

基于深圳天气移动互联网渠道用户行为分析的 智能气象服务模型构建

■ 李海龙 曹梅 文凤洁

通过对深圳天气微信公众号和深圳天气手机客户端2017年9月—2018年9月期间的活跃用户进行用户行为数据进行分析,再结合移动端日常运行所需的气象数据,建立基于用户行为数据和气象数据相结合的智能气象服务模型,为气象服务从传统的被动获取向个性化、智能化、主动推送式的大城市精细化气象服务转变提供借鉴。

DOI: 10.3969/j.issn.2095-1973.2019.03.035

移动互联网目前已经成为了公众获取天气信息的主要渠道,利用移动互联网的便利性展开个性化的气象信息服务将有助于推进整个社会的科技创新能力及数字城市的建设。

近年来,深圳市气象局的公众气象服务也已将重心放在深圳天气微信公众号和深圳天气客户端等移动互联网公众服务渠道产品。深圳天气微信公众号、深圳天气手机客户端是深圳市气象局在“移动互联网+气象服务”形势下的新型气象服务模式。本文选用深圳天气微信公众号和深圳天气手机客户端2017年9月—2018年9月后台记录的用户行为数据信息,通过对其进行数据分析,得出以公众需求角度的智能气象服务模型,以期为大城市精细化气象服务提供借鉴。

1 深圳天气移动互联网渠道和气象大数据服务现状

1.1 深圳天气移动互联网渠道服务产品现状

近年来许多气象部门通过移动软件来传递天气信息,进而改善传统的公共气象服务的传播方式,充分弥补传统的气象预报信息媒介如广播电视以及报纸等方式的不足。虽然行业和公众获取气象服务的渠道日趋多样化,但是目前绝大多数气象服务模式仍然是“气象服务=天气预报+简单提示”的传统模式,能够体现差异化、精细化的服务产品少之又少。从2018年5月中国互联网信息中心(CNNIC)发布的《中国互联网络发展状况统计报告》显示,截至2017年12月我国网民使用手机上网的比例达97.5%,较2016年底提升了2.4%,移动互联网渠道已经成为用户的主阵地。

在移动互联网气象渠道服务产品方面,深圳天气微信公众号和深圳天气客户端在用户数量方面始终处于行业领先地位。自2011年陆续上线截至2018年10月,深圳天气微信公众号粉丝总数已达98万,点击总数114.5亿;深圳天气客户端累积下载量达1900万次。在功能方面,深圳天气微信公众号和深圳天气客户端立足于移动互联网的公众气象服务产品,通过GPS定位功能实现显示用户所在位置最临近自动站的实况和临近预报,提供以用户为中心精细到街道级别的气象服务产品。深圳天气微信公众号和深圳天气客户端(以下简称移动端)不仅在深圳市获得较高公众的认可,也在深圳市气象局连续5年获得广东省气象服务公众评价第一名,所以选取上述移动端的用户作为研究对象无论从数据上还是从效果上均具有移动互联网气象服务的代表意义。

1.2 国内外气象大数据研究现状

目前国外对于大数据的主要研究方向是将其作为一种具有更高的处理速度和更优处理能力的信息资产,它是在密集型数据的情况下,对数据进行更进一步的探索和思考的一种新模式。大数据分析的核心在于通过一种更加有效的方式来管理庞大的数据库并且从海量数据中汲取所需要的价值,它是伴随着数据科学的发展而产生的一种最为先进的分析思维和方法。目前国外的研究较为重视个性化细小区域的数据分析,并通过数据分析掌握用户需求达到提供以用户为中心的信息服务;而国内对于气象大数据的分析绝大部分指的是观测数据的探究,但目前可参考和借鉴的

收稿日期:2018年11月30日;修回日期:2019年4月14日

用于气象服务的案例相对较少。尽管目前国内外对于大数据分析有众多的分析成果，但是在实际应用于气象服务或基于用户使用习惯的气象服务更少。气象行业本身具有大量的大数据资源，目前随着气象数据量不断翻番，每年的气象数据已接近PB量级。通过结合气象数据和用户行为数据进行气象服务将成为一种新的服务模式。气象数据主要可分为两大类：一类是实况数据，即利用各种观测手段观测到的“过去的的数据”；另一类为模式数据，利用观测到的数据通过模式计算得到的“未来的数据”用于预测未来天气。如何有效地将气象数据与用户行为数据相结合，产生基于用户行为的气象数据将会为气象服务提供一个新的方式。

2 获取移动端活跃用户的数据

本研究通过对2017年9月—2018年9月期间同一天内3次或以上访问移动端且位置在深圳的用户作为研究样本主体，共筛选出1072582人次，并定义该样本主体为移动端过去一年的活跃用户。通过在移动端后台各功能模块插入监听软件，即设置“埋点”和借助手机GPS获取定位信息来记录活跃用户的日活跃时段、点击习惯以及所在位置的数据形成用户标签系统，建立相应的用户行为数据库，形成精细化气象服务模型。本研究搜集的用户数据主要基于用户位置的点击次数，不涉及用户敏感信息，如：身份证号、银行卡号、家庭住址等。研究过程中及研究结束后未有任何用户位置信息泄露，不会对用户隐私以及安全造成影响。

2.1 用户活跃时段数据

用户活跃时间段的获取即通过监测同一用户一天内3次或以上访问移动端的时间段。经过统计2017年9月—2018年9月1072582名活跃用户每天集中访问时间得出用户活跃时段分布（图1），可以看出大多数活跃用户获取气象信息的时段主要集中在07—08

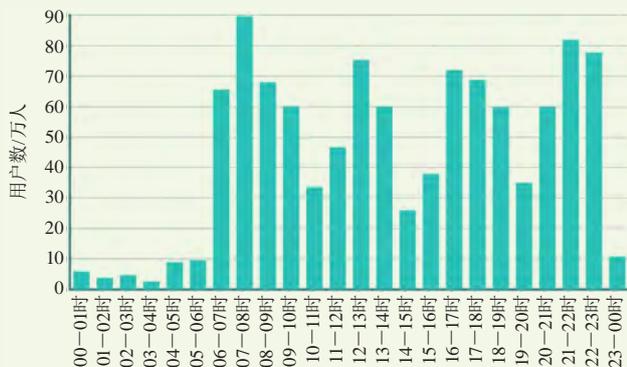


图1 活跃用户时间段分布

时、12—13时、16—17时、21—22时。现针对不同ID的用户基于不同活跃时间段记录为该用户的活跃时段标签。

2.2 活跃用户日点击习惯大数据

通过在移动端各功能模块插入监听软件，即设置“埋点”来记录1072582名活跃用户在一天内的点击习惯，并针对不同功能模块用户的点击习惯获取活跃用户的日点击习惯大数据（图2），可以看出，活跃用户全天点击习惯排名前三的功能模块是“天气提示”“交通服务”“天气实况”，说明用户通过移动端获取“天气实况”和“天气提示”等常规气象信息的同时，也关注天气对交通的影响。而每天12—18时点击“定点预报”和“雷达图像”的用户明显增多，则是因为深圳午后经常出现局地短时阵雨，天气不稳定的情况下用户更关注短时临近预报。现针对不同ID的用户基于不同功能模块的点击习惯记录为该用户的活跃时段标签。

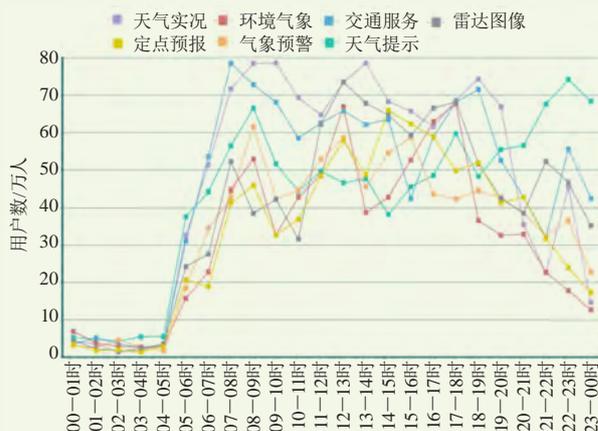


图2 活跃用户日点击习惯

2.3 活跃用户位置数据

利用手机GPS定位功能选取用户在一天内访问3次以上移动端的地点，根据经纬度选取其所在地半径1 km范围以内的公共场所如“住宅区”“车站”“公园”标记为该用户的位置，可以得出活跃用户一天内的位置变化情况（图3），可以看出大多数用户习惯于早晨或傍晚在车站或公园点击移动端获取气象信息。在住宅区的活跃用户数量从20时起呈明显增长趋势，说明用户习惯于在家睡前访问移动端。在商场和餐厅的活跃用户数量相较于其他位置的用户数量偏少且变化率不大，说明用户在休闲场所较少关注天气信息。而在写字楼的活跃用户在16—18时数量较多，说明用户在工作场所更加关注下班天气。现针对不同ID的用户基于不同位置记录该用户一天内基于时间的位

置标签。

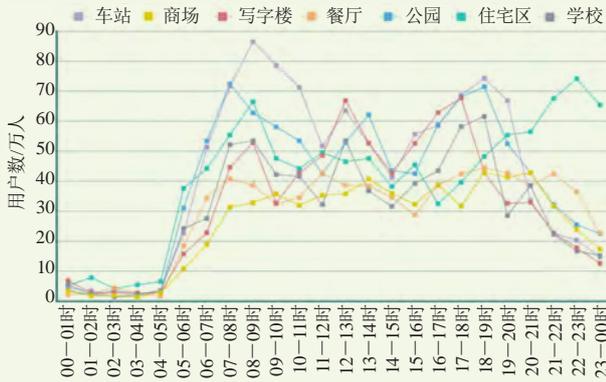


图3 活跃用户一天内位置变化

3 建立以移动端为渠道的智能气象服务系统

以某一用户活跃时段标签为“06—07时”，活跃地点在“公园”，日点击习惯为“天气实况”和“环境气象”为例，系统判定该用户较大可能关注早晨时段公园附近的天气实况信息，提取时间标签“06—07时”、地点标签“公园”为触发条件，并匹配用户点击习惯“天气实况”功能分发给该用户，移动端则会根据系统判定在06—07时向该用户推送公园附近天气实况信息。即用户标签系统以“提取用户标签”为触发条件，以“匹配特征用户”为分发机制向用户提供定制化服务，并且在该用户获取推送的天气信息时再搜集该用户其他的行为数据，做到提供服务的同时再次获取的双向服务方式来获取更多的用户数据，形成以移动端为服务渠道基于用户行为分析的服务闭环，从而不断优化和丰富现有的用户数据库形成用户标签系统。再以用户标签系统为服务基础、移动互联网渠道为服务载体、标签化处理的用户为服务对象，并采用交互式服务方式，形成基于移动端用户行为数据智能服务的基本构架（图4）。在上述智能服务构架的基础上不断优化不同场景服务过程中记录到的数据，获得具有大量用户行为标签的气象服务产品集合，以此搭建能够匹配特征用户的智能气象服务产品数据库。通过将该数据库进行端口开放，建立能够依据用户标签匹配用户需求且具有多渠道服务能力的智能气象服务系统。

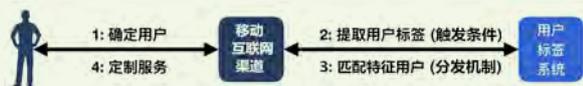


图4 基于移动端用户行为数据智能服务的基本构架

4 结语

经过对2017年9月—2018年9月期间的活跃用户行为数据的分析得出，大多数用户习惯于在早晨上班

前或夜晚睡觉前点击移动端获取气象信息，且移动互联网用户更倾向于获取所在位置的天气实况、短临预报和交通气象服务信息。但为满足用户在不同时间、不同地点、不同场景的精细化气象服务的需求，需要将用户行为数据和气象数据进行整合并建立基于移动互联网的气象服务用户标签。移动互联网渠道在借助气象服务用户标签对外服务的同时不断搜集用户行为数据，形成“边服务、边获取、边分析”的智能气象服务数据库，并提供可扩展的开放式端口，实现多渠道、个性化、智能化的气象服务构架（图5）。

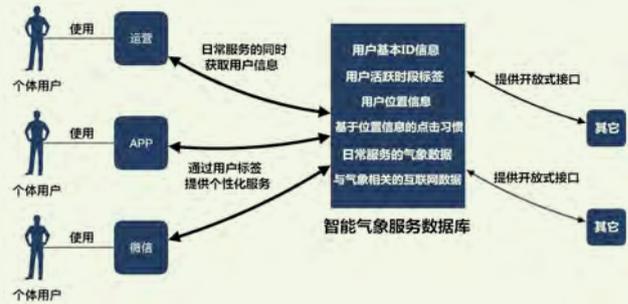


图5 基于用户行为数据的个性化、智能化的气象服务模型

气象行业本身即是具有大数据属性的行业，每天会产生大量的天气实况和预报模式数据，具有进行大数据分析和研究的行业优势。但从现阶段来看，庞大的数据并没有得到充分开发和利用，也未做到基于公众需求提供。本文采集了过去一年的活跃用户数据并建立基于深圳天气移动互联网的服务构架，但现阶段获取的用户数据仍不足以使智能气象服务数据库达到真正的智慧化，所以用户的数据分析和精准度训练需要一个长期的积累过程。通过开放式接口来对接不断更新的气象数据，未来也可新增其他行业的数据，如交通、旅游等，随着第三方数据和服务渠道的不断增加，最终实现跨行业、多渠道的智能气象服务数据库，提升气象公众服务的整体质量，实现基于用户需求的大城市精细化气象服务。

深入阅读

- 姜青山, 2018. 浅谈气象服务App的开发与应用. 科技风, (1): 124-124.
- 刘东华, 张琳琳, 2012. 深圳市气象信息服务个性化的设计与展望. 广东气象, 34(3): 57-59.
- 杨承睿, 任芳, 马楠, 2016. 试论大数据在气象服务中的应用. 农业网络信息, (8): 53-55.
- Provost F, Fawcett T, 2013. Data science and its relationship to big data and data-driven decision making. Big Data, 1(1): 51-59.
- 王丽, 李云鹏, 甄熙, 2018. 浅析互联网大数据在气象行业的应用. 电脑知识与技术, 14(24): 218-219.

(作者单位: 李海龙、文凤洁, 深圳市气象服务中心; 曹梅, 广东省气象公共服务中心)