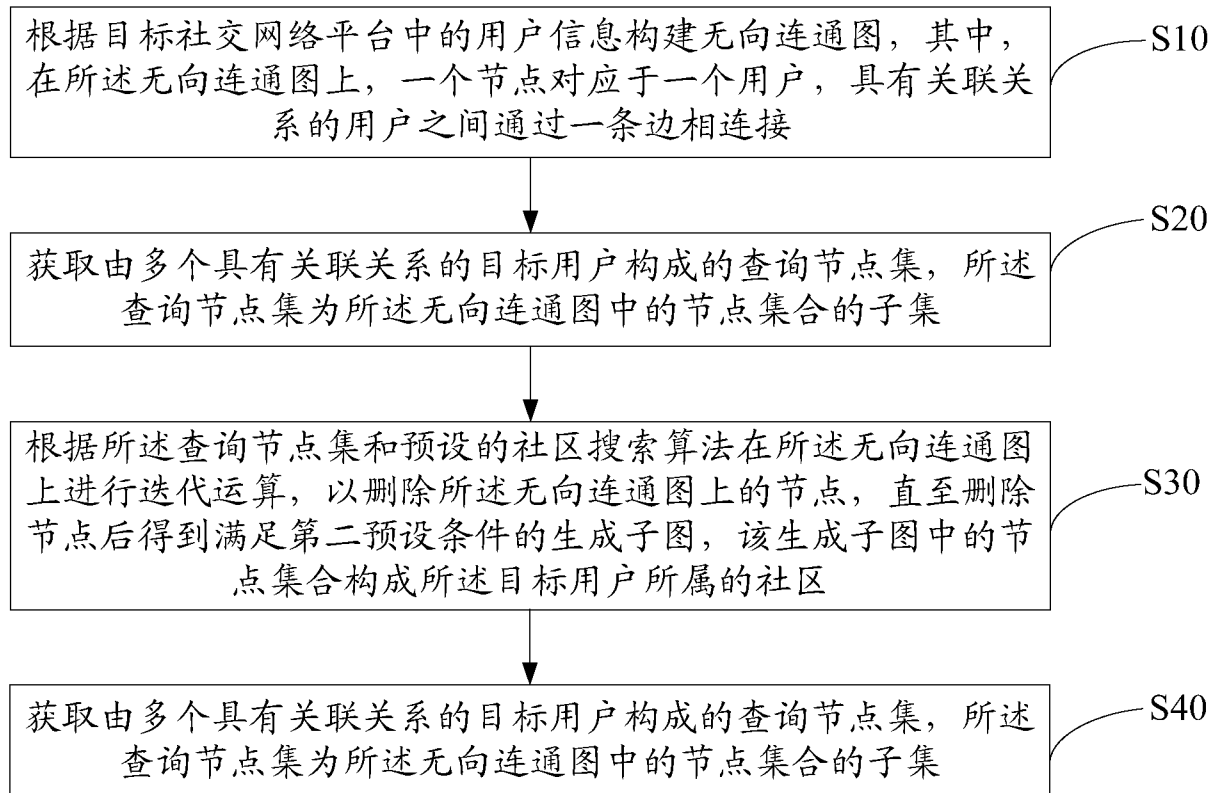


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/102112

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/9536(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)

WPI, EPODOC, CNPAT, CNKI: 社区, 社团, 群体, 无向图, 相似, 类似, 相同, 社交网络, 用户, 节点, 边, 删除, 去除, 划分, graph, group, user, similar, node, edge, delete, social network, SNS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 105721279 A (CHINA UNITED NETWORK COMMUNICATIONS LIMITED, GUANGDONG BRANCH ET AL.) 29 June 2016 (2016-06-29) description, paragraphs [0042]-[0046] and [0058]	1, 5, 8, 12, 15, 19
A	CN 104731962 A (CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 24 June 2015 (2015-06-24) entire document	1-20
A	CN 104102675 A (RENMIN UNIVERSITY OF CHINA) 15 October 2014 (2014-10-15) entire document	1-20
A	US 8606787 B1 (GOOGLE INC.) 10 December 2013 (2013-12-10) entire document	1-20
A	US 2007226248 A1 (TIMOTHY PAUL DARR) 27 September 2007 (2007-09-27) entire document	1-20



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

10 January 2019

Date of mailing of the international search report

29 January 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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/102112

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	105721279	A	29 June 2016	None			
CN	104731962	A	24 June 2015	None			
CN	104102675	A	15 October 2014	None			
US	8606787	B1	10 December 2013	None			
US	2007226248	A1	27 September 2007	WO	2007109726	A2	27 September 2007

国际检索报告

国际申请号

PCT/CN2018/102112

A. 主题的分类

G06F 16/9536 (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))

WPI, EPODOC, CNPAT, CNKI: 社区, 社团, 群体, 无向图, 相似, 类似, 相同, 社交网络, 用户, 节点, 边, 删除, 去除, 划分, graph, group, user, similar, node, edge, delete, social network, SNS

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
X	CN 105721279 A (中国联合网络通信有限公司广东省分公司 等) 2016年 6月 29日 (2016 - 06 - 29) 说明书第[0042]-[0046], [0058]段	1, 5, 8, 12, 15, 19
A	CN 104731962 A (重庆邮电大学) 2015年 6月 24日 (2015 - 06 - 24) 全文	1-20
A	CN 104102675 A (中国人民大学) 2014年 10月 15日 (2014 - 10 - 15) 全文	1-20
A	US 8606787 B1 (GOOGLE INC.) 2013年 12月 10日 (2013 - 12 - 10) 全文	1-20
A	US 2007226248 A1 (TIMOTHY PAUL DARR) 2007年 9月 27日 (2007 - 09 - 27) 全文	1-20

☐ 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 1月 10日

国际检索报告邮寄日期

2019年 1月 29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

受权官员

吴琼

电话号码 86-(10)-53961338

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/102112

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	105721279	A	2016年 6月 29日	无			
CN	104731962	A	2015年 6月 24日	无			
CN	104102675	A	2014年 10月 15日	无			
US	8606787	B1	2013年 12月 10日	无			
US	2007226248	A1	2007年 9月 27日	WO	2007109726	A2	2007年 9月 27日