



(12) 发明专利

(10) 授权公告号 CN 109388649 B

(45) 授权公告日 2022. 05. 13

(21) 申请号 201811121360.1

G06Q 50/26 (2012.01)

(22) 申请日 2018.09.28

审查员 夏容

(65) 同一申请的已公布的文献号  
申请公布号 CN 109388649 A

(43) 申请公布日 2019.02.26

(73) 专利权人 土流集团有限公司  
地址 410000 湖南省长沙市天心区芙蓉南路新时空1号401房

(72) 发明人 欧阳琦 刘珍珍 王祥飞

(74) 专利代理机构 长沙市和协专利代理事务所  
(普通合伙) 43115  
专利代理师 熊晓妹

(51) Int. Cl.

G06F 16/242 (2019.01)

G06F 16/2458 (2019.01)

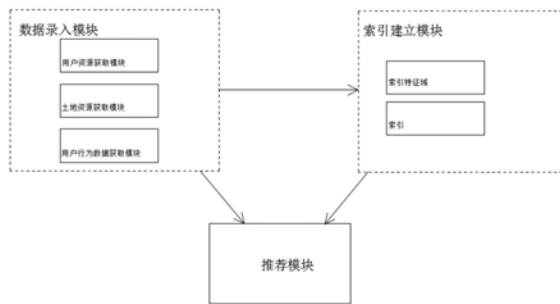
权利要求书2页 说明书4页 附图1页

(54) 发明名称

一种土地智能推荐方法及系统

(57) 摘要

本发明提供一种土地智能推荐方法及系统,该系统包括数据录入模块、索引建立模块以及推荐模块。本发明还提供了根据该系统实现土地智能推荐的方法。本发明提供的一种土地智能推荐方法及系统,能够对可利用开发的土地资源进行检索,适用于建立大型土地资源检索数据库,还可以利用该系统分析用户行为的特点为其进行土地资源的智能化推荐,操作方便、实用性强。



1. 一种土地智能推荐方法,其特征在于,包括以下步骤:

步骤1,获取基础数据,包括用户基础数据、土地基础数据以及用户行为数据的获取;

步骤2,对目标用户的搜索行为数据进行分析处理,得到所有的搜索分词项;

步骤3,将所有分词项映射到土地所有标签属性并进行加权计算得到每个标签的加权值;

步骤4,利用Lucene对所有土地资源根据标签以及权重建立索引库;

步骤5,分析单个活跃用户最近的搜索行为数据并分析得到所有分词项对应的标签属性;

步骤6,根据步骤5所得的标签属性为参数,通过Lucene检索出土地资源数据并利用Lucene内部机制打分排序,选择分数最高的土地资源;

步骤7,根据步骤6所得的推荐土地资源通过第三方短信平台发送至用户手机上;

所述步骤2中搜索行为数据进行分析处理包括分词处理、活跃用户统计以及次活跃用户统计,其中:

分词处理,使用开源的搜索引擎构建工具Lucene,对所有的搜索词汇进行分词处理,得到所有的分词项,并过滤掉单个字,以及无意义的词汇;

活跃用户统计,基于存储于关系型数据库中的登录日志,通过结构化的查询语句SQL统计一个月内至少2次登陆过平台的需求用户群体并查处所有的联系电话;

次活跃用户统计,基于存储于关系型数据库的登录日志,通过结构化的查询语句SQL统计半年内至少5次登陆过平台的需求用户群体,需要排除活跃用户;

所述步骤3中标签的加权值计算,包括如下步骤:

步骤一:计算分词项的总数totalNumber;

步骤二:将所有的分词项通过文字匹配,或含义匹配映射到土地高精属性值;

步骤三:计算可映射的分词项的重复次数,记为repeats;

步骤四:计算分词项的频度 $frequency = (repeats / totalNumber) * 100$ ;

步骤五:分词项的频度为对应高精属性值的权重值并存储,数据将进行周期性的更新,因为用户的行为在不停的增长,属性值集合的内容及其权重值也在不停的变化。

2. 根据权利要求1所述的一种土地智能推荐方法,其特征在于,为实现所述土地智能推荐方法采用的系统包括数据录入模块、索引建立模块以及推荐模块,其中数据录入模块包括:

用户资源获取模块,用户资源包括以下两类:通过手机号注册的需求用户,是智能推荐系统的短信推送渠道的目标用户;通过下载、安装平台手机APP应用并注册的需求用户,是智能推荐系统短信推送渠道以及消息通知推送渠道的目标用户;

土地资源获取模块,用于收集门户网站、手机应用或服务中心平台注册的用户发布的待流转土地;

用户行为数据获取模块,用于获取至少包括需求用户在平台门户网站或手机APP上搜索土地资源的检索词汇以及用户的土地资源浏览轨迹等数据;

索引建立模块包括:用户行为数据分析模块,用于包括分词处理、活跃用户统计以及次活跃用户统计;

标签权重计算模块,用于处理用户行为数据分析模块所得到的所有分词项;

索引构建查询模块,用于利用开源的搜索引擎Lucene对所有土地资源即待推荐资源构建索引库;

推荐模块,用于针对用户进行土地资源的推荐。

## 一种土地智能推荐方法及系统

### 技术领域

[0001] 本发明属于网络技术领域,具体地涉及一种土地智能推荐方法及系统。

### 背景技术

[0002] 随着国家土地政策改革,土地确权工作的展开,全国范围内可流通于市场的土地资源越来越多,流通速度越来越快。面对海量的可流转土地资源,有需求的企业家生产家在土地相关资源平台上往往需要花费大量的时间和精力才能找到符合需求的土地资源,这直接阻碍了土地资源的流转,严重影响了经济效益的产生速度,不符合国家发展战略。

[0003] 同时,土地资源平台现有的土地资源检索推荐方案效率低下,限制较多,无法快速响应用户需求,严重影响用户体验,降低用户粘性,不但不利于新用户的产生,更容易导致已有用户的流失。

[0004] 现有技术中进行资源推荐的平台技术开发大多都是基于 MySQL、Oracle、MSSQL 等关系型数据库管理系统,通过领域建模,以结构化的二维表和行记录的方式来存储资源数据。利用标准的结构化查询语句 (Structured Query Language, SQL) 对资源数据进行增加、删除、修改以及查询等操作。通过对二维表关键属性列添加索引在一定程度上提高性能。存在的问题是当检索条件变得复杂时,SQL 的查询效率将直线降低,而当数据库数据增加到百万千万级别时,SQL 语句的检索速率更是成几何级数下降,而且通过 SQL 来进行数据维护成本将成倍增长。

[0005] 现有技术方案的实施方式是通过用户在资源平台提供的信息检索界面主动录入有限的信息检索条件,平台利用用户录入的条件与数据库表中相关属性列进行数据相似性匹配,然后被动的将最终查询结果反馈给平台用户。

[0006] 因此,有必要开发一种轻量级、智能化的土地智能推荐方法及系统。

### 发明内容

[0007] 本发明的目的在于提供一种土地智能推荐系统,该系统包括数据录入模块、索引建立模块以及推荐模块,其中数据录入模块包括:

[0008] 用户资源获取模块,用户资源包括以下两类:通过手机号注册的需求用户,是智能推荐系统的短信推送渠道的目标用户;通过下载、安装平台手机 APP 应用并注册的需求用户,是智能推荐系统短信推送渠道以及消息通知推送渠道的目标用户;

[0009] 土地资源获取模块,用于收集门户网站、手机应用或服务中心平台注册的用户发布的待流转土地;

[0010] 用户行为数据获取模块,用于获取至少包括需求用户在平台门户网站或手机 APP 上搜索土地资源的检索词汇以及用户的土地资源浏览轨迹等数据;

[0011] 索引建立模块包括:用户行为数据分析模块,用于包括分词处理、活跃用户统计以及次活跃用户统计;

[0012] 标签权重计算模块,用于处理用户行为数据分析模块所得到的所有分词项;

[0013] 索引构建查询模块,用于利用开源的搜索引擎Lucene对所有土地资源即待推荐资源构建索引库;

[0014] 推荐模块,用于针对用户进行土地资源的推荐。

[0015] 本发明另一个目的是采用上述系统提供一种土地智能推荐方法,包括以下步骤:步骤1,获取基础数据,包括用户基础数据、土地基础数据以及用户行为数据的获取;

[0016] 步骤2,对目标用户的搜索行为数据进行分析处理,得到所有的搜索分词项;

[0017] 步骤3,将所有分词项映射到土地所有标签属性并进行加权计算得到每个标签的加权值;

[0018] 步骤4,利用Lucene对所有土地资源根据标签以及权重建立索引库;

[0019] 步骤5,分析单个活跃用户最近的搜索行为数据并分析得到所有分词项对应的标签属性;

[0020] 步骤6,根据步骤5所得的标签属性为参数,通过Lucene检索出土地资源数据并利用Lucene内部机制打分排序,选择分数最高的土地资源;

[0021] 步骤7,根据步骤6所得的推荐土地资源通过第三方短信平台发送至用户手机上。

[0022] 优选的,所述步骤3中标签的加权值计算,包括如下步骤:

[0023] 步骤一:计算分词项的总数totalNumber,例如我们得到了10000 个相同或不同的分词项;

[0024] 步骤二:将所有的分词项通过文字匹配,或含义匹配映射到土地高精属性值;

[0025] 步骤三:计算可映射的分词项的重复次数,即repeats,例如,农用地这一分词项在总数为10000的分词项合集中出现了50次,即得到repeats=50;

[0026] 步骤四:计算分词项的频度frequency=(repeats/totalNumber)\* 100;

[0027] 步骤五:分词项的频度即对应高精属性值的权重值并存储,数据将进行周期性的更新,因为用户的行为在不停的增长,属性值集合的内容及其权重值也在不停的变化。

[0028] 优选的,所述步骤2中搜索行为数据进行分析处理包括分词处理、活跃用户统计以及次活跃用户统计,其中:

[0029] 分词处理,使用开源的搜索引擎构建工具Lucene,对所有的搜索词汇进行分词处理,得到所有的分词项,并过滤掉单个字,以及无意义的词汇;

[0030] 活跃用户统计,基于存储于关系型数据库中的登录日志,通过结构化的查询语句SQL统计一个月内至少2次登陆过平台的需求用户群体并查处所有的联系电话;此为推荐系统目标用户之一,推荐的目的是帮助用户快速定位需求资源,促成土地流转,增加用户粘性,提高用户留存率;活跃用户将得到更多的推荐资源;

[0031] 次活跃用户统计,基于存储于关系型数据库的登录日志,通过结构化的查询语句SQL统计半年内至少5次登陆过平台的需求用户群体,需要排除活跃用户;此为推荐系统目标用户之一,推荐的目的是唤醒用户需求,帮助用户快速定位需求资源,促成土地流转。

[0032] 本发明的原理:通过分析大量用户的行为数据以及土地数据,得到一套标准的标签属性以及对应的权重值,再对土地数据打标,然后针对单个用户的行为轨迹得到标准标签属性的一个子集,最后通过该子集从Lucene索引库中查找最匹配的土地数据并打分排序,将最符合的数据推送给用户。

[0033] 与相关技术相比,本发明提供一种土地智能推荐方法及系统,能够对可利用开

发的土地资源进行检索,适用于建立大型土地资源检索数据库,还可以利用该系统分析用户行为的特点为其进行土地资源的智能化推荐,操作方便、实用性强。

## 附图说明

[0034] 图1为本发明提供的土地智能推荐系统的结构框图;

## 具体实施方式

[0035] 以下将参考附图并结合实施例来详细说明本发明。

[0036] 如图1所示,本实施例提供的土地智能推荐系统,该系统包括数据录入模块、索引建立模块以及推荐模块,首先通过录入模块获取基础数据,包括获取用户基础数据,即智能推荐系统目标用户,用户来源的方式有两种:第一种是在平台门户网站通过手机号注册的需求用户,是智能推荐系统的短信推送渠道的目标用户;第二种是通过下载、安装平台手机APP应用并注册的需求用户,是智能推荐系统短信推送渠道以及消息通知推送渠道的目标用户。其中,短信推送渠道和消息通知推送渠道皆为系统通过网络调用第三方平台API接口来将推荐资源送达到目标用户的手机上并呈现在手机短信或系统通知栏目中。

[0037] 获取土地基础数据,包括土地数据,土地分类数据以及土地分类对应的高精属性数据。

[0038] 获取土地数据,即待推荐资源,来自于门户网站、手机应用或服务中心平台注册的“地主”用户发布的待流转土地,“地主”用户可以是真正的土地所有者,也可以是第三方代理商、中介机构或服务中心等。其中,土地数据除了基础信息(包括:位置、描述、联系人、面积价格等)外,还有很多额外的高精属性,高精属性主要用于描述土地资源自身的性质,属性信息越完善,将会有更大的概率被推荐并促成交易。

[0039] 获取土地分类数据,来自于产品经理,运营以及土地评估师对国家公布的相关文件整理得出。

[0040] 获取土地分类高精属性数据,来自于产品经理,运营以及土地评估师通过对国家公布的相关文件整理,实地考察,技术咨询等方式得来。

[0041] 获取用户行为数据,至少包括需求用户在平台门户网站或手机APP上搜索土地资源的检索词汇以及用户的土地资源浏览轨迹等数据。需求用户在门户网站或手机应用APP上的搜索栏,填写检索词汇发起搜索土地资源的行为,其中用户的检索词汇,请求的IP地址,请求时间等信息会被存储日志数据库中。同时,每个用户的登录时间,在平台上的访问轨迹等数据也会存储在日志数据库中,作为用户行为数据分析的基础。

[0042] 再通过索引建立模块建立索引文件,用户行为分析,包括分词处理,活跃用户统计以及次活跃用户统计。

[0043] 分词处理,使用开源的搜索引擎构建工具Lucene,对所有的检索词汇进行分词处理,得到全部的分词项,并过滤掉单个字,以及无意义的词汇。

[0044] 活跃用户统计,基于存储于关系型数据库中的登录日志,通过结构化的查询语句SQL统计一个月内至少2次登陆过平台的需求用户群体并查处所有的联系电话。此为推荐系统目标用户之一,推荐的目的是帮助用户快速定位需求资源,促成土地流转,增加用户粘性,提高用户留存率。活跃用户将得到更多的推荐资源。

[0045] 次活跃用户统计,基于存储于关系型数据库的登录日志,通过结构化的查询语句SQL统计半年内至少5次登陆过平台的需求用户群体,需要排除活跃用户。此为推荐系统目标用户之一,推荐的目的是唤醒用户需求,帮助用户快速定位需求资源,促成土地流转。

[0046] 标签权重计算,即进一步处理用户行为分析模块所得到的所有分词项。

[0047] 步骤一:计算分词项的总数totalNumber,例如我们得到了10000 个相同或不同的分词项;

[0048] 步骤二:将所有的分词项通过文字匹配,或含义匹配映射到土地高精属性值;

[0049] 步骤三:计算可映射的分词项的重复次数,即repeats,例如,农用地这一分此项在总数为10000的分词项合集中出现了50次,即得到repeats=50;

[0050] 步骤四:计算分词项的频度frequency=(repeats/totalNumber)\* 100;

[0051] 步骤五:分词项的频度即对应高精属性值的权重值并存储,次数据将进行周期性的更新,因为用户的行为在不停的增长,属性值集合的内容及其权重值也在不停的变化。

[0052] 构建索引库,利用开源的搜索引擎Lucene对所有土地资源,即待推荐资源构建索引库,每一条土地资源拥有的高精属性值如果出现权重表中,就对该属性值所对应的索引域加权,权重值即上一步得到的频度值。次索引库将周期性的更新,因为流转土地在不停的增长,索引也会随之增长,而且权重值也周期性的变化。

[0053] 单个(次)活跃用户行为分析,即单个目标用户,采用开源的搜索引擎构建工具Lucene,对该用户所有记录的搜索词汇进行分词处理,得到全部的分词项,并过滤掉单个字,以及无意义的词汇并去重。

[0054] 对单个用户的有效分词项,即上个步骤得到的分词集合通过文字匹配,含义匹配完成到高精属性值的替换。

[0055] 将完成替换的分词项集合作为Lucene搜索引擎查询工具的参数查处最匹配的土地资源数据并打分排序;

[0056] 将打分最高的前N个土地资源选出作为该需求用户的待推荐资源并存储,活跃用户存储8个,次活跃用户存储4个,N值可以自由设置。

[0057] 重复以上4个步骤得到所有目标用户的待推荐资源并存储;

[0058] 推荐模块,定位到数据库的土地资源推荐记录及其文件的存放位置进行用户的个性化推荐。将目标用户和对应的待推荐资源以每隔N 天发送一条的方式,作为参数通过调用第三方短信或消息推送平台的 API将待推荐资源发送到目标用户的手机上,以短信或通知消息的方式显示。

[0059] 当所有得目标用户的待推荐资源完成推送后,再重复以上步骤开始新一轮的智能推荐行为。

[0060] 以上所述仅为本发明的实施例,并非因此限制本发明的专利范围,凡是利用本发明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它相关的技术领域,均同理包括在本发明的专利保护范围内。

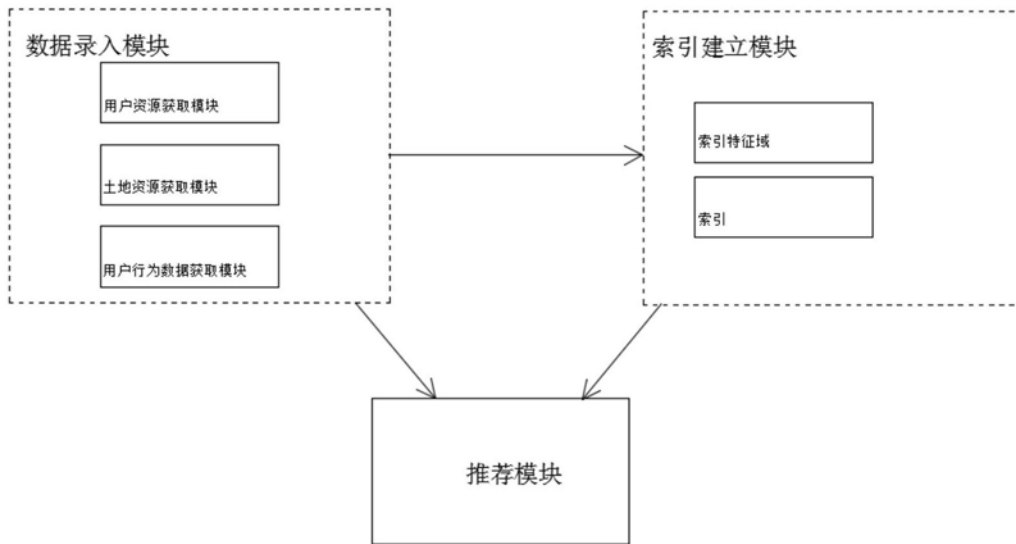


图1