



(12) 发明专利申请

(10) 申请公布号 CN 102789508 A

(43) 申请公布日 2012. 11. 21

(21) 申请号 201210262796. 9

(22) 申请日 2012. 07. 27

(71) 申请人 吴建辉

地址 100094 北京市海淀区西北旺镇永丰屯村及第公寓 003 室

(72) 发明人 吴建辉

(51) Int. Cl.

G06F 17/30 (2006. 01)

H04L 29/08 (2006. 01)

H04L 12/58 (2006. 01)

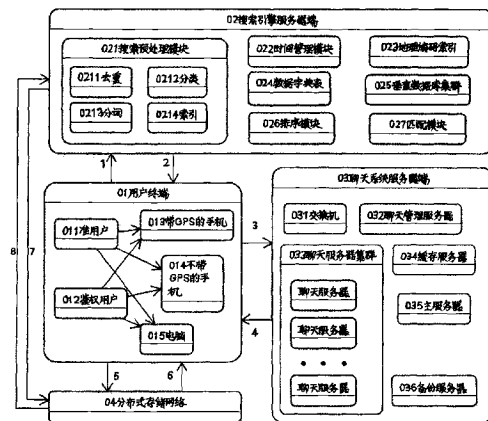
权利要求书 2 页 说明书 6 页 附图 1 页

(54) 发明名称

基于地理位置的分布式实况搜索引擎及聊天系统

(57) 摘要

一种基于地理位置的实况搜索引擎及聊天系统,包括用户终端,搜索引擎服务器端,聊天系统服务器端和分布式存储网络。当用户在用户终端输入想要搜索的信息后,这些信息被发送到搜索引擎服务器端,提取出关键词,并且用户的地理位置信息,用户 ID, IP 地址,端口号和签名后的关键词被存入垂直数据库集群,经过匹配,数据库选出和这些关键词相关联的和用户地理位置最近的一些用户的用户 ID, IP 地址,端口号等信息,返回用户终端,用户根据这些信息连接聊天系统服务器端,和这些用户聊天交流,当用户在用户终端输入想要搜索的信息的同时,这些信息也被发送到分布式网络进行备份存储。



1. 一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述的基于地理位置的分布式实况搜索引擎及聊天系统的构架包括用户终端,搜索引擎服务器端,聊天系统服务器端,无论搜索引擎服务器端还是聊天系统服务器端的用户的所有数据都要进行分布式存储,当搜索人输入想要搜索的信息进行搜索时,所述的搜索引擎搜索到的结果是一定地理范围内对这些信息也感兴趣的用户以及这些用户发布的含有这些信息的关键词的信息,点击这些用户中的任意一个用户的头像,就可以和这个用户进行聊天交流。

2. 根据权利要求1所述的一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述的实况搜索引擎及聊天系统的用户终端可以是带有GPS的手机,或者是不带GPS的手机,或者是电脑,其中只有带有GPS的手机能够提供搜索人所在的地理位置信息,这些地理位置信息可以是GPS采集到的经纬度,也可以是GPS采集到的经纬度编码后的数据。

3. 根据权利要求1或2所述的一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述的实况搜索引擎及聊天系统的搜索引擎服务器端包括搜索预处理模块,时间管理模块,垂直数据库集群,匹配模块和排序模块,其中搜索预处理模块由去重,分类,分词,索引四个功能模块组成,当搜索人输入想要搜索的信息后,这些信息被发送到搜索引擎服务器端,搜索预处理模块根据数据字典表对这些信息进行去重,分类,分词,索引等预处理后,分离出关键词,并且把这些关键词连同搜索人的用户ID,用户终端的IP地址,端口号等数据存放在垂直数据库集群中,垂直数据库集群是分布式集群,可以根据地理编码索引快速查找垂直数据库集群中的表,然后匹配模块会从垂直数据库集群中查到和预处理后的关键词相关联的一些跟搜索人距离最近的用户的用户ID,用户终端的IP地址,端口号,再把这些查到的结果进行排序后,发送到用户终端。

4. 根据权利要求1或2或3所述的一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述的实况搜索引擎及聊天系统的聊天系统服务器端包括交换机,聊天服务器集群,聊天管理服务器,缓存服务器,主服务器和备份服务器,用户终端和聊天系统服务器端交互的数据可以是文字,语音,图片,视频,当用户A输入想要搜索的信息,并从搜索引擎服务器端获得用户B的用户ID,用户终端的IP地址,端口号后,可以根据这些信息连接聊天系统服务器端,聊天系统服务器端会根据用户A以及用户B的用户终端的IP地址,端口号等信息把聊天任务安排给聊天服务器集群中的一台服务器进行处理,这一台服务器完成用户A以及用户B的聊天过程的会话搭建,用户A以及用户B交互的数据暂时存放在缓存服务器上,然后把聊天记录备份到备份服务器中,用户A和用户B聊天的方式可以是私聊,也可以是群聊,即允许其他人加入,如果是用户A发起的群聊,聊天管理服务器会把请求发送给主服务器,然后主服务器会向用户A提供的想要群聊的用户群组发送广播包,接收到广播包的用户作出许可应答后,就可以加入到群聊中。

5. 根据权利要求1或2或3或4所述的一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述的实况搜索引擎及聊天系统的用户分为准用户以及鉴权用户,鉴权用户还可以细分为几个等级,准用户具备搜索及聊天权限,鉴权用户不仅仅能搜索关键词,还可以发布一些自己想被别人搜索到的信息,比如广告,自己组织的兴趣活动的描述,鉴权用户的权限值不同,能够发布的信息的数据类型(指文字,语音,图片,视频)及数据长度,数据数量不同。

6. 根据权利要求 1 或 2 或 3 或 4 或 5 所述的一种基于地理位置的分布式实况搜索引擎及聊天系统,其特征在于:所述实况搜索引擎及聊天系统不仅仅在搜索引擎服务器端和聊天系统服务器端把用户的数据存储到分布式数据库集群中,而且用户输入的搜索想要搜索的信息及发布的文字广告会被签名成 <key, value> 数据对,分布式存储在网络节点中,当用户想要搜索的信息发送到搜索引擎服务器端后经过搜索预处理提取出的关键词也会被签名成 <key, value> 数据对,分布式存储在网络节点中。

## 基于地理位置的分布式实况搜索引擎及聊天系统

### 技术领域

[0001] 本发明涉及一种移动互联网的搜索引擎及聊天系统,尤其涉及一种基于地理位置的分布式实况搜索引擎及聊天系统。

### 背景技术

[0002] 目前的搜索引擎是利用网络蜘蛛程序抓取网站的外部链接,然后建立起数据库存储这些从互联网提取的各个网站的以网页文字为主的信息,并能检索与用户查询条件相匹配的记录,按一定的排列顺序返回结果,这样的搜索引擎搜索的是网站,搜索到的内容专业性不够,时效性和地域性不强,缺乏互动。

[0003] 生活中,搜索人经常遇到的是寻找特定目标人群的问题,比如有房要出租的房东想要寻找打算租房的人,同时打算租房的人也在寻找有房要出租的房东,类似的关系还有要做家教的老师和要补课的学生等,对于租房或者家教这类有交易发生的事件,无论现实中还是互联网上都有中介机构为交易的双方提供对方的信息,并且从中收取中介费。

[0004] 生活中,搜索人与被搜索人还可能构成另一类关系,即志同道合者,比如计划周末约着出去旅游的人,想一起学习乐器的音乐爱好者等,这类关系的双方并不发生交易,双方找到对方的联系方式的途径很有限。

[0005] 目前还没有搜索引擎为以上两类关系中的搜索者提供寻找到适合自己的目标人群的解决方案,而且以上两类关系中的搜索者和被搜索者往往在很距离很近的地理范围内,这就要求搜索引擎对搜索的结果根据地理位置进行筛选。

### 发明内容

[0006] 本发明的目的是:提供一种实况搜索引擎及聊天系统,当用户在用户终端输入想要搜索的信息(即查询条件)后,搜到的结果是一定地理范围内对这个信息感兴趣的人,以及这些人发布的包含这个信息的关键词的信息,然后点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流。

[0007] 比如,搜索者输入“人大西门租房”,那么搜索到的结果是人大西门附近有房出租的用户以及同样也想在人大西门租房的用户,或者搜索到的结果仅仅是人大西门附近有房出租的用户,同时可以看到这些用户发布的信息,比如“出租人大西门附近苏州街地铁口小南庄主卧,2000/月”,然后用户点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流。

[0008] 本发明是这样实现的:一种基于地理位置的分布式实况搜索引擎及聊天系统,其构架包括用户终端,搜索引擎服务器端,聊天系统服务器端以及分布式存储网络,无论搜索引擎服务器端还是聊天系统服务器端的用户的所有数据都要存储到分布式数据库集群中。

[0009] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端的功能是,接受到从搜索者用户终端发送的查询条件之后,匹配出和查询条件相关的离搜索者最近的一些用户的信息,这些信息包括用户名,用户 ID,用户终端的 IP 地址,用户终端的端口号等,并且把这些

信息返回给搜索者的用户终端。

[0010] 跟传统的搜索引擎相比,所述的搜索引擎服务器端存储的不是网页的文字信息,而是一些跟用户相关的信息,这些信息主要包括两类:一类是用户的查询记录,当一个用户输入某个查询条件进行搜索的时候,他的搜索记录也被保存在搜索引擎服务器端的垂直数据库集群中,成为别人查询的数据源,另一类信息是用户发布的信息,比如某个餐馆发布的今日特价菜单,或者是周末组织的某个兴趣活动的消息。

[0011] 所述的实况搜索引擎及聊天系统的用户分为准用户以及鉴权用户,鉴权用户还可以细分为几个等级,准用户具备搜索及聊天权限,鉴权用户不仅仅能搜索信息,还可以发布一些自己想被别人搜索到的信息,比如广告,自己组织的兴趣活动的描述,鉴权用户的权限值不同,能够发布的信息的数据类型(指文字,语音,图片,视频)及数据长度,数据数量不同。

[0012] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端在匹配用户的查询条件的时候,要根据搜索者用户所在的地理位置筛选出离搜索者最近的一些用户的信息,这就要求用户终端提供地理位置信息。

[0013] 所述的实况搜索引擎及聊天系统的用户终端可以是带有 GPS 的手机,或者是不带 GPS 的手机,或者是电脑,其中只有带有 GPS 的手机能够提供搜索人所在的地理位置信息,这些地理位置信息可以是 GPS 采集到的经纬度数据,也可以是 GPS 采集到的经纬度编码后的数据。

[0014] 对于所述的不带 GPS 的手机,或者是电脑,用户终端设备无法获得准确经纬度,但是用户可以在用户终端输入自己所在的地理位置的名称,发送到搜索引擎服务器端,这些地理位置的名称也被作为是匹配的条件之一。

[0015] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端包括搜索预处理模块,时间管理模块,垂直数据库集群,匹配模块和排序模块,其中搜索预处理模块包括去重,分类,分词,索引四个功能模块。

[0016] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端的工作流程是:当搜索人在用户终端输入要搜索的查询条件后,查询条件被发送到搜索引擎服务器端,搜索预处理模块根据数据字典表对查询条件进行去重,分类,分词,索引等预处理后,提取出关键词,把这些关键词存放在垂直数据库集群中,然后匹配模块会根据地理位置信息从垂直数据库集群中查到和预处理后的关键词相关的信息,这些信息包括用户 ID,用户终端的 IP 地址,端口号,然后排序模块对匹配后的结果进行排序,再把排序后的结果发送到用户终端。

[0017] 所述的搜索引擎服务器端的时间管理模块的主要功能是把时间划分成一些时间片,根据时间片更新垂直数据库集群。

[0018] 所述的搜索引擎服务器端把搜索到的用户 ID,用户终端的 IP 地址,端口号返回给搜索者的用户终端后,用户终端并不显示这些信息,搜索者只能从 UI 界面上看到搜索到的用户的头像和该用户发布的包含有搜索者输入的查询条件中的关键词的信息,然后点击搜索到的用户的头像按钮,搜索者根据搜索到的用户 ID,用户终端的 IP 地址,端口号连接聊天系统服务器端,并建立和搜索到的用户的会话。

[0019] 所述的实况搜索引擎及聊天系统,用户终端与聊天系统服务器端交互的信息包括文字,图片,语音,视频等多种格式。

[0020] 所述的实况搜索引擎及聊天系统,用户可以点击一个搜索到的用户的头像进行私聊,也可以点击几个搜索到的用户进行群聊,或者是创建讨论组。

[0021] 所述的实况搜索引擎及聊天系统的聊天系统服务器端包括交换机,聊天服务器集群,聊天管理服务器,缓存服务器,主服务器和备份服务器。

[0022] 所述的实况搜索引擎及聊天系统的聊天系统服务器端的工作流程是:当用户 A 输入想要搜索的信息,并从搜索引擎服务器端获得用户 B 的用户 ID,用户终端的 IP 地址,端口号后,可以根据这些信息连接聊天系统服务器端,聊天系统服务器端会根据用户 A 以及用户 B 的用户终端的 IP 地址,端口号等信息把聊天任务安排给聊天服务器集群中的一台服务器进行处理,这一台服务器完成用户 A 以及用户 B 的聊天过程的会话搭建,用户 A 以及用户 B 交互的数据暂时存放在缓存服务器上,然后会把聊天记录备份到备份服务器中。

[0023] 所述的实况搜索引擎及聊天系统,用户 A 和用户 B 聊天的方式可以是私聊,也可以是群聊,即允许其他人加入,如果是用户 A 发起的群聊,聊天管理服务器会把请求发送给主服务器,然后主服务器会向用户 A 提供的想要群聊的用户群组发送广播包,接受到广播包的用户作出许可应答后,就可以加入到群聊中。

[0024] 所述的实况搜索引擎及聊天系统,用户输入的查询条件会被签名成 <key, value> 数据对,存储在分布式网络节点中,其中的 key 值是查询条件经过签名运算得到的结果,value 就是查询条件。

[0025] 在所述的实况搜索引擎及聊天系统的搜索引擎服务器端,搜索人的查询条件经过搜索预处理后分解出关键词,对各个关键词进行签名运算,计算得到的一个 key 值,原来的关键词为 value,这样的每一对 <key, value> 值也存储在分布式网络节点中。

[0026] 所述的分布式网络存储的意义是节省了搜索引擎服务器端备份数据所需要的服务器集群,节约了成本,一个网络节点中存储的 <key, value> 数据对中的 key 值必须是和本地路由表中的 nodeID 非常接近。

[0027] 由于采用了本发明的结构,当用户在用户终端输入想要搜索的信息(即查询条件)后,搜到的结果是一定地理范围内对这个信息感兴趣的人,以及这些人发布的包含这个信息的关键词的信息,然后点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流,这样搜索到的信息地域性强,时效性强,增加了互动性,也免去了中介费,降低了成本。

## 附图说明

[0028] 附图 1 是本发明的系统构架图。

## 具体实施方式

[0029] 下面结合附图及实施例对本发明作进一步说明。

[0030] 如附图 1 所示,提供一种实况搜索引擎及聊天系统,当用户在用户终端 01 输入想要搜索的信息(即查询条件)后,搜到的结果是一定地理范围内对这个信息感兴趣的人,以及这些人发布的包含这个信息的关键词的信息,然后点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流。

[0031] 比如,搜索者输入“人大西门租房”,那么搜索到的结果是人大西门附近有房出租

的用户以及同样也想在人大西门租房的用户,或者搜索到的结果仅仅是人大西门附近有房出租的用户,同时可以看到这些用户发布的信息,比如“出租人大西门附近苏州街地铁口小南庄主卧,2000/月”,然后用户点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流。

[0032] 本发明是这样实现的:一种基于地理位置的分布式实况搜索引擎及聊天系统,其构架包括用户终端 01,搜索引擎服务器端 02,聊天系统服务器端 03 以及分布式存储网络 04,无论搜索引擎服务器端 02 还是聊天系统服务器端 03 的用户的所有数据都要存储到分布式数据库集群中。

[0033] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端 02 的功能是,接受到从搜索者用户终端发送的查询条件之后,匹配出和查询条件相关的离搜索者最近的一些用户的信息,这些信息包括用户名,用户 ID,用户终端的 IP 地址,用户终端的端口号等,并且把这些信息返回给搜索者的用户终端 01。

[0034] 跟传统的搜索引擎相比,所述的搜索引擎服务器端 02 存储的不是网页的文字信息,而是一些跟用户相关的信息,这些信息主要包括两类:一类是用户的查询记录,当一个用户输入某个查询条件进行搜索的时候,他的搜索记录也被保存在搜索引擎服务器端 02 的垂直数据库集群 025 中,成为别人查询的数据源,另一类信息是用户发布的信息,比如某个餐馆发布的今日特价菜单,或者是周末组织的某个兴趣活动的消息。

[0035] 所述的实况搜索引擎及聊天系统的用户分为准用户 011 以及鉴权用户 012,鉴权用户 012 还可以细分为几个等级,准用户 011 具备搜索及聊天权限,鉴权用户 012 不仅仅能搜索信息,还可以发布一些自己想被别人搜索到的信息,比如广告,自己组织的兴趣活动的描述,鉴权用户 012 的权限值不同,能够发布的信息的数据类型(指文字,语音,图片,视频)及数据长度,数据数量不同。

[0036] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端 02 在匹配用户的查询条件的时候,要根据搜索者用户所在的地理位置筛选出离搜索者最近的一些用户的信息,这就要求用户终端提供地理位置信息。

[0037] 所述的实况搜索引擎及聊天系统的用户终端 01 可以是带有 GPS 的手机 013,或者是不带 GPS 的手机 014,或者是电脑 015,其中只有带有 GPS 的手机 013 能够提供搜索人所在的地理位置信息,这些地理位置信息可以是 GPS 采集到的经纬度数据,也可以是 GPS 采集到的经纬度编码后的数据。

[0038] 对于所述的不带 GPS 的手机 014,或者是电脑 015,用户终端 01 设备无法获得准确经纬度,但是用户可以在用户终端 01 输入自己所在的地理位置的名称,发送到搜索引擎服务器端 02,这些地理位置的名称也被作为是匹配的条件之一。

[0039] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端 02 包括搜索预处理模块 021,时间管理模块 022,垂直数据库集群 025,匹配模块 027 和排序模块 026,其中搜索预处理模块 021 包括去重 0211,分类 0212,分词 0213,索引 0214 四个功能模块。

[0040] 所述的实况搜索引擎及聊天系统的搜索引擎服务器端 02 的工作流程是:当搜索人在用户终端 01 输入要搜索的查询条件后,查询条件被发送到搜索引擎服务器端 02,搜索预处理模块 021 根据数据字典表 024 对查询条件进行去重 0211,分类 0212,分词 0213,索引 0214 等预处理后,提取出关键词,把这些关键词存放在垂直数据库集群 025 中,然后匹配

模块 027 会根据地理位置信息从垂直数据库集群 025 中查到和预处理后的关键词相关的信息, 这些信息包括用户 ID, 用户终端的 IP 地址, 端口号, 然后排序模块 026 对匹配后的结果进行排序, 再把排序后的结果发送到用户终端。

[0041] 所述的搜索引擎服务器端 02 的时间管理模块 022 的主要功能是把时间划分成一些时间片, 根据时间片更新垂直数据库集群 025。

[0042] 所述的搜索引擎服务器端 02 把搜索到的用户 ID, 用户终端的 IP 地址, 端口号返回给搜索人的用户终端 01 后, 用户终端 01 并不显示这些信息, 搜索者只能从 UI 界面上看到搜索到的用户的头像和该用户发布的包含有搜索者输入的查询条件中的关键词的信息, 然后点击搜索到的用户的头像按钮, 搜索者根据搜索到的用户 ID, 用户终端的 IP 地址, 端口号连接聊天系统服务器端 03, 并建立和搜索到的用户的会话。

[0043] 所述的实况搜索引擎及聊天系统, 用户终端 01 与聊天系统服务器 03 端交互的信息包括文字, 图片, 语音, 视频等多种格式。

[0044] 所述的实况搜索引擎及聊天系统, 用户可以点击一个搜索到的用户的头像进行私聊, 也可以点击几个搜索到的用户进行群聊, 或者是创建讨论组。

[0045] 所述的实况搜索引擎及聊天系统的聊天系统服务器端 03 包括交换机 031, 聊天服务器集群 033, 聊天管理服务器 032, 缓存服务器 034, 主服务器 035 和备份服务器 036。

[0046] 所述的实况搜索引擎及聊天系统的聊天系统服务器端 03 的工作流程是: 当用户 A 输入想要搜索的信息, 并从搜索引擎服务器端 02 获得用户 B 的用户 ID, 用户终端的 IP 地址, 端口号后, 可以根据这些信息连接聊天系统服务器端 03, 聊天系统服务器端 03 会根据用户 A 以及用户 B 的用户终端的 IP 地址, 端口号等信息把聊天任务安排给聊天服务器集群 033 中的一台服务器进行处理, 这一台服务器完成用户 A 以及用户 B 的聊天过程的会话搭建, 用户 A 以及用户 B 交互的数据暂时存放在缓存服务器 034 上, 然后会把聊天记录备份到备份服务器 036 中。

[0047] 所述的实况搜索引擎及聊天系统, 用户 A 和用户 B 聊天的方式可以是私聊, 也可以是群聊, 即允许其他人加入, 如果是用户 A 发起的群聊, 聊天管理服务器 032 会把请求发送给主服务器 035, 然后主服务器 035 会向用户 A 提供的想要群聊的用户群组发送广播包, 接收到广播包的用户作出许可应答后, 就可以加入到群聊中。

[0048] 所述的实况搜索引擎及聊天系统, 用户输入的查询条件会被签名成 <key, value> 数据对, 存储在分布式存储网络 04 的网络节点中, 其中的 key 值是查询条件经过签名运算得到的结果, value 就是查询条件。

[0049] 在所述的实况搜索引擎及聊天系统的搜索引擎服务器端 02, 搜索人的查询条件经过搜索预处理后分解出关键词, 对各个关键词进行签名运算, 计算得到的一个 key 值, 原来的关键词为 value, 这样的每一对 <key, value> 值也存储在分布式网络存储网络 04 的网络节点中。

[0050] 所述的分布式存储网络 04 的意义是节省了搜索引擎服务器端 02 备份数据所需要的服务器集群, 节约了成本, 一个网络节点中存储的 <key, value> 数据对中的 key 值必须是和本地路由表中的 nodeID 非常接近。

[0051] 由于采用了本发明的结构, 当用户在用户终端 01 输入想要搜索的信息 (即查询条件) 后, 搜到的结果是一定地理范围内对这个信息感兴趣的人, 以及这些人发布的包含这



个信息的关键词的信息,然后点击其中一个搜索到的人的头像按钮,就可以和这个人进行聊天交流,这样搜索到的信息地域性强,时效性强,增加了互动性,也免去了中介费,降低了成本。

[0052] 应当指出,在不脱离本发明原理的情况下,本领域的技术人员可以对本发明的搜索引擎服务器端 02,聊天系统服务器端 03,分布式存储网络 04 以及用户终端 01 进行各种改动和变型,这些改动和变型,也被视为本发明的保护范围。

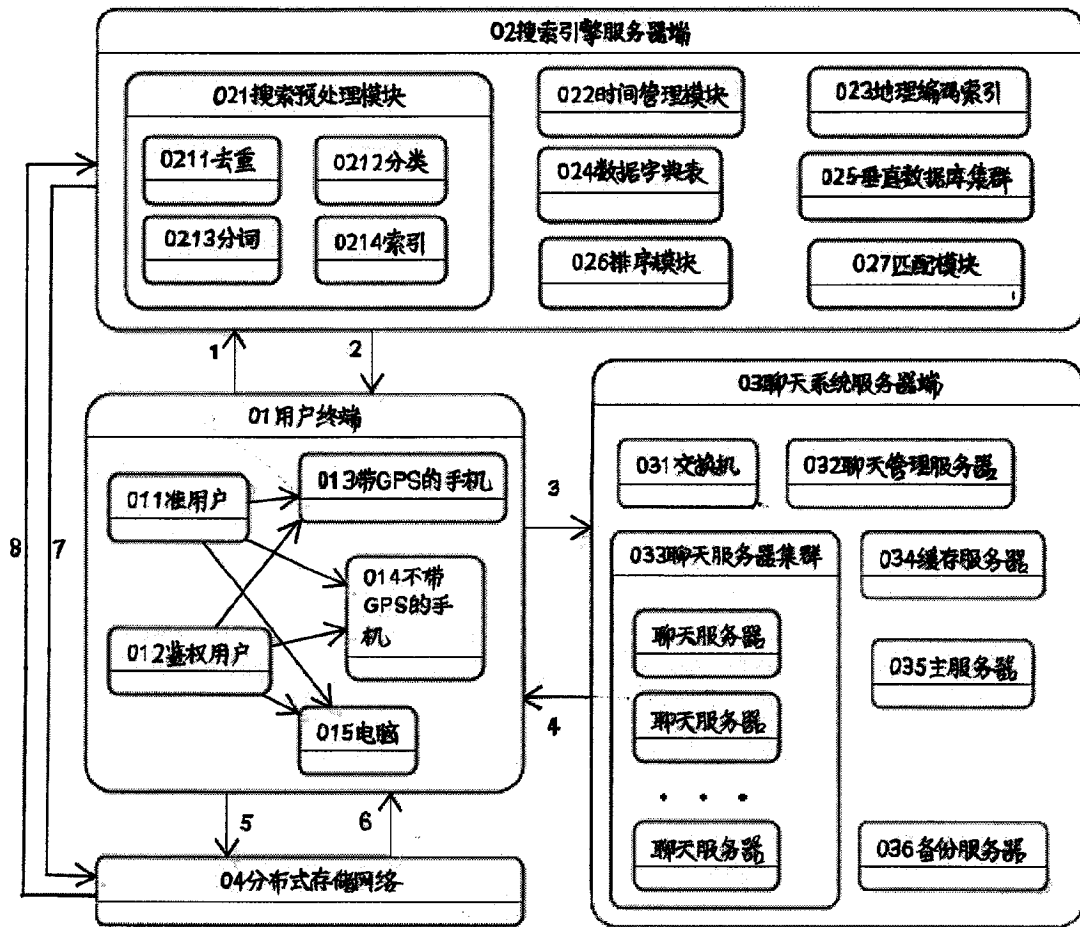


图 1