



(12) 发明专利申请

(10) 申请公布号 CN 112101999 A

(43) 申请公布日 2020.12.18

(21) 申请号 202010971485.4

(22) 申请日 2020.09.16

(71) 申请人 智慧足迹数据科技有限公司
地址 100000 北京市西城区西单北大街甲
131号西单大悦城写字楼1105B

(72) 发明人 刘祖军 陶周天 张岩 李振军
邹炎炎

(74) 专利代理机构 北京超凡宏宇专利代理事务
所(特殊普通合伙) 11463
代理人 张欣欣

(51) Int.Cl.
G06Q 30/02 (2012.01)
G06K 9/62 (2006.01)

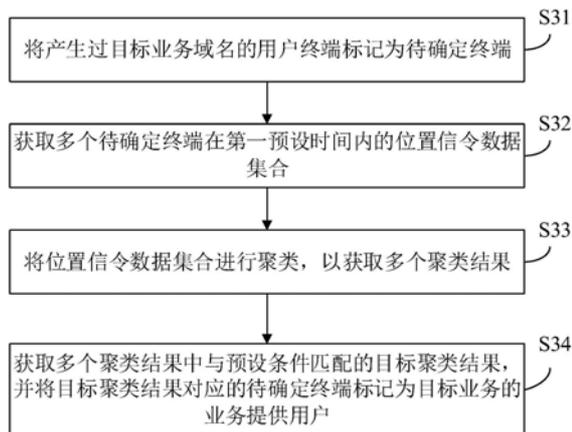
权利要求书2页 说明书7页 附图6页

(54) 发明名称

用户识别方法、装置、电子设备和计算机可读存储介质

(57) 摘要

本申请提供一种用户识别方法、装置、电子设备和计算机可读存储介质,涉及网络数据处理领域。所述方法包括:将产生过目标业务域名的用户终端标记为待确定终端;获取多个所述待确定终端在第一预设时间内的位置信令数据集合;将所述位置信令数据集合进行聚类,以获取多个聚类结果;获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通过聚类获得业务提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。



1. 一种用户识别方法,其特征在于,所述方法包括:
将产生过目标业务域名的用户终端标记为待确定终端;
获取多个所述待确定终端在第一预设时间内的位置信令数据集合;
将所述位置信令数据集合进行聚类,以获取多个聚类结果;
获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

2. 根据权利要求1所述的方法,其特征在于,将产生过目标业务域名的用户终端标记为待确定终端,包括:

获取包括多个业务域名的域名列表;

将所述域名列表中与所述目标业务匹配的业务域名作为所述目标业务域名;

获取所述目标业务的多个历史使用信令;

将所述多个历史使用信令中包括所述目标业务域名的使用信令的用户终端标记为所述待确定终端。

3. 根据权利要求1所述的方法,其特征在于,获取多个所述待确定终端在第一预设时间内的位置信令数据集合,包括:

获取所述待确定终端在所述第一预设时间内的驻留点位置;所述驻留点位置为所述待确定终端的停留时间大于或等于第二预设时间的位置信息;

获取第三预设时间内,所述待确定终端在所述驻留点位置的驻留次数;所述第一预设时间大于所述第三预设时间,所述第三预设时间大于所述第二预设时间;

将每个所述待确定终端的所述驻留点位置和所述驻留点位置对应的驻留次数作为所述位置信令数据集合。

4. 根据权利要求3所述的方法,其特征在于,将每个所述待确定终端的所述驻留点位置和所述驻留点位置对应的驻留次数作为所述位置信令数据集合,包括:

将所述驻留点位置和所述驻留点位置对应的驻留次数进行最大池化处理,以获取所述待确定终端在第四预设时间内的驻留数据;所述第一预设时间大于所述第四预设时间,所述第四预设时间大于所述第三预设时间;

将多个所述驻留数据进行聚合,以得到所述位置信令数据集合。

5. 根据权利要求1所述的方法,其特征在于,将所述位置信令数据集合进行聚类,以获取多个聚类结果,包括:

使用时部法则确定所述位置信令数据集合的类别数量;

根据所述类别数量对所述位置信令数据集合进行聚类,以获取多个聚类结果;所述聚类结果的个数为所述类别数量。

6. 一种用户识别装置,其特征在于,所述装置包括:

标记模块,用于将产生过目标业务域名的用户终端标记为待确定终端;

获取模块,用于获取多个所述待确定终端在第一预设时间内的位置信令数据集合;

聚类模块,用于将所述位置信令数据集合进行聚类,以获取多个聚类结果;

处理模块,用于获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

7. 根据权利要求6所述的装置,其特征在于,所述标记模块还用于获取包括多个业务域

名的域名列表；

所述标记模块还用于将所述域名列表中与所述目标业务匹配的业务域名作为所述目标业务域名；

所述标记模块还用于获取所述目标业务的多个历史使用信令；

所述标记模块还用于将所述多个历史使用信令中包括所述目标业务域名的使用信令的用户终端标记为所述待确定终端。

8. 根据权利要求6所述的装置,其特征在於,所述获取模块还用于获取所述待确定终端在所述第一预设时间内的驻留点位置;所述驻留点位置为所述待确定终端的停留时间大于或等于第二预设时间的位置信息;

所述获取模块还用于获取第三预设时间内,所述待确定终端在所述驻留点位置的驻留次数;所述第一预设时间大于所述第三预设时间,所述第三预设时间大于所述第二预设时间;

所述获取模块还用于将每个所述待确定终端的所述驻留点位置和所述驻留点位置对应的驻留次数作为所述位置信令数据集。

9. 一种电子设备,其特征在於,包括处理器和存储器,所述存储器存储有能够被所述处理器执行的机器可执行指令,所述处理器可执行所述机器可执行指令以实现权利要求1-5中任意一项所述的方法。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在於,所述计算机程序被处理器执行时实现权利要求1-5中任意一项所述的方法。

用户识别方法、装置、电子设备和计算机可读存储介质

技术领域

[0001] 本申请涉及网络数据处理领域,具体而言,涉及一种用户识别方法、装置、电子设备和计算机可读存储介质。

背景技术

[0002] 随着社会的进步和互联网的发展,网约车和外卖业务逐渐占据了公共交通和餐饮行业的主导。

[0003] 对于外卖业务,为了确定外卖骑手的位置,通过解析手机用户的应用程序(Application,APP)使用情况,以便抓取有使用外卖骑手APP的用户标记为外卖骑手。但是,APP解析的方法是抓包APP记录并进行域名解析。但是往往由于外卖平台同时运营外卖骑手平台与外卖点餐平台导致不同APP的域名解析结果相同(例如同一个厂商下的外卖与骑手最终都指向该厂商),所以会造成将点外卖的人群误判为外卖骑手的情况;且,用户使用APP存在不确定性;例如用户操作手机时的误触或者APP在手机后台持续运行会导致错误记录,继而导致外卖骑手与点餐者产生错误识别;网约车也存在和外卖业务相同的问题。因此,如何针准确的识别业务供给用户是目前亟需解决的问题。

发明内容

[0004] 本申请的目的包括,例如,提供了一种用户识别方法、装置、电子设备和计算机可读存储介质,其能够通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通过聚类获得业务提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。

[0005] 本申请的实施例可以这样实现:

[0006] 第一方面,本申请实施例提供一种用户识别方法,所述方法包括:将产生过目标业务域名的用户终端标记为待确定终端;获取多个所述待确定终端在第一预设时间内的位置信令数据集合;将所述位置信令数据集合进行聚类,以获取多个聚类结果;获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

[0007] 在可选的实施方式中,将产生过目标业务域名的用户终端标记为待确定终端,包括:获取包括多个业务域名的域名列表;将所述域名列表中与所述目标业务匹配的业务域名作为所述目标业务域名;获取所述目标业务的多个历史使用信令;将所述多个历史使用信令中包括所述目标业务域名的使用信令的用户终端标记为所述待确定终端。

[0008] 在可选的实施方式中,获取多个所述待确定终端在第一预设时间内的位置信令数据集合,包括:获取所述待确定终端在所述第一预设时间内的驻留点位置;所述驻留点位置为所述待确定终端的停留时间大于或等于第二预设时间的位置信息;获取第三预设时间内,所述待确定终端在所述驻留点位置的驻留次数;所述第一预设时间大于所述第三预设时间,所述第三预设时间大于所述第二预设时间;将每个所述待确定终端的所述驻留点位

置和所述驻留点位置对应的驻留次数作为所述位置信令数据集合。

[0009] 在可选的实施方式中,将每个所述待确定终端的所述驻留点位置和所述驻留点位置对应的驻留次数作为所述位置信令数据集合,包括:将所述驻留点位置和所述驻留点位置对应的驻留次数进行最大池化处理,以获取所述待确定终端在第四预设时间内的驻留数据;所述第一预设时间大于所述第四预设时间,所述第四预设时间大于所述第三预设时间;将多个所述驻留数据进行聚合,以得到所述位置信令数据集合。

[0010] 在可选的实施方式中,将所述位置信令数据集合进行聚类,以获取多个聚类结果,包括:使用肘部法则确定所述位置信令数据集合的类别数量;根据所述类别数量对所述位置信令数据集合进行聚类,以获取多个聚类结果;所述聚类结果的个数为所述类别数量。

[0011] 第二方面,本申请实施例提供一种用户识别装置,所述装置包括:标记模块,用于将产生过目标业务域名的用户终端标记为待确定终端;获取模块,用于获取多个所述待确定终端在第一预设时间内的位置信令数据集合;聚类模块,用于将所述位置信令数据集合进行聚类,以获取多个聚类结果;处理模块,用于获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

[0012] 在可选的实施方式中,所述标记模块还用于获取包括多个业务域名的域名列表;所述标记模块还用于将所述域名列表中与所述目标业务匹配的业务域名作为所述目标业务域名;所述标记模块还用于获取所述目标业务的多个历史使用信令;所述标记模块还用于将所述多个历史使用信令中包括所述目标业务域名的使用信令的用户终端标记为所述待确定终端。

[0013] 在可选的实施方式中,所述获取模块还用于获取所述待确定终端在所述第一预设时间内的驻留点位置;所述驻留点位置为所述待确定终端的停留时间大于或等于第二预设时间的位置信息;

[0014] 所述获取模块还用于获取第三预设时间内,所述待确定终端在所述驻留点位置的驻留次数;所述第一预设时间大于所述第三预设时间,所述第三预设时间大于所述第二预设时间;

[0015] 所述获取模块还用于将每个所述待确定终端的所述驻留点位置和所述驻留点位置对应的驻留次数作为所述位置信令数据集合。

[0016] 第三方面,本申请实施例提供一种电子设备,包括处理器和存储器,所述存储器存储有能够被所述处理器执行的机器可执行指令,所述处理器可执行所述机器可执行指令以实现前述实施方式中任意一项所述的方法。

[0017] 第四方面,本申请实施例提供一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现前述实施方式中任意一项所述的方法。

[0018] 相较于现有技术,本申请提供一种用户识别方法、装置、电子设备和计算机可读存储介质,涉及网络数据处理领域。所述方法包括:将产生过目标业务域名的用户终端标记为待确定终端;获取多个所述待确定终端在第一预设时间内的位置信令数据集合;将所述位置信令数据集合进行聚类,以获取多个聚类结果;获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通过聚类获得业务

提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。

附图说明

[0019] 为了更清楚地说明本申请实施例的技术方案,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申请的某些实施例,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。

[0020] 图1为本申请实施例提供的一种电子设备的方框示意图;

[0021] 图2为本申请实施例提供的一种用户识别方法的流程示意图;

[0022] 图3为本申请实施例提供的另一种用户识别方法的流程示意图;

[0023] 图4为本申请实施例提供的另一种用户识别方法的流程示意图;

[0024] 图5为本申请实施例提供的另一种用户识别方法的流程示意图;

[0025] 图6为本申请实施例提供的一种位置信令数据集合的折线图;

[0026] 图7为本申请实施例提供的另一种位置信令数据集合的折线图;

[0027] 图8为本申请实施例提供的一种聚类结果的示意图;

[0028] 图9为本申请实施例提供的一种用户识别装置的方框示意图。

具体实施方式

[0029] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请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申请一部分实施例,而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。

[0030] 因此,以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。此外,若出现术语“第一”、“第二”等仅用于区分描述,而不能理解为指示或暗示相对重要性。需要说明的是,在不冲突的情况下,本申请的实施例中的特征可以相互结合。

[0031] 为了至少解决背景技术提出的不足,本申请实施例提供一种用户识别方法,应用于电子设备,请参见图1,图1为本申请实施例提供的一种电子设备的方框示意图。该电子设备20包括存储器21、处理器22和通信接口23。该存储器21、处理器22和通信接口23相互之间直接或间接地电性连接,以实现数据的传输或交互。例如,这些元件相互之间可通过一条或多条通讯总线或信号线实现电性连接。存储器21可用于存储软件程序及模块,如本申请实施例所提供的用户识别方法对应的程序指令/模块,处理器22通过执行存储在存储器21内的软件程序及模块,从而执行各种功能应用以及数据处理。该通信接口23可用于与其他节点设备进行信令或数据的通信。在本申请中该电子设备20可以具有多个通信接口23。

[0032] 其中,存储器21可以是但不限于,随机存取存储器(Random Access Memory,RAM),只读存储器(Read Only Memory,ROM),可编程只读存储器(Programmable Read-Only Memory,PROM),可擦除只读存储器(Erasable Programmable Read-Only Memory,EPRM),

电可擦除只读存储器 (Electric Erasable Programmable Read-Only Memory, EEPROM) 等。

[0033] 处理器22可以是一种集成电路芯片,具有信号处理能力。该处理器可以是通用处理器,包括中央处理器 (Central Processing Unit, CPU)、网络处理器 (Network Processor, NP) 等;还可以是数字信号处理器 (Digital Signal Processing, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现场可编程门阵列 (Field-Programmable Gate Array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。

[0034] 电子设备20可以实现本申请提供的任一种用户识别方法。该电子设备20可以是,但不限于手机、平板电脑、可穿戴设备、车载设备、增强现实 (Augmented Reality, AR) /虚拟现实 (Virtual Reality, VR) 设备、笔记本电脑、超级移动个人计算机 (Ultra-Mobile Personal Computer, UMPC)、上网本、个人数字助理 (Personal Digital Assistant, PDA)、服务器等,本申请实施例对电子设备的具体类型不作任何限制。

[0035] 为了至少解决背景技术提出的不足,下面在图1示出的电子设备20的基础上,本申请实施例提供一种用户识别方法,应用于电子设备,请参见图2,图2为本申请实施例提供了一种用户识别方法的流程示意图,该用户识别方法可以包括以下步骤:

[0036] S31,将产生过目标业务域名的用户终端标记为待确定终端。

[0037] 例如,该目标业务域名可以是,但不限于“meituan”、“eleme”、“ali”等。该用户终端可以是,但不限于手机、平板电脑、笔记本电脑、可穿戴设备、车载设备、AR/VR设备、笔记本电脑、UMPC、上网本、PDA等。

[0038] S32,获取多个待确定终端在第一预设时间内的位置信令数据集合。

[0039] 该第一预设时间可以是30天,15天等,该位置信令数据集合可以包括,但不限于待确定终端的驻留点 (停留的位置坐标)、待确定终端在驻留点的驻留次数、驻留时间等信息。

[0040] S33,将位置信令数据集合进行聚类,以获取多个聚类结果。

[0041] 应理解,聚类时可以选用多种聚类方式,如无监督的K-means聚类等。在一种可能的情况下,上述的S33可以包括:使用肘部法则确定位置信令数据集合的类别数量;根据类别数量对位置信令数据集合进行聚类,以获取多个聚类结果,该聚类结果的个数为聚类的类别数量。

[0042] 应理解,肘部法则的计算原理是成本函数,成本函数是类别畸变程度之和,每个类的畸变程度等于每个变量点到其类别中心的位置距离平方和,若类内部的成员彼此间越紧凑则类的畸变程度越小,反之,若类内部的成员彼此间越分散则类的畸变程度越大。若聚类选用的是无监督的K-means聚类,则使用肘部法则确定的为聚类的K值,如 $K=4$ 。

[0043] S34,获取多个聚类结果中与预设条件匹配的目标聚类结果,并将目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

[0044] 该预设条件可以根据不同的目标业务进行调整,业务提供用户也随目标业务的变化而改变。如,若目标业务为外卖送餐服务,预设条件可以是在用餐高峰期的中午12点和下午6点位置变换频繁情况,该业务提供用户可以是外卖骑手;若目标业务为网约车服务,预设条件可以是在用车高峰期的下午17点和下午22点的位置变换频繁情况,该业务提供用户可以是网约车司机。

[0045] 可以理解的,通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通

过聚类获得业务提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。

[0046] 在可选的实施方式中,为了确定待确定终端,在图2的基础上,给出一种可能的实现方式,请参见图3,图3为本申请实施例提供的另一种用户识别方法的流程示意图,上述的S31可以包括:

[0047] S311,获取包括多个业务域名的域名列表。

[0048] 例如,可以将多个业务的业务域名进行收集和存储,而对于同一业务,由于其提供厂商不同,使用的APP也不相同,可以将其均收录至域名列表中。

[0049] S312,将域名列表中与目标业务匹配的业务域名作为目标业务域名。

[0050] 例如,若目标业务为外卖业务,则与外卖业务匹配的业务域名可以包括:“meituan”、“eleme”等;若目标业务为网约车业务,则与网约车业务匹配的业务域名可以包括:“didi”、“uber”等。

[0051] S313,获取目标业务的多个历史使用信令。

[0052] 该历史使用信令可以是用户终端针对该目标业务产生的业务域名信令,如,若目标业务为外卖业务,用户终端为手机,则该历史使用信令可以是用户终端在一定时间内(如7天)的产生的APP使用信令。

[0053] S314,将多个历史使用信令中包括目标业务域名的使用信令的用户终端标记为待确定终端。

[0054] 如,抓取厂商(如外卖APP提供的运营商)用户使用手机产生的APP使用信令,解析其中域名,筛选有产生过外卖域名信令的人群,记为外卖骑手疑似人群(待确定终端)。

[0055] 在可选的实施方式中,为了获取位置信令数据集合,在图2的基础上,给出一种可能的实现方式,请参见图4,图4为本申请实施例提供的另一种用户识别方法的流程示意图,上述的S32可以包括:

[0056] S321,获取待确定终端在第一预设时间内的驻留点位置。

[0057] 该驻留点位置为待确定终端的停留时间大于或等于第二预设时间的位置信息。例如,该第二预设时间可以是1分钟,3分钟等;位置信息可以通过用户终端的导航系统进行确定,并以地理坐标的形式进行存储。

[0058] S322,获取第三预设时间内,待确定终端在驻留点位置的驻留次数。

[0059] 该第一预设时间大于第三预设时间,第三预设时间大于第二预设时间。例如,该第一预设时间可以是30天,第二预设时间可以是3分钟,第三预设时间可以是1小时。

[0060] S323,将每个待确定终端的驻留点位置和驻留点位置对应的驻留次数作为位置信令数据集合。

[0061] 例如,可以将每个待确定终端的驻留点位置和驻留点位置对应的驻留次数通过表格的形式(如折线图)进行存储和保留。

[0062] 应理解,通过获取待确定终端在第一预设时间内的驻留点位置,以及待确定终端在第三预设时间内于驻留点位置的驻留次数,可以获取待确定终端的位置信令数据,以便电子设备对位置信令数据进行聚类,将符合预设条件的待确定终端作为业务提供用户。

[0063] 在可选的实施方式中,为了进一步的对位置信令数据进行获取,在图4的基础上,给出一种可能的实现方式,请参见图5,图5为本申请实施例提供的另一种用户识别方法的

流程图示意图,上述的S323可以包括:

[0064] S323a,将驻留点位置和驻留点位置对应的驻留次数进行最大池化处理,以获取待确定终端在第四预设时间内的驻留数据。

[0065] 该第一预设时间大于第四预设时间,第四预设时间大于第三预设时间。如,该第一预设时间可以是30天,第四预设时间可以是24小时,第三预设时间可以是1小时。

[0066] S323b,将多个驻留数据进行聚合,以得到位置信令数据集合。

[0067] 应理解,聚合的时候可以通过表格或者图形的形式进行存储。

[0068] 为了便于理解上述实施例提供的用户识别方法,本申请实施例给出一种可能的具体实施例,以目标业务是外卖业务、用户终端是手机为例,将用户识别方法分为三个阶段:

[0069] 第一阶段,通过手机的APP使用情况寻找外卖骑手的疑似人群(待确定终端):首先,建立有外卖业务的域名列表,记为外卖域名(目标业务域名);其次,抓取运营用户使用手机产生的APP使用信令,解析其中域名,筛选有产生过外卖域名信令的人群,记为外卖骑手疑似人群(待确定终端)。

[0070] 第二阶段,建立疑似人群的标准数据集合(位置信令数据集合):首先,选取每个疑似用户(待确定终端)在目标时间窗口(第一预设时间)内全部的位置信令数据中的驻留点数据,按日度切分数据。

[0071] 在该位置信令数据集合内,每个用户在每个时刻都应存在一个驻留点位置,按照驻留点位置再次切割每个用户的数据。依据不同驻留点,统计该用户于该小时内在该驻留点所驻留的次数。例如,小李于2020年6月的每一天凌晨1点都会在家睡觉,则小李对于所居住的居民楼在一点的时刻的次数为30次。请参见图6,图6为本申请实施例提供的一种位置信令数据集合的折线图,横坐标为按照时间进行“0~24”小时的分布,纵坐标为停留于驻留点的次数(0~30次),不同的折线代表不同驻留点的子数据集。

[0072] 将图6示出的位置信令数据集合进行最大池化处理,如,对于每个小时,选取该横坐标下的最大纵坐标值,获得一个长度为24的向量,针对图6示出的位置信令数据集合进行最大池化处理,得到的结果为图7,图7为本申请实施例提供的另一种位置信令数据集合的折线图,横坐标为按照时间进行“0~24”小时的分布,纵坐标为停留于驻留点的次数(0~30次)。应理解,根据图6和图7的方式,可以对全量的外卖骑手疑似人群产生长度为24的标准数据集合,即多个待确定终端的位置信令数据集合。

[0073] 第三阶段,聚类并产生对应标签:对第二阶段所得的数据集合使用K-Means聚类,利用肘部法则得到最优K值;如,对于甲市某月数据,最优K选取为4,四个子类的类中心如图8所示,图8为本申请实施例提供的一种聚类结果的示意图,横坐标为按照时间进行“0~24”小时的分布,纵坐标为停留于驻留点的次数(0~30次)。针对图8示出的曲线1~4,由于曲线3于正常工作时间段对于单一工作地的依赖性较弱,且此特征于用餐高峰期的中午12点和下午6点最为明显,故曲线3对应的待确定终端为实际外卖骑手人群(业务提供用户),进而将曲线3对应的待确定终端标记上外卖骑手标签。

[0074] 为了实现上述任意一个实施例提供的用户识别方法,本申请实施例提供一种用户识别装置,请参见图9,图9为本申请实施例提供的一种用户识别装置的方框示意图,该用户识别装置包括:标记模块41、获取模块42、聚类模块43和处理模块44。

[0075] 标记模块41用于将产生过目标业务域名的用户终端标记为待确定终端;获取模块

42用于获取多个待确定终端在第一预设时间内的位置信令数据集合;聚类模块43用于将位置信令数据集合进行聚类,以获取多个聚类结果;处理模块44用于获取多个聚类结果中与预设条件匹配的目标聚类结果,并将目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。

[0076] 在可选的实施方式中,标记模块41还用于获取包括多个业务域名的域名列表;标记模块41还用于将域名列表中目标业务匹配的业务域名作为目标业务域名;标记模块41还用于获取目标业务的多个历史使用信令;标记模块41还用于将多个历史使用信令中包括目标业务域名的使用信令的用户终端标记为待确定终端。

[0077] 在可选的实施方式中,获取模块42还用于获取待确定终端在第一预设时间内的驻留点位置;驻留点位置为待确定终端的停留时间大于或等于第二预设时间的位置信息;获取模块42还用于获取第三预设时间内,待确定终端在驻留点位置的驻留次数;第一预设时间大于第三预设时间,第三预设时间大于第二预设时间;获取模块42还用于将每个待确定终端的驻留点位置和驻留点位置对应的驻留次数作为位置信令数据集合。

[0078] 应理解,标记模块41、获取模块42、聚类模块43和处理模块44可以协同实现上述任意一个实施例提供的用户识别方法及其可能的子步骤。通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通过聚类获得业务提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。

[0079] 本申请实施例还提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现上述任意一个实施例提供的用户识别方法。该计算机可读存储介质可以是,但不限于,U盘、移动硬盘、ROM、RAM、PROM、EPROM、EEPROM、磁碟或者光盘等各种可以存储程序代码的介质。

[0080] 综上,本申请提供一种用户识别方法、装置、电子设备和计算机可读存储介质,涉及网络数据处理领域。所述方法包括:将产生过目标业务域名的用户终端标记为待确定终端;获取多个所述待确定终端在第一预设时间内的位置信令数据集合;将所述位置信令数据集合进行聚类,以获取多个聚类结果;获取所述多个聚类结果中与预设条件匹配的目标聚类结果,并将所述目标聚类结果对应的待确定终端标记为目标业务的业务提供用户。通过获取待确定终端,根据待确定终端的位置信令数据集合,继而通过聚类获得业务提供用户,解决了单纯的解析用户信令域名来进行用户识别的弊端,准确的识别业务提供用户。

[0081] 以上所述,仅为本申请的具体实施方式,但本申请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申请揭露的技术范围内,可轻易想到的变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

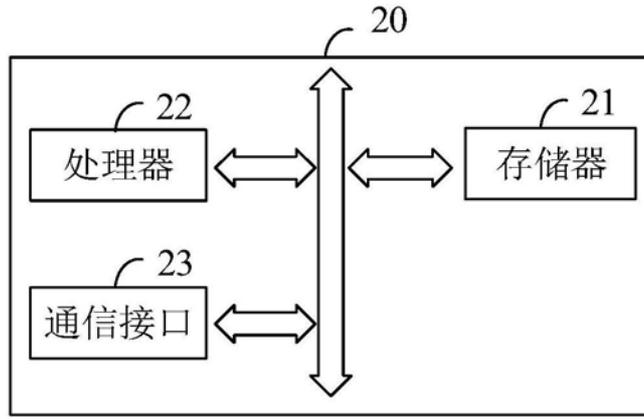


图1

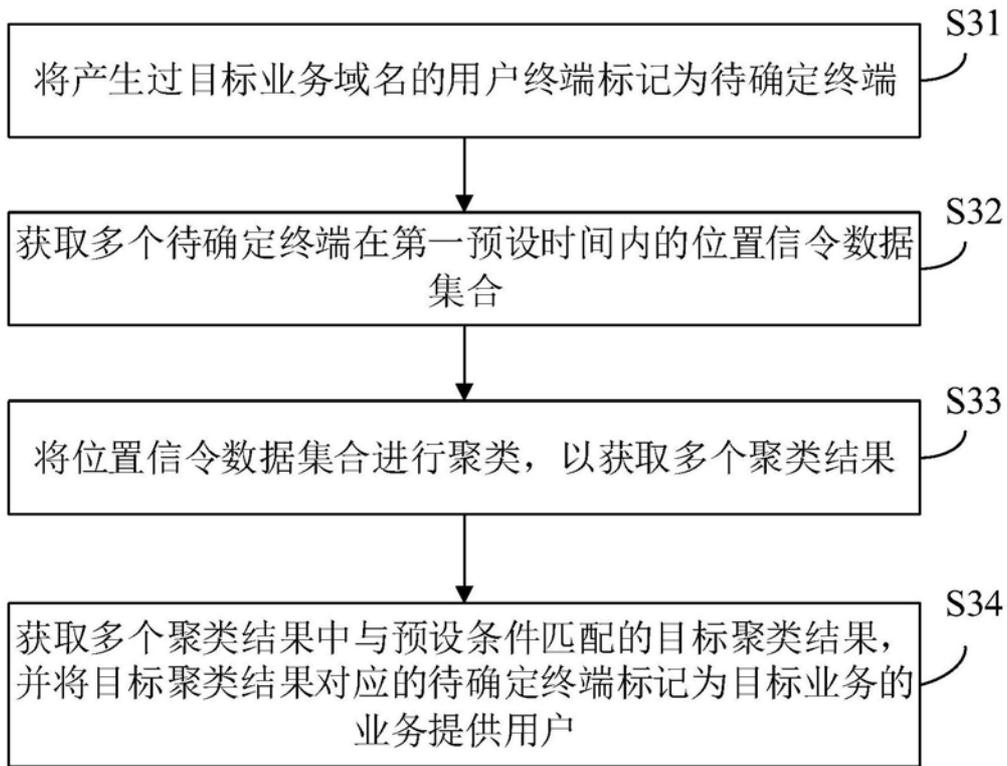


图2

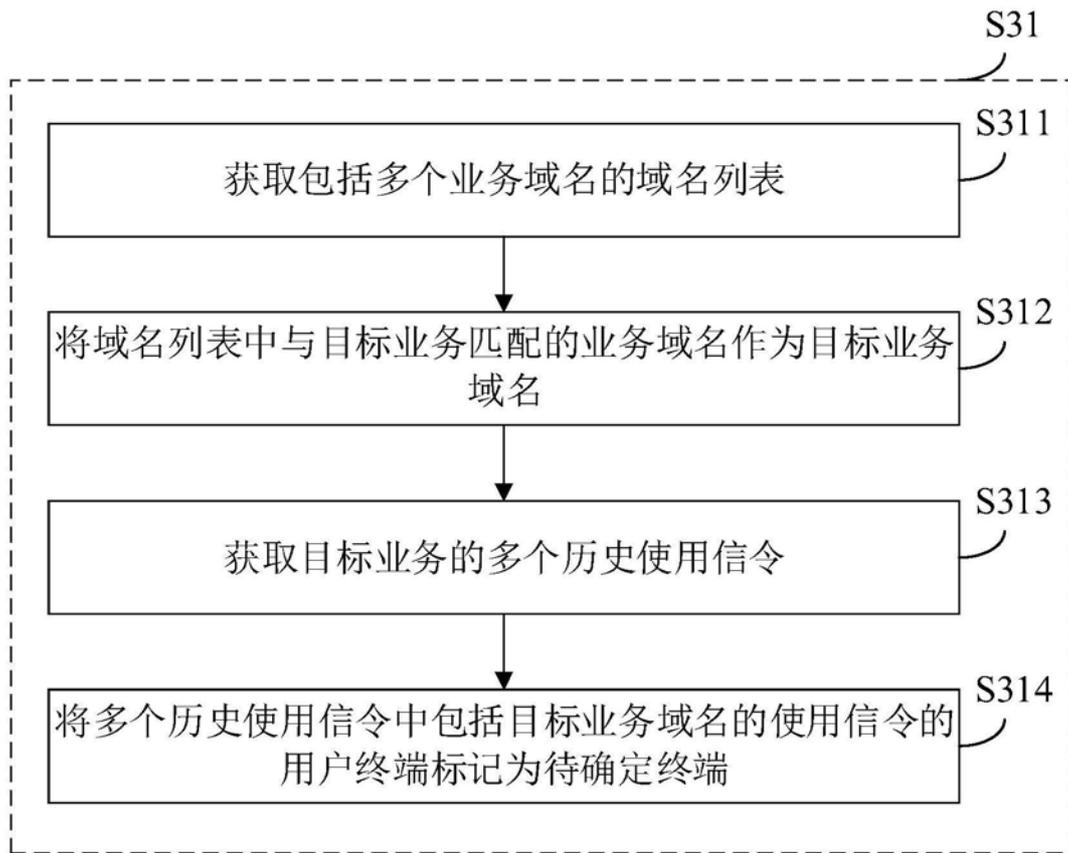


图3

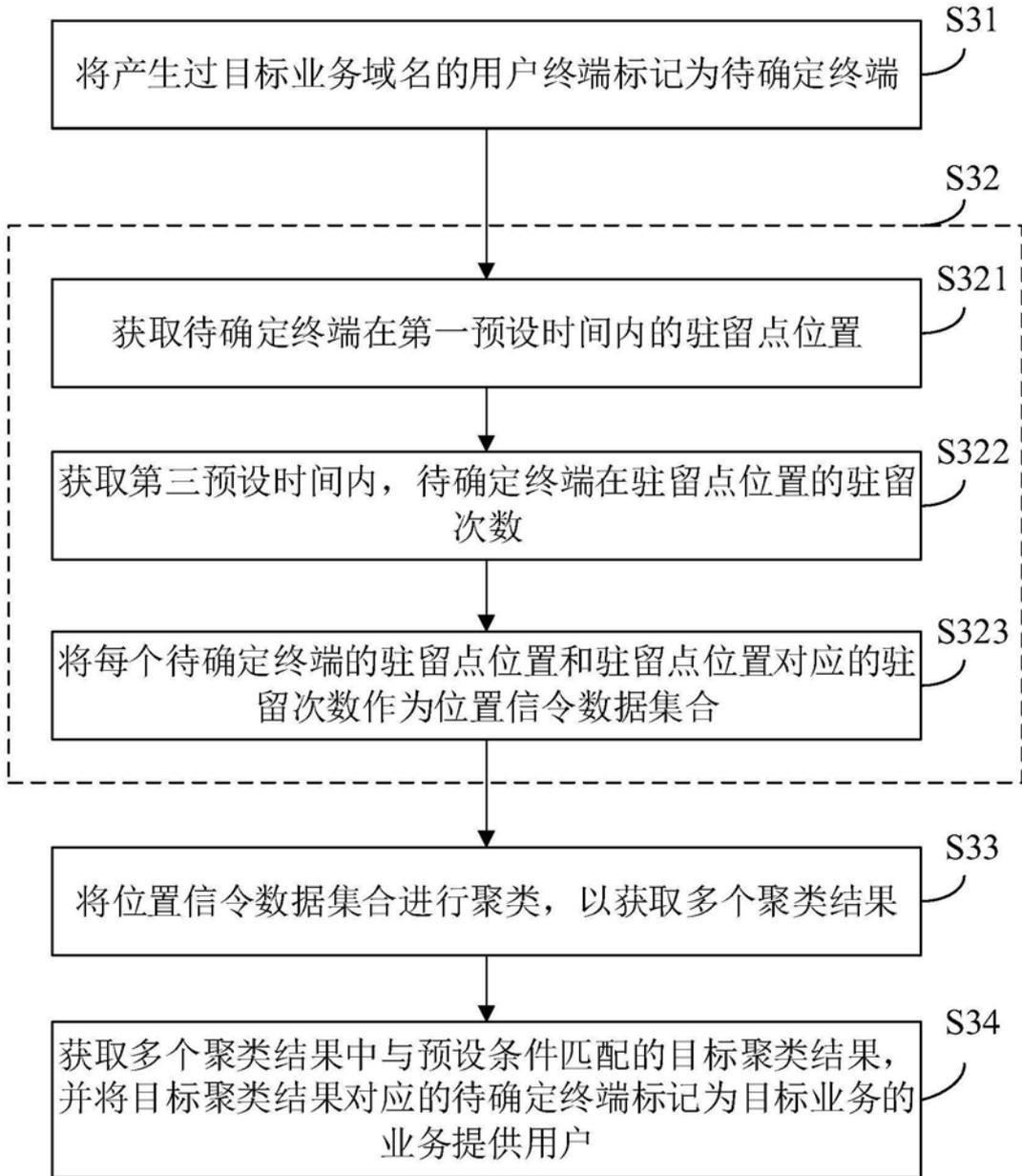


图4

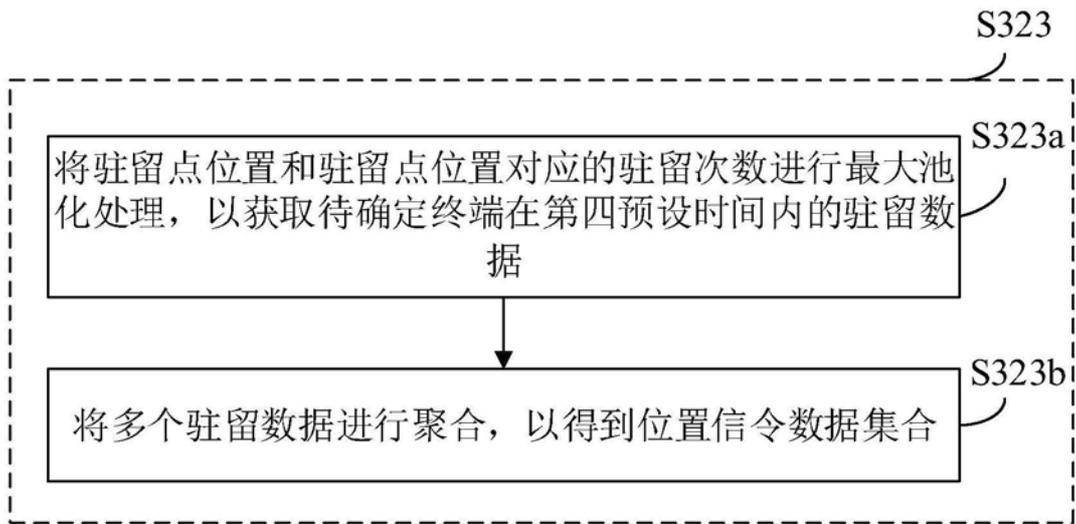


图5

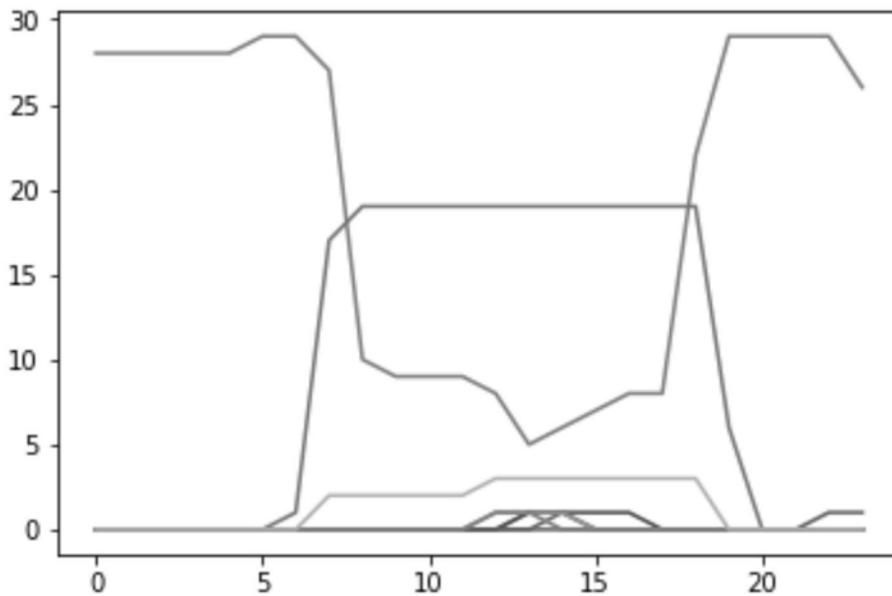


图6

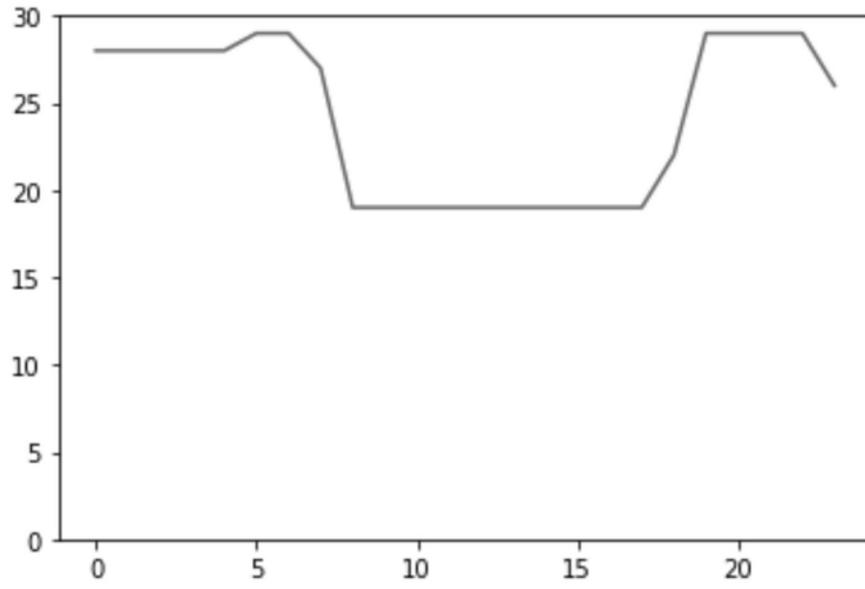


图7

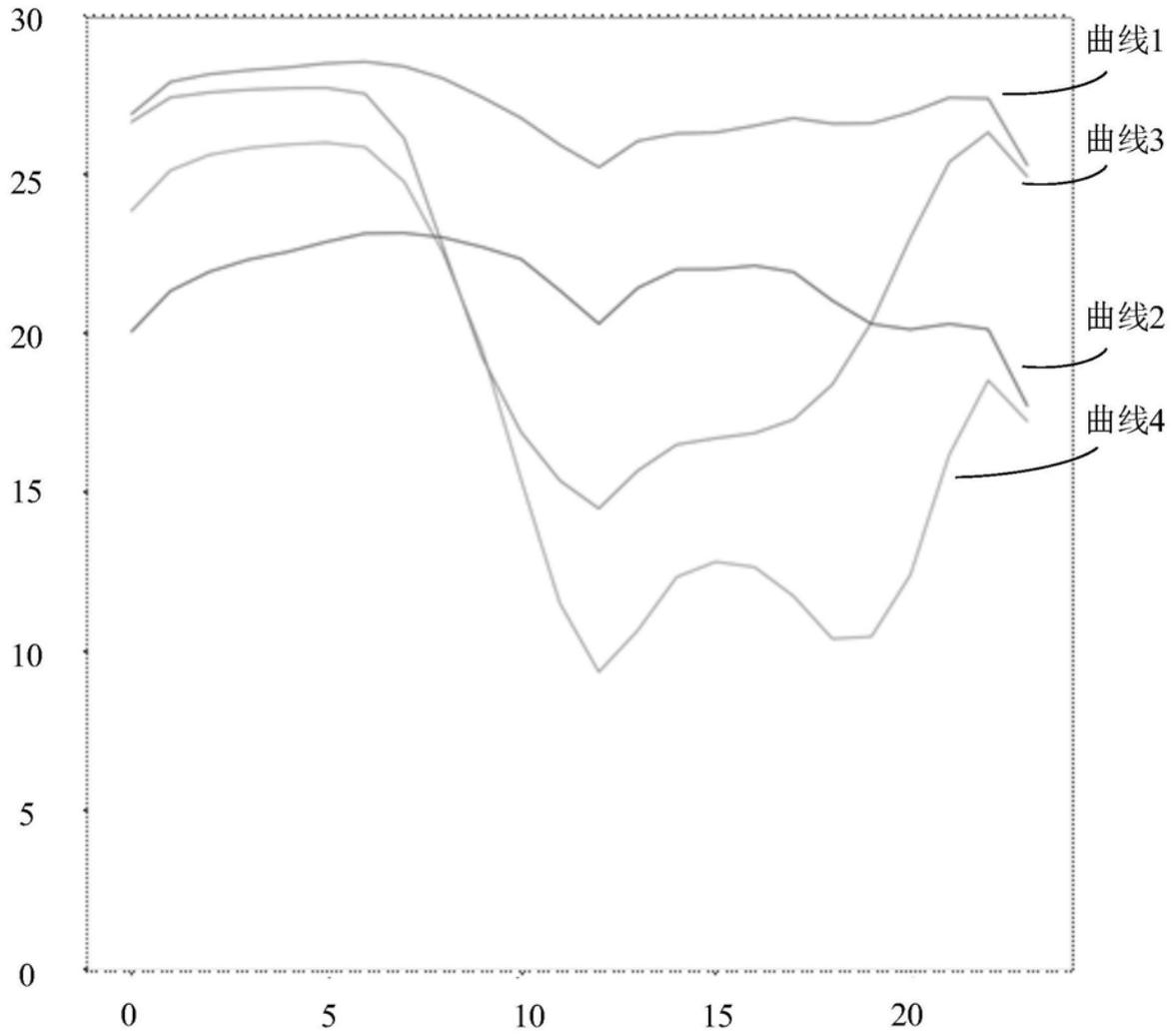


图8

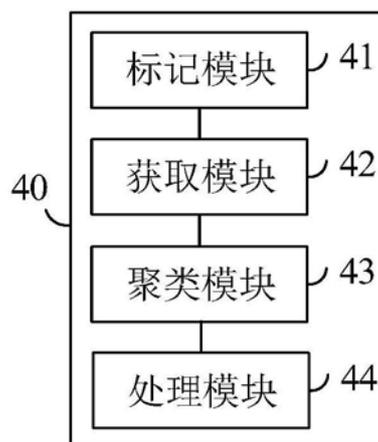


图9