

基于移动特征数据的内容推送技术研究与应用

吴明礼,杨双亮

(北方工业大学,北京 100144)

摘要:随着移动互联网的快速发展以及移动设备的普及,移动信息内容推送已经成为当前的热点之一。根据移动终端用户的行为习惯确定内容推送,提高内容推送的性能和用户满意度,已成为移动内容推送的主要任务之一。为此,在概括分析传统推送技术及自动推送技术优势的基础上,根据移动数据的动态特征,提出了一种基于移动特征数据变化的内容推送模型。该模型根据用户所处位置和时间变化,通过支持向量机(Support Vector Machine, SVM)算法来预测用户的类别标签偏好,并从内容库中选取相应内容,推送给用户,并通过实验的方式对基于移动特征数据变化的内容推送模型的推送效果进行了验证。实验验证结果表明,所提出的模型有效、可行且适用性好。

关键词:移动特征数据;支持向量机;内容库;用户画像;推送模型

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2017)09-0155-06

doi:10.3969/j.issn.1673-629X.2017.09.034

Research and Application on Content Push Technology with Mobile Feature Data

WU Ming-li, YANG Shuang-liang

(North China University of Technology, Beijing 100144, China)

Abstract: With the rapid development of mobile Internet and the popularity of mobile devices, mobile content delivery has become one of the hottest topics in the pushing field. How to push content for users based on the behaviors of the mobile users has become the main task of mobile content push to improve its performance and user satisfaction. On the basis of analysis on traditional push technology and advantages of automatic push technology as well as the dynamic characteristics of mobile data, a content push model based on the changes of mobile feature data has been presented, which can predict the user's category label preference by SVM and select the relevant content from the content library for commendation to the users according to the change of user's location and time. The experiments for its verification is conducted, which show that it is effective and feasible with good applicability.

Key words: mobile feature data; SVM; content library; user profile; push model

0 引言

随着移动互联网的迅速发展,移动互联用户的数量得到迅速增长。据 2016 年互联网数据中心(IDC)最新研究报告显示:2016 年全球互联网用户数将达到 32 亿,约占总人口的 44%,其中移动互联用户总数将达到 20 亿,再者由于移动终端设备的应用普及以及软硬件方面的不断改进,使得移动互联网成为人们获取信息的一个绝佳平台。用户可以根据需求在移动的过程中获取 Internet 的服务,例如天气预报、娱乐资讯、股票投资等最新的服务内容信息。但是随着信息量的不断暴增,使得用户手机经常收到一些对自己无用的内

容信息或者难以搜索到自己感兴趣的信息,严重影响了用户体验,致使大量用户的流失。如何有效地根据移动用户需求,自动为用户推送其感兴趣的内容,吸引新用户,留住老用户,成为亟待解决的重要问题。移动个性化服务是解决该问题的有效途径,其目标是创新服务模式,用智能、主动的、信息找人的推送模式,代替陈旧的、被动的、人找信息的搜索模式^[1]。

目前个性化服务研究领域用户需求获取技术的研究还处于探索阶段^[2],而移动用户需求获取技术方面的研究则更少。与传统网络相比,移动通讯网络有其自身特点:带宽有限,连接稳定性差,延迟长;与电脑相比,移动终端有屏幕相对较小,CPU 处理能力较低,电

收稿日期:2016-10-15

修回日期:2017-01-18

网络出版时间:2017-07-11

基金项目:北京市教育科技计划面上项目(KM201510009008)

作者简介:吴明礼(1978-),男,博士,讲师,研究方向为数据库技术和数据挖掘;杨双亮(1988-),男,硕士,研究方向为数据挖掘、推荐系统。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170711.1455.058.html>

池续航能力较短等不足;而且在移动网络环境下,用户周围的上下文复杂多变,对用户需求的影响更加明显^[3]。因此如何结合这些特点,实时、准确获取用户在不同上下文(例如位置、时间等)影响下用户的需求,推送用户感兴趣的内容,实现服务按需提供,内容按需推送,将有助于实现以用户为中心的服务模式,提高用户的满意度。

为此,在研究分析用户需求获取技术及推送技术的基础上,提出了基于移动特征数据变化的内容推送模型。该模型包含移动用户需求的获取、内容库的设计及内容的选择、内容自动推送等。实验证明了该模型的适用性和有效性。

1 相关技术及其原理

1.1 用户需求获取技术

用户需求获取技术是指在复杂、融合、协作、泛在的移动网络环境下,通过跟踪、学习用户的兴趣、偏好以及性格特征等信息,实时、准确地发现不同用户对各种移动网络服务的需求,并对其变化做出适应和调整^[4]。所涉及的用户需求是包含用户偏好、兴趣在内的,是广义上的业务需求。传统的用户需求获取技术一般有以下几种:

(1)TF-IDF 方法。Yeung 等^[5]将用户感兴趣的文档,通过分词,表示成关键词的向量,然后计算出用户对各个关键词的权重。

(2)聚类方法。常慧君等^[6]运用聚类的方法对用户行为进行聚类,获取各个人群的行为规律。

(3)决策树归纳。范琳等^[7]将用户偏好的获取过程表达成一棵决策树,用户从根节点开始,被引导来回答一系列问题。一旦达到叶子节点,可得到对用户偏好的完整描述。

(4)朴素贝叶斯分类。

随着时间的推移,用户需求会发生变化,如何监测这些变化并进行适应,对个性化服务具有重要价值。用户模型更新技术主要分为三类:信息增补技术^[5];神经网络技术:用户的兴趣偏好发生变化时,神经网络的连接权重也会跟着变化,最终导致不同的输出结果;如周朴雄等结合情景的用户偏好模型,运用 BP 神经网络方法来预测不同情景下的用户偏好^[8]。

1.2 推送技术

所谓的推送技术就是一种基于客户端/服务器端(C/S)的机制,服务器端主动将内容信息发送到客户端的技术。其优点在于内容信息发送的主动性和及时性,可随时将信息推送到用户面前(客户端)^[9]。那么如何使客户端能够接收到服务器端的内容,主要有两种方式:

(1)Pull(拉):它是客户端每过一定的时间间隔到服务器端获取消息,通过比较信息是否有更新,这是一种伪推送方式,因为并不是服务器主动向客户端发送内容消息,而是客户端每隔一定的间隔向服务器端发送一次查询来获取新的内容消息^[10]。

(2)Push(推送):服务器端有新信息后,就自动将最新的内容信息 Push 到客户端^[11]。

目前主要的推送技术有:简单轮询、SMS(Short Message Service,短信群发系统)、持久连接方式。而这些推送技术都有其不足之处,而且都是主动进行推送,当有新的内容需要推送时,才进行手动操作,没有考虑到用户需求。因此结合移动数据的特点,运用移动用户获取技术,获取移动用户的需求,然后从内容库中选取移动用户感兴趣的内容,并推送给用户。这种自动推送的方式既具有传统推送技术的优点,又考虑了用户需求,并且节省了大量的人力资源。例如郑小雪为了使政府部门快速且正确处理公众的诉求和意见,提出了基于知识地图的政府诉求文件自动推送模型,从而减少了人工分拣文件的作业,提高了政府的行政效率^[12]。

2 推送模型

2.1 系统架构

图 1 为整体的系统架构。

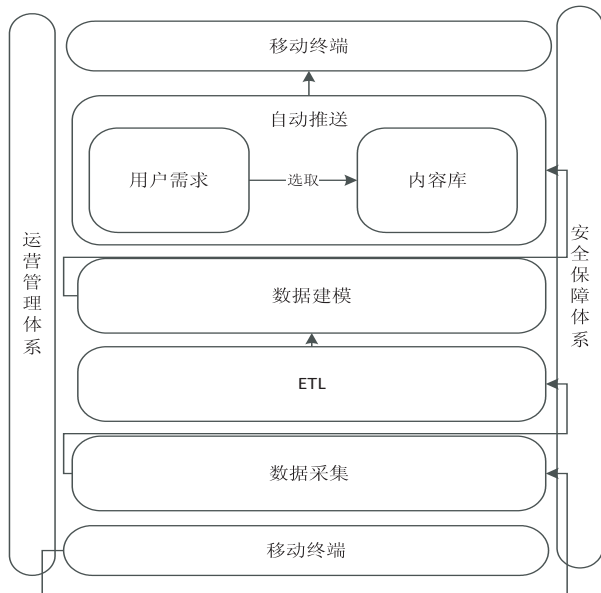


图 1 系统整体架构

首先通过移动终端的 APP 软件,获取移动终端用户的相关信息,包含静态信息和动态信息等,然后经过一系列的数据抽取(Extract)、转换(Transform)、加载(Load)(简称 ETL),用于上一层的数据建模,预测出用户的需求,从而自动从内容库中选择用户可能感兴趣的内容服务,推送给用户。

2.2 移动特征数据

采用的数据来源于智能终端的 APP 软件,它能够自动收集用户的一些显式信息和隐式反馈数据。相比于显式信息,隐式反馈数据的使用具有更多优势,其采集成本更低,对用户的干扰更小,拥有丰富的数据量^[13]。显式信息包含用户的性别和年龄,由用户注册,隐式信息包含用户的位置信息、使用手机的信息、使用手机 APP 软件的信息以及用户上网浏览的网址信息等,用户的主要信息分为两种:

(1)用户的静态信息:一般是指用户的信息不发生变化的,或者很长时间内不会发生变化的信息。例如用户的性别、年龄等信息,一般情况下是由用户自己注册的,因此这些信息的获取相比而言较为困难。

(2)用户的动态信息:移动用户在长期使用智能终端的基础上形成的行为习惯信息。

①用户的位置信息。
位置信息是由经度属性 LON 和纬度属性 LAT 以及当前的时间点两部分组成,即 $LOCATION = \langle U, LON, LAT, T \rangle$,其含义就是用户 U 在 T 时刻处于 $\langle LON, LAT \rangle$ 这个位置。历史轨迹数据集记录了用户在不同时间点上所处的地理位置点,是由一系列的连续位置序列组成^[14]。用户在不同的时间点,所处的位置可能不同,那么用户周围的位置语义也是不同的,用户在不同的位置、不同的时间点所需要的内容信息也不尽相同。

②用户使用手机信息。
移动通信可以通过无处不在的移动通信网络,在任何地点任何时间介入到网络中。只要有移动网络覆盖的区域就可以随时随地使用移动智能终端。通过获取智能终端开关屏的信息,来表示用户使用手机的情况,其使用智能终端的信息用 $\langle U, STATE, T \rangle$ 表示,其中 STATE 有两种状态:0 代表智能终端屏幕关闭,1 代表智能终端屏幕打开。那么用户使用智能终端一次的信息,就可以用两条记录表示,即 $\langle U, 1, T_1 \rangle, \langle U, 0, T_2 \rangle$,那么就可以用 $T_2 - T_1$ 表示该用户这次使用手机的时间长短,那么也可以通过这些记录来获取用户使用智能终端的频次。

根据人的日常行为规律进行离散化分析,将每天二十四小时分成不同的时间段(T 来代表),每个时间段的时长不等(见表 1),这种方式能够得到用户在各个时段的兴趣爱好。

针对时间信息,Yuan 等发现用户在某些给定时段里的行为具有一定的规律^[15],而且呈现出一定的峰段。在此基础上,研究了用户一天 24 小时内使用智能终端的情况,如图 2 所示。对连续的时间,采用离散化方式进行分析,研究其规律。

表 1 时间段划分

时间段	名称
00:00-05:59	T_1
06:00-06:59	T_2
07:00-08:59	T_3
09:00-11:59	T_4
12:00-12:59	T_5
13:00-17:59	T_6
18:00-19:59	T_7
20:00-23:59	T_8

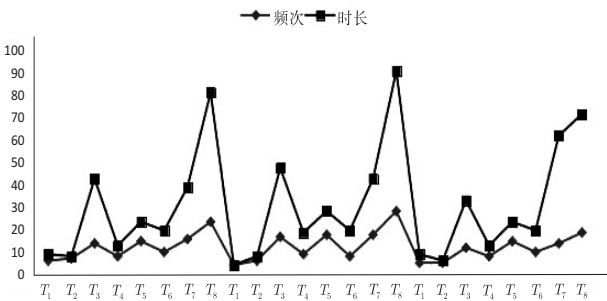


图 2 用户每天使用手机情况

无论是使用频次,还是使用时长,在 T_3 、 T_5 、 T_8 时段都会出现相应的峰值,在 T_7 时段也会出现较高的次数。从图中可以看出,这些也都符合人们的作息规律, T_3 、 T_7 为上下班时间段, T_5 为中午吃饭时间段, T_8 为晚上时间段,这些时间段用户有更多的时间使用手机。那么这些时间段对于内容提供者而言就是一种很好的即时推送时刻,如果能在用户经常使用智能终端的时间段里,为用户推送感兴趣的内容,那么将会起到事半功倍的效果。

③用户使用 APP 信息。
移动智能终端上,用户还安装了很多 APP 软件,用户的每次启动都会智能终端的操作系统上启动相应的进程,可以通过该进程来获取用户使用智能终端上 APP 的情况,以及相应 APP 软件每次使用所消耗的流量,可以通过 $\langle U, APPname, flow, T \rangle$ 来表示。其中 flow 代表用户该次使用所消耗的流量。通过记录用户所有使用 APP 的记录,可以知道用户对哪种 APP 软件更为感兴趣。

2.3 需求预测模型

用户信息影响用户自身标签的形成,也影响用户在特定情景下的意图,尤其移动端用户,因为用户可以随时随地地使用移动终端来搜索自己需要的内容。相应移动终端用户的因素可能会有很多,如位置、时间、天气等,而在这些因素中,时间、位置是最重要的因素,因此在情景信息主要考虑时间、位置这两个因素。

约定 1:情景信息 Context 用 C 表示, $C = \{Lon, Lat, T\}$,其中 T 表示预先约定好的时间段,Lon,Lat 分

别表示用的经纬度。不同的 T , 或者不同的经纬度, 都代表着不同的情景。那么 $U = \{C_1, C_2, \dots, C_n\}$ 表示用户的 N 个不同的情景信息。

约定 2: 用户的情景偏好用四元组表示为 $UCP = \{U, C, P, W\}$ 。其中, U 表示具体的用户, C 表示情景, P 表示某一特定领域本体中的类别标签, W 表示用户 U 在 C 情景下, 对于 P 标签的偏好的权重大小, 该值是通过该用户的历史数据计算出来的。

在四元组的基础上, 首先通过用户每天的行为数据, 计算出用户在各个情景下对于各个类别标签的权重即 W , 那么就组成了用户在不同时间, 不同地点, 对于不同的类别标签, 拥有不同的喜好程度。由于时间和位置两个特征属性的多样性, 使得对于多分类问题的求解较为困难。

首先将用户分为两类: 一类是有用户行为数据的用户; 另一类是没有任何行为数据的用户, 即新用户。对于不同的用户应当拥有不同的处理方式。

1) 有行为数据的用户。用户拥有相应历史行为数据, 首先根据用户的历史数据计算出每个用户对应场景下类别标签的权重大小, 即 W 。然后根据用户不断使用所产生的历史行为数据, 采用支持向量机模型 (Support Vector Machine, SVM) 预测用户在未来某个场景下用户的类别标签偏好。之所以选用 SVM, 是因为移动用户所处的位置特征和时间特征太多样化了, 而 SVM 能很好地处理多分类问题。

SVM 是一种新的统计学习算法, 其学习原则是使结构风险最小化, 通过二次规划问题需将数据分为两类的最佳超平面^[16]。核心内容为: 对于输入空间中的非线性可分问题, 选择适当的映射, 将映射空间中的样本点映射到一个高维特征空间, 使得样本点在该空间线性可分, 而且通过核函数使其计算在原空间进行, 降低映射高维特征空间计算的复杂性。SVM 算法步骤如下:

(1) 已知训练集 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中 $x_i \in X, y_i \in Y, i = 1, 2, \dots, N$ 。首先假设样本集 $z = \{(x_i, y_i), i = 1, 2, \dots, k\}$, 其中, 输入向量 $x_i = (x^{(1)}, x^{(2)}, \dots, x^{(n)}) \in R^n$ 代表不同的属性或指标; y_i 代表样本类别。对于二分类问题, 所有的样本被分成 A、B 两类, 以 $y_i \in \{-1, +1\}, i = 1, 2, \dots, k$ 来表示, 每一个样本对应一个 y_i , 若 $x_i = +1$, 则将 x_i 分到 A 类, 若 $x_i = -1$, 则将 x_i 分到 B 类。

对于 $y_i = +1$, 对于 $y_i = -1$, 可以得到:

$$y_i [w^T \varphi(x_i) + b] \geq 1, i = 1, 2, \dots, k$$

其中, $R^n \rightarrow R^d$ 表示映射函数, 可将低维的样本集映射到高维空间, 从而实现线性可分, 等号成立的样本点 x_i 被称为支持向量。

由最佳超平面可以计算出两类样本的分类间隔 $2/\|w\|^2$, 要保证分类间隔最大, 等价于使 $\|w\|^2/2$ 最小。另外, 在利用核函数进行分类时, 假定数据是线性可分的, 虽然说将原始数据映射到高维空间后, 能够使线性可分的概率大大增加, 但并非所有的数据都是线性可分的, 因为数据有噪音, 即使在映射后的高维空间也会判别错误。因此, 支持向量机模型在优化问题中引入了松弛变量, 即使映射在高维空间也会被判别错误的概率。因此, 支持向量机模型加上一个松弛变量, 且 $\varepsilon > 0$, 最终的模型为:

$$\begin{cases} \min_{w, b, \varepsilon} (\frac{1}{2} (\|w\|^2) + C \sum_{i=1}^k \varepsilon_i) \\ \text{s. t. } y_i \varphi(x_i) + b \geq 1 - \varepsilon_i \\ \varepsilon_i \geq 0, i = 1, 2, \dots, k \end{cases}$$

其中, $C (> 0)$ 表示样本点被错误分类时的惩罚系数, C 值越大, 代表对错误惩罚越严重。

(2) 选取的核函数为高斯核函数 $k(x_1, x_2) = \exp(-\frac{\|x_1 - x_2\|^2}{2\sigma^2})$, 该核函数会将原始空间映射为无穷维空间。不过 σ 选得过大的话, 高次特征上的权重实际衰减得非常快, 实际上就相当于一个低维的子空间; 反之, 如果 σ 选得非常小, 则可以将任意的数据映射为线性可分, 但这会出现非常严重的过拟合问题。因此, 可以通过参数 σ 进行调控, 高斯函数具有相当高的灵活性。

2) 无行为数据的用户。

对于没有任何历史行为数据的用户来说, 可以通过用户的静态属性信息, 例如性别、年龄, 对用户进行聚类分析, 从而得到与用户同性别、同年龄段的用户群体较为感兴趣的内容, 然后再从该群体的兴趣爱好中选择 Top- k 作为推荐列表。

2.4 内容库设计

内容库用于存储要推送给用户的内容, 推送系统可以根据用户所处的位置、使用手机的习惯等, 从内容库中选取相应的推送内容, 进而推送给用户。一般来说, 内容类别标签都是以层级的形式组织的, 可以有一级维度、二级维度等。其主要来源有已有内容的标签、网络抓取流行标签、对运营内容进行关键词的提取。

(1) 基于位置的内容库。

由于移动终端的便捷性, 移动终端设备 (智能手机) 便于携带, 现有的 GPS 等技术可以方便、快速地定位移动终端用户, 而相同的用户在不同的位置具有不同的爱好, 那么就应当根据用户的具体位置来为用户推送相应服务。推送的内容库是基于位置的内容库, 它是根据中国的行政单位来划分的, 在最低级的行政单位下, 要推送相应的具体内容序列。

(2) 基于图片偏好的内容库。

图 3 为图片的爱好内容库,总共十七个大类,每个大类下平均十几个二级分类,每个二级分类下,拥有多

维度管理 ×			
添加同级 添加下级 修改 删除			
名称	编码	序列	时长(秒)
1 地理位置	001		120
2 使用时长	002		120
8 用户喜好	003		120
9 体育	003001		120
10 篮球	003001001	体育_篮球	120
11 CBA	003001001001	CBA	1800
12 NBA	003001001002	NBA	1800
13 世界男篮锦标赛	003001001003	体育_篮球	1800
14 足球	003001002	体育_足球	120

图 3 图片爱好库

(3) 基于 APP 偏好的内容库。

表 2 为 APP 分类,总共十九大的分类,每种下又有多种二级分类,然后以树状的形式进行组织。这种 APP 分类的类别都是网络爬虫从 APPStore 上爬下来的,从而保持 APP 分类的更新。

表 2 App 偏好内容库

编码	一级标签	二级标签
004001	社交通讯	
004001001		聊天
004001002		社交
004001003		婚恋
.....
004002	系统工具	
004002001		优化
004002002		安全
004002003		输入法
.....

3 自动推送机制

前面分析了用户使用手机的时间情景、位置情景,二者都会影响手机内容是否被用户所接受,并且用户使用手机在时间上具有一定的规律性,如果能够在用户经常使用手机的时段,为用户推送感兴趣的内容,那么就能被用户即时地接触到;当然对于处在不同位置的用户,如果能够根据用户的不同位置,为用户推送附近用户感兴趣的内容,也能达到增强内容的效果。因此要基于这些动态特征数据,触发相应的用户感兴趣的推送内容。触发方式可以是位置、动作、时间、环境属性等,这里最重要的当然是位置和时间了。韩吉等提出了通过无线传感器感知外界气象的变化,实时地向 Android 手机端推送预警通知^[15]。因此,基于位置情景和时间情景两个维度的动态变化来触发相应的推送内容服务,并从相应的内容库中选取相应的内容序列,进而推送给用户。其基于动态特征数据变化的推送算法如

组的要推送的内容序列。如图所示,体育、财经为一级分类,足球、篮球等为体育下的二级分类,篮球下的 CBA、NBA 为事先编辑好的有一定意义的内容序列。

(1) 获取当前用户的经纬度以及当前的时间点,用 $\langle U, U_{\text{LON}}, U_{\text{LAT}}, T \rangle$ 四元组表示。

(2) 判断用户的位置是否发生变化。以用户的位置优先,对于移动智能终端来说,由于其便捷性,用户可以很方便地携带在身上,随时记录用户的位置属性。如果用户的位置发生变化,则优先根据用户的当前所在位置,从内容库中选择相应的推送内容。使用中国的行政单位的边界为界限,来判定用户的位置是否发生变化。每个行政区边界可以使用一组经纬度表示:边界 = $\langle \langle \text{Lon}_1, \text{Lat}_1 \rangle, \langle \text{Lon}_2, \text{Lat}_2 \rangle \cdots \rangle$,可以根据需求来确定边界的大小。那么就可以通过比较用户的位置是否在相应的区域内,来判定位置区域的变化。

(3) 对于推送内容该如何选择。每个位置区域下又有多个事先编辑好的具有一定意义的内容,可以使用 $\text{Item} = \langle I_1, I_2, I_3 \cdots \rangle$ 表示。对于内容的选取采用轮询算法,从中选取用户感兴趣的,又与当前位置相关的内容服务。

(4) 如果用户的位置不发生变化,则会基于用户的时间情景来触发用户的内容推送服务。对于移动终端用户而言,如果用户不使用移动终端,那么即使为用户推送了内容,用户也看不到,当然如果能够在用户使用移动终端的时刻为用户推送感兴趣的内容,那么内容被看到的次数增多,内容被接受的概率就会提高。

(5) 有了触发推送的时刻,使得用户有更大的机会看到推送的内容,然后就是要推送用户感兴趣的东西了,这个要根据用户需求预测的结果来为用户提供个性化推送,使得推送的内容为用户感兴趣的内容。图 4 的 UTP 库代表用户在不同的情景下,用户对某一类别标签的偏好程度。

在用户频繁使用移动终端的基础上,为用户推送该时间段下用户感兴趣的类别标签,类别标签可能有多个,而每个标签又占有不同的权重。如用户 U 感兴趣的类别标签为: $B = \langle c_1, c_2, c_3, c_4 \rangle$,用户对每个类别标签的权重分别记为 $W = \langle w_1, w_2, w_3, w_4 \rangle$,采用

加权轮询算法,该算法首先根据列表标签相应的权重,生成对应的一组新序列: $\langle c_1, c_1, \dots, c_1, c_2, c_2, \dots, c_2, c_3, c_3, \dots, c_3, c_4, c_4, \dots, c_4 \rangle$ 。其中这组序列中有 w_1 个 c_1 ,有 w_2 个 c_2 ,有 w_3 个 c_3 ,有 w_4 个 c_4 ,那么选取的时候,就会从新生成的这组序列中选取。这样使得占有权重较大的标签被选中的概率更大,而且让该类别下的所有内容序列都可以无重复地推送给用户。

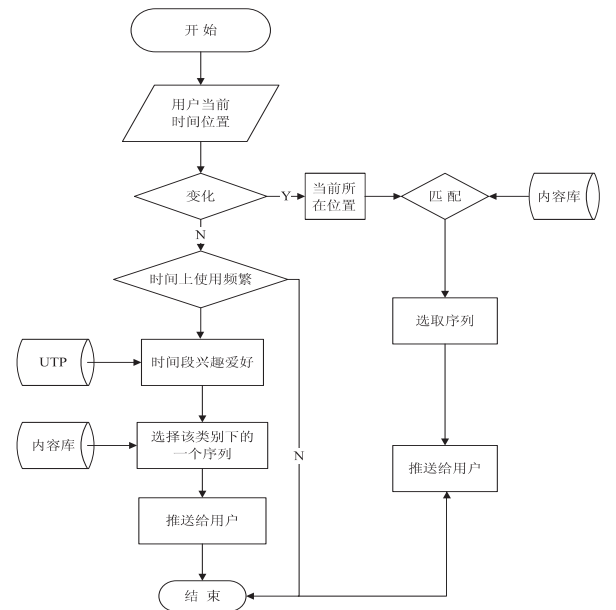


图 4 推送流程

4 实验结果及分析

实验一:针对 SVM 算法惩罚因子 C 以及核函数中 σ 的确定。采用从 200 多位移动端用户收集来的近 400 000 条行为数据,300 000 条用户使用 APP 软件数据。将数据的 80% 作为训练集,剩余的 20% 作为测试集。通过预测的准确率来确定 C 和 σ 的值,见表 3。

表 3 因子与准确率的关系

C	σ	准确率/%
7	0.041	94.3
8	3	93
245	34	95
37	5	92

实验二:测试推送效果。使用两种方式来测试推送服务的效果,一种是模拟仿真,模拟真实的用户;另一种是以用户为中心的方式,让真实用户参与到实验中^[16]。采用第二种方式来进行测试,邀请了 8 位同学作为此次实验的真实用户,给他们的手机上安装该款 APP 软件,并注册使用一个月的时间,在这一个月的时间中,用户的生活和平常一样,分别为其推送相应的图片序列;然后下一个月将其切换为传统的推送技术进行图片推送。用如下公式评估推送的内容被用户接受的情况。万方数据

$$\text{Precision} = \frac{\text{Num}_{\text{slide}}}{\text{Num}_{\text{push}}}$$

其中, $\text{Num}_{\text{slide}}$ 表示推送后用户浏览观看的类别标签数; Num_{push} 表示推送系统为用户推送的类别标签数。

八位用户对于推送内容的浏览情况如图 5 所示。与传统的内容推送相比,这种自动推送方式被用户所浏览的次数普遍要高。由此可见,基于移动特征数据变化的推送更加高效。

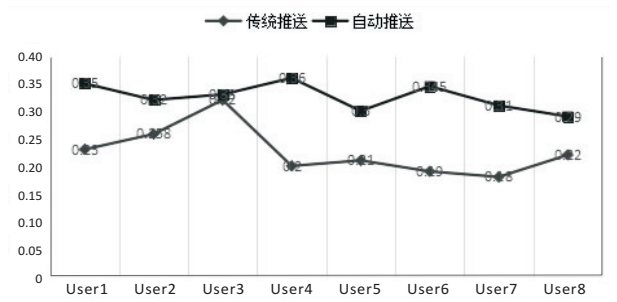


图 5 浏览情况

5 结束语

为解决提高移动终端推送准确性的问题,提出了一种基于移动特征数据的内容推送模型。该模型针对移动特征数据进行分析,预测移动用户需求,并根据移动特征数据的动态变化,动态地为用户推送感兴趣的内容,使得内容推送服务更加及时和准确。当然,推送内容是否精准更多地取决于用户需求预测的准确性,预测得越准确,推送内容也就越精准,用户浏览和接受的概率就越大。

参考文献:

[1] Pan B, Wang X, Song E, et al. Campsf: cloud-assisted mobile service provision framework supporting personalized user demands in pervasive computing environment[C]//9th international wireless communications and mobile computing conference. [s. l.]: IEEE, 2013: 649-654.

[2] Adomavicius G, Tuzhilin A. Context-aware recommender systems[M]//Recommender systems handbook. [s. l.]: Springer, 2011: 217-253.

[3] Ai D X, Zuo H, Yang J. Personalized mobile catering recommender system based on context ontology model and rule inference[C]//Advanced materials research. [s. l.]: Trans Tech Publications, 2013: 708-713.

[4] 孟祥武, 王 凡, 史艳翠, 等. 移动用户需求获取技术及其应用[J]. 软件学报, 2014, 25(3): 439-456.

[5] Yeung K F, Yang Y, Ndzi D. A proactive personalised mobile recommendation system using analytic hierarchy process and Bayesian network[J]. Journal of Internet Services and Appli-

对上述问题,为满足定量与定性混合过程信息的融合信息处理要求,在分析研究基础上,提出了结合过程神经网络的定量信息处理技术与模糊逻辑推理方法的模糊推理过程神经网络模型。实验结果表明,该模型对于处理沉积微相模式识别问题具有很好的识别效果,显著提高了预测精度,是一种比较实用的判别模型。

参考文献:

[1] 宋翔. 基于神经网络与图形处理技术的沉积微相识别系统[D]. 长春: 吉林大学, 2004.

[2] 耿茜. 安塞油田桥 19 井区长 3 油层组沉积相与油气聚集规律研究[D]. 西安: 西安石油大学, 2015.

[3] 程超, 胡望水, 秦刚, 等. 精细油藏描述中的沉积微相研究—以吉林油田大 208 区黑帝庙油层为例[J]. 石油地质与工程, 2009, 23(2): 26-28.

[4] 高宪伟. 应用测井资料定量识别碳酸盐岩沉积微相的方法研究[D]. 成都: 成都理工大学, 2013.

[5] 孙铁军, 冯庆付, 刘浩, 等. 利用测井资料计算机自动识别沉积微相方法研究[J]. 西北地质, 2010, 43(3): 132-140.

[6] 许少华, 陈妍, 许辰, 等. 一种基于正向云变换的混合计算神经网络及其应用[J]. 计算机与数字工程, 2014, 42(12): 2284-2288.

[7] 许少华, 毕聪聪, 张宇, 等. 基于径向基过程神经网络的油田开发指标预测[J]. 计算技术与自动化, 2015, 34(3): 52-54.

[8] 张亚光. 基于云变换的信息融合方法及智能推理模型的研究[D]. 大庆: 东北石油大学, 2014.

[9] 陈惟昌, 王自强, 陈志华, 等. 神经信息学的原理与展望[J]. 生物物理学报, 2001, 17(4): 613-620.

(上接第 160 页)

cations, 2012, 3(2): 195-214.

[6] 常慧君, 单洪, 满毅. 基于分段、聚类 and 时序关联分析的用户行为分析[J]. 计算机应用研究, 2014, 31(2): 526-531.

[7] 范琳, 王忠民. 穿戴位置无关的手机用户行为识别模型[J]. 计算机应用研究, 2015, 32(1): 63-66.

[8] 周朴雄, 张兵荣, 赵龙文. 基于 BP 神经网络的情境化信息推荐服务研究[J]. 情报科学, 2016, 34(3): 71-75.

[9] 倪红军. 基于 Android 平台的消息推送研究与实现[J]. 实验室研究与探索, 2014, 33(5): 96-100.

[10] 孙泽军, 常新峰. 基于 XMPP 推送技术在移动 OA 中的应用研究[J]. 实验室研究与探索, 2015, 34(7): 130-134.

[11] 律智坚, 吴广财. 消息推送在移动高级应用中的研究与实现[J]. 广东电力, 2014, 27(2): 117-120.

[10] 吴雅娟, 许少华. 一种模糊计算过程神经元网络及其应用[J]. 东北林业大学学报, 2008, 36(8): 43-44.

[11] Shang L, Liu G, Zhang R, et al. An information fusion algorithm for integrated autonomous orbit determination of navigation satellites[J]. Acta Astronautica, 2013, 85(4): 33-40.

[12] 李小红. 信息 Agent 在 CSCW 中的应用[D]. 大庆: 东北石油大学, 2006.

[13] 许少华, 张亚光, 李学贵. 一种基于逆向云变换的混合推理神经网络[J]. 计算机技术与发展, 2015, 25(3): 118-121.

[14] Paylin G, Oude P D, Maris M, et al. A multi-agent systems approach to distributed Bayesian information fusion[J]. Information Fusion, 2010, 11(3): 267-282.

[15] 李盼池, 许少华. 加权模糊推理网络及在水淹层识别中的应用[J]. 计算机应用, 2004, 24(10): 105-107.

[16] Tiruneh A T. Higher order aitken extrapolation with application to converging and diverging gauss-seidel iterations[J]. Journal of Applied Mathematics & Physics, 2013, 5(11): 128-143.

[17] 衣治安, 牟春苗, 孙寅萍. 基于改进粒子群算法的模糊神经网络[J]. 计算机系统应用, 2014, 23(6): 135-140.

[18] 董焕忠. 基于加权模糊推理网络的河流—三角洲储层沉积微相识别[J]. 大庆石油学院学报, 2005, 29(1): 24-27.

[19] 裴琴娟. 解线性方程组的共轭梯度法[J]. 新乡学院学报: 自然科学版, 2011, 28(4): 309-310.

[20] 王金海, 丁锋. CARMA 模型离线最小二乘迭代辨识方法[J]. 科学技术与工程, 2007, 7(23): 5998-6003.

[21] 徐士良. 计算机常用算法[M]. 第 2 版. 北京: 清华大学出版社, 1995: 95-101.

[22] Messaoudi S, Messaoudi K, Dagtas S. Bayesian data fusion for smart environments with heterogenous sensors[J]. Journal of Computing Sciences in Colleges, 2010, 25(5): 140-146.

[12] 郑小雪. 基于知识地图的政府诉求文件自动推送模型研究[J]. 现代情报, 2015, 35(8): 43-46.

[13] 孟祥武, 纪威宇, 张玉洁. 大数据环境下的推荐系统[J]. 北京邮电大学学报, 2015, 38(2): 1-15.

[14] 章少平, 梁雪春. 优化的支持向量集成分类器在非平衡数据集分类中的应用[J]. 计算机应用, 2015, 35(5): 1306-1309.

[15] 韩吉, 周杰, 杜景林. 基于 Android 的气象 WSN 监测系统设计与实现[J]. 计算机工程与设计, 2014, 35(8): 2709-2714.

[16] Hawalah A, Fasli M. A multi-agent system using ontological user profiles for dynamic user modelling[C]//Proceedings of the 2011 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology. [s. l.]: IEEE Computer Society, 2011: 430-437.