

基于气象服务大数据洞察平台的用户行为分析

■ 耿蓉

上海市气象服务中心的“气象服务大数据洞察平台”依托“多终端跨屏大数据”优势建立，平台基于囊括PC（常用气象类服务网站）、IPTV（气象类电视栏目）、手机终端（常用气象类手机APP）以下简称PC、IPTV、Mobile多终端数据来源的用户流量数据，主要提供PC、IPTV、Mobile三终端的上海市民每日访问的PV以及UV数据。

DOI: 10.3969/j.issn.2095-1973.2021.06.013

传统气象信息管理和共享服务系统在可复用资源、系统集约化程度、信息管理能力和加工处理能力、共享服务能力等方面的局限性已经无法适应新型气象业务需求。以用户需求为导向，立足于市民生活、企业生产、政府管理的实际需求，需要不断拓展大数据、云计算、物联网、移动互联网等最新信息技术在气象领域应用的广度和深度，切实增强信息化对提升气象服务业务水平的支撑、引领和带动效用。随着大数据时代的到来，大数据和气象服务的结合越来越紧密。大数据应用于各专业气象服务的理念，核心是大数据技术获取非气象类的大数据，及其与气象数据的碰撞，表现为专业气象服务水平的提升和服务方式的革新。将“气象行业大数据”和“气象互联网大数据”有效融合，能够更加有效地应对各种业务服务需求和社会服务需求。喻迎春等利用气象相关的互联网新闻、舆情大数据和新媒体发展数据，分析了江西省公共气象服务需求，发现大部分公众更关注当天天气，对其他气象相关问题的参与度较低。对于公众关注气象信息的渠道来说，于庚康等研究发现江苏省公众获取气象信息的主要渠道是电视、手机和网络。曹梅等利用深圳天气官方微博、微信服务号、手机客户端为研究对象，结合不同天气过程进行对比，分析了各种服务产品的用户数量变化和公众关注度变化进行分析，发现客户端的点击数和下载量具有明显的年变化特征且在重大天气过程频繁期对客户端的需求量较大。

本文对上海市气象服务中心气象服务大数据洞察平台的后台用户行为大数据进行深入挖掘分析，实现对上海市民气象服务使用情况定性、定量分析。通过

总体洞察、分析用户画像用户标签，并结合气象要素变化和用户行为变化之间的相关性分析，真实还原了上海市民在不同具体时期的气象服务使用情况，充分发挥了数据信息的服务功能。

1 数据

上海市气象服务中心的“气象服务大数据洞察平台”依托“多终端跨屏大数据”优势建立，平台基于囊括PC（常用气象类服务网站）、IPTV（气象类电视栏目）、手机终端（常用气象类手机APP）以下简称PC、IPTV、Mobile多终端数据来源的用户流量数据，主要提供PC、IPTV、Mobile三终端的上海市民每日访问的PV以及UV数据。其中PV（Page View）访问量数据定义为浏览量或点击量，旨在衡量用户访问的网页（电视节目、APP）的数量，在每日的统计周期内用户每打开或刷新一个页面就记录1次，多次打开或刷新同一页面则浏览量累计。UV（Unique Visitor）代表独立访客数量，可以理解成一天内访问某网站（电视节目、APP）的终端的数量。在多次打开但不清除cookies，再次访问相关内容时，统计的UV数是不变的。此外该平台还提供基于电信集团用户大数据库的气象服务使用者的用户标签、气象网站评分以及实时气象热词等内容。

本文主要使用数据为气象服务大数据洞察平台监测到的多端渠道数据（PC网站、IPTV栏目、手机APP）的PV以及UV数据，它基于对热门气象服务网站、手机APP以及气象类电视栏目的总体使用日常流量数据监测统计。此外还有依托上海电信公司大数据库中的上海市使用气象服务人群的用户标签，该数据标签库囊括了用户的年龄、性别以及车辆使用情况等

收稿日期：2019年12月30日；修回日期：2020年11月30日

作者：耿蓉，Email: 415129318@qq.com

信息。平台数据能够较为有代表性地反映上海市民对气象服务的使用情况。该大数据平台于2017年年底投入运行，本文选取气象服务大数据洞察平台2018年3月1日至2019年2月28日整年的数据。文中所使用的气象数据是上海市区国家站徐家汇站点（站号58367）的气象数据。

2 数据获取及分析

2.1 总体洞察

统计2018年3月1日至2019年2月28日期间每日的PC、IPTV、Mobile三端之和PV、UV数据，时间跨度包含春夏秋冬四个季节。由于这是目前公众主要获取三种气象渠道的来源，认为其代表上海市民对气象信息的关注度。从2018年3月1日至2019年2月28日期间每日的PC、IPTV、Mobile三端之和的PV、UV时间序列来看（图1），PV、UV时间序列变化整体趋势一致，上海市民对气象服务使用情况存在季节，这与曹梅等发现的深圳市民对气象服务的使用情况较为相似。具体表现为上海市民在春、夏、秋季对气象信息关注度较高，冬季关注度相对来说略低，这可能是由于冬季相对其他季节所发生的天气过程相对平稳，所以市民对气象信息的关注度偏低。从UV、PV时间序列来看，在某些时间段，上海市民对天气关注度变化存在波动，下文还将针对这些关注度变化比较大的部分做个例分析。

对前文中提到的UV、PV数据进行计算，粗略认为用每日的PV/UV后可以得到对应的用户关注气象信息的次数。用户关注气象信息的次数大约在5~25次不等，没有存在明显的季节变化。关注度较高的两个峰值出现在2018年“十一”长假前期和12月上旬。关于12月上旬，出现了一次寒潮过程，将在后面具体分析。将Mobile、PC、IPTV三个终端数据分开来看，发现用户使用手机端每日访问次数最高。平均来说，Mobile用户每天使用7次、PC端使用3次、IPTV大约在1次。可见，比起传统的网站和电视栏目，手机客户端占有明显的市场优势，这是由于目前手机网民的高普及率。根据中国互联网络信息中心发布的第44次《中国互联网络发展状况统计报告》，截至2019年6月，我国网民使用手机上网的比例达99.1%。由此可以看出，做好手机端的用户气象服务还有很大的空间和市场。

为了对比分析不同终端渠道用户使用气象服务的区别，将三端PV数据分开来看（图2），绝大多数流

量来自手机端（Mobile）来访问天气，占比96%，PC端访问量只有3%，而IPTV仅仅为1%。这印证了目前气象影视收益下降的根本原因是由于市民使用电视这一终端获取气象服务的比例陡降。因此，气象部门在利用好气象类APP、官方微信号、官方微博的优势做好服务的同时，应当考虑气象影视节目的转型升级来吸引更多的电视用户。

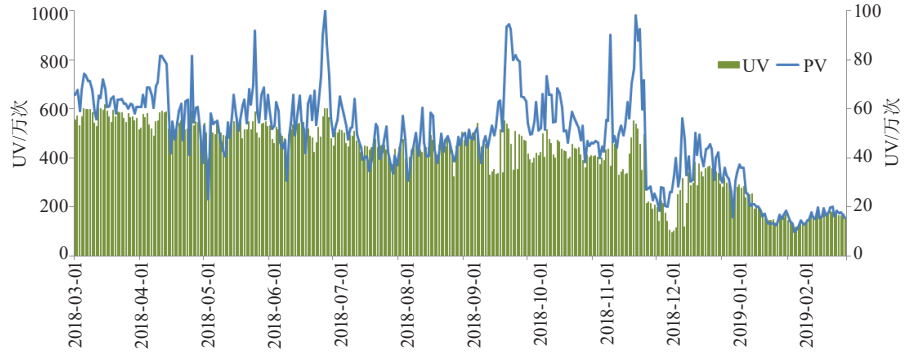


图1 PC、IPTV、Mobile每日PV、UV之和的时间序列

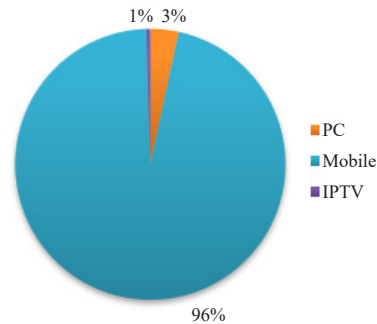


图2 PC、Mobile、IPTV终端使用气象服务的用户占比

2.2 用户画像浅析

用户画像是以用户的资料、行为、偏好分析结果为基础，直观描述用户的基本特征及其偏好情况，能够为气象服务使用行为洞察等重点应用提供用户层面的信息支撑。对于2018年3月至2019年2月这一时间区间内，结合UV数据统计结果及电信大数据提供的部分基于PC、Mobile使用互联网气象服务的用户人群画像标签，包括性别、年龄、家庭车辆拥有情况等属性内容进行分析。

对于上海市使用气象服务用户的市民年龄进行统计可以发现（图3），用户群体集中分布在30岁以上的成年人群，30岁以下的用户仅仅占到用户总数的11%。

对上海市关注气象信息市民的性别进行分析可以发现，女性用户占比高达72%，远高于男性，因此女性比男性更加关注天气情况，后期可考虑针对成年女性用户推送穿衣搭配、育儿信息等精准广告投放。

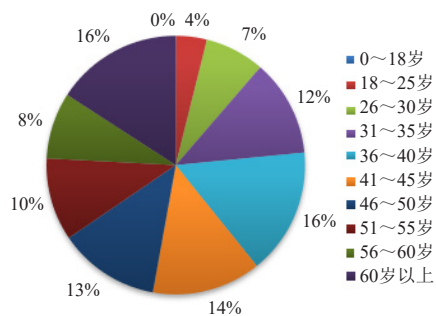


图3 气象服务用户年龄分布情况

对上海市关注气象信息市民的汽车拥有情况进行分析可以发现（图略），在使用上海市气象服务的用户群体中，29%的用户拥有汽车，占总用户的近三分之一。根据上海市交通行业发展报告，2018年全市注册机动车千人拥有率为161辆/千人。后期可以考虑为车主定向推送囊括出行指数、洗车指数的专属的气象服务产品。

有效掌握和利用上述年龄性别以及车辆拥有情况的用户画像，有助于全面掌握上海气象服务用户互联网生活特征，感知不同天气场景下不同标签用户的行为变化，以支撑相关气象公共服务工作开展。未来将更加精准地对人群实现分类服务，进一步开辟更多围绕“气象大数据”应用服务的新市场。

2.3 气温和降水对用户行为的影响

为了更加深入洞察气象要素的变化与上海市民对气象信息关注度变化之间的对应关系，探究每当气象要素变化时，公众对气象信息的关注度情况，选取文章第一部分中时间序列变化较大期间的流量数据并结合气温和降水两个公众主要关心的气象要素来进行个例分析。

对于降水，采用上海市区徐家汇站点的24小时累计降水量。对于气温，由于温度的变化对百姓的实际生活更有影响，因此，用日变温（即每日平均气温之差）以下简称日变温，来代替日平均气温，考虑其对公众对气象信息关注度的影响。

2.3.1 暴雨及台风天气过程

表1列出了部分PV变化率较高的时间段的雨量、日变温、PV、PV变化率、UV、UV变化率及PV/UV数据。对于4月23日当日的一次降水过程，徐家汇站点降水量为53.1 mm，日变温为4.5 °C。在4月22日，PV增长率就比前一天增长33.4%，UV则增长了9%，PV/UV的值也有所提高。4月24日，当此次降水过程结束，PV、UV变化率又有了明显的下降，与之对应的PV/UV的值也下降了。在5月25日，徐家汇站点降水量为44.2 mm，当天日变温-1.3 °C，当日，用户PV

提高了30.9%，UV增长了7.6%，PV/UV也显著增长。

表1 夏季时段雨量、日变温及对应的PV、PV变化率，UV、UV变化率及PV/UV

日期	雨量/mm	日变温/°C	PV/次	PV变化率/%	UV/次	UV变化率/%	PV/UV
2018-04-21	0	0.4	4725929		471603		10.021
2018-04-22	0	2.9	6304436	33.40	514072	9.01	12.264
2018-04-23	53.1	-4.5	6377475	1.16	516517	0.48	12.347
2018-04-24	1.5	-5.9	4102671	-35.67	484971	-6.11	8.460
2018-05-24	0.2	0.1	7001584		546726		12.806
2018-05-25	44.2	-1.3	9166781	30.92	588111	7.57	15.587
2018-06-18	0	2.5	4188057		411520		10.177
2018-06-19	31.5	0.4	6192499	47.86	516204	25.44	11.996
2018-08-02	0	-0.4	4721061		470462		10.035
2018-08-03	56.5	-3.5	5709617	20.94	505882	7.53	11.286
2018-08-04	0	3.5	4763654	-16.57	396276	-21.67	12.021
2018-08-10	0	0.6	4080524		420554		9.703
2018-08-11	0	-0.4	4756977	16.58	431910	2.70	11.014
2018-08-12	64.4	-3.7	6023647	26.63	493135	14.18	12.215
2018-08-13	9.8	1.6	4611249	-23.45	458880	-6.95	10.049
2018-08-14	0.3	0.2	4042941	-12.32	442943	-3.47	9.127
2018-08-15	0.6	0.6	4081038	0.94	440613	-0.53	9.262
2018-08-16	11.9	-1.6	5826454	42.77	491836	11.63	11.846
2018-08-17	70.6	-1.6	5708191	-2.03	472374	-3.96	12.084
2018-08-18	0	2.2	4313844	-24.43	441382	-6.56	9.773
2018-11-20	0	2.3	7644272		550470		13.887
2018-11-21	31.3	-0.1	9793143	28.11	595647	8.21	16.441

对于2018年的台风季来说，台风和降水过程所带来的气象信息关注度的变化也十分明显。2018年8月3日，第12号台风云雀（热带风暴级）给徐家汇站点带来的降水量达到56.5 mm，当天日变温为-3.5 °C，PV上升了20.9%，UV也提高了7.5%。8月12日第14号台风摩羯（热带风暴级）给徐家汇站点带来64.4 mm的降水量，日变温为-3.7 °C。当天，PV变化率为26.6%，UV提高了14.2%，PV/UV也有所提高。2018年8月17日，第18号台风温比亚（热带风暴级）清晨在上海浦东新区南部沿海登陆，这是近一个月以内第3个登陆上海的台风，在上海140多年气象记录史上首次出现，当日，城区徐家汇站降水量达到70.6 mm。可能是由于媒体及时的宣传报道，早在8月16日台风登陆的前一天，PV变化率就上升了42.8%，UV变化率也上升了11.6%，PV/UV也明显提高。

由此可见，每当暴雨、台风等重大天气过程发生时，市民对气象信息的关注度骤升，无论是PV代表的用户次数以及UV代表的用户量以及PV/UV都有明显的提高，在一些重要天气过程来临时，由于官方媒体的信息宣传发布到位，关注度的提升甚至出现在天气过程发生之前。作为气象服务部门，每当重大天气过

程发生时，应当加密公众气象服务，及时做好信息发布和群众生产生活引导工作。

2.3.2 冬季寒潮及连阴雨天气过程

2018年12月，上海发生了一次寒潮天气过程，选取2018年12月2—9日的连续降温天气过程，列出期间雨量、日变温、PV、PV变化率、UV、UV变化率以及PV/UV数据。从表2可以看出，12月4日起当寒潮天气过程发生时，PV变化率和UV变化率都有提高，其中PV提高了53.2%、UV提高了57.9%。12月7日当天，降水量达到19.2 mm，日变温为-5.5 °C，当天PV增长了30.43%，UV增长了29.43%。值得注意的是，从12月6日开始，PV/UV的值明显提高，而当寒潮过程结束时，由于UV骤降，而PV却变化不大，导致12月8日、12月9日PV/UV显著上升，具体原因还有待进一步探究。

表2 冬季时段雨量、日变温及对应的PV、PV变化率，UV、UV变化率及PV/UV

日期	雨量/mm	日变温/°C	PV/次	PV变化率/%	UV/次	UV变化率	PV/UV
2018-12-02	9.3	1	1826565		141400		12.918
2018-12-03	3.8	0.4	2798268	53.20	223208	57.86	12.537
2018-12-04	5.8	-3.5	2764997	-1.19	214604	-3.85	12.884
2018-12-05	4.9	-2.5	2026635	-26.70	173973	-18.93	11.649
2018-12-06	27.9	-1.8	1993824	-1.62	141617	-18.60	14.079
2018-12-07	19.2	-5.5	2600513	30.43	183292	29.43	14.188
2018-12-08	2.2	-2.1	2598217	-0.09	93670	-48.90	27.738
2018-12-09	2.1	0.8	2765207	6.43	100929	7.75	27.398

2018—2019年上海发生了百年一遇的阴雨湿冷天气。在徐家汇有气象记录以来的145年中，2018—2019年的连绵阴雨仅次于同时期的1918—1919年，居于第二高位。选取2019年2月7—22日（共计16 d）时间段的雨量、日变温、PV、PV变化率、UV、UV变化率及PV/UV数据并与之前相对降水量较少的16天进行比对。2月7—22日的PV累计26291480次，比1月24日到2月6日的累计19637720次增长高达34.6%，而UV则高出26.7%。1月24日—2月22日期间，当2月7日连阴雨过程开始后，无论是PV还是UV都呈现整体递增的趋势。

由此可见，当冬季的连续降温过程发生或连阴雨持续时间较长时，用户对于气象信息的关注度明显提高。回顾2018—2019年期间上海市发生的连阴雨天气过程，新闻媒体频频爆出干衣机、内衣内裤等季节性用品一度脱销的新闻，今后可考虑在气象APP栏目进行定向广告投放，实现精准营销。

3 结论及展望

基于气象服务大数据洞察平台的用户PV、UV数

据并结合用户标签库以及气象数据进行研究可以发现，总体来看，上海市民对气象服务的使用情况呈现季节变化，在春、夏、秋三季的关注度明显高于冬季。大部分用户倾向于使用手机来获取气象服务。用户画像方面，气象服务的用户以成年人群为主，女性用户居多，关注天气的市民中大部分没有私家车。当暴雨、台风、降温等天气过程发生时，市民对气象信息的需求情况骤升，当天气过程结束后，需求会有所下降。

大数据的核心是预测，利用气象服务大数据洞察平台所监测的用户气象服务使用情况以及实况观测气象要素数据，将人的行为数据和气象数据深度融合分析，旨在打造创新性的大数据气象服务产品，为精准气象服务提供实时数据导向和支持，提升对以用户需求为导向的精细化智能服务产品的服务能力。

由于气象服务大数据洞察平台投入应用伊始，随着时间序列的增长，今后长时间序列用户行为大数据结合气象数据能够更加深入的探究气象要素的变化对用户行为的影响。除了常规的降水量、日变温以外下一步我们还将探究空气质量、能见度等与生产生活密切气象要素变化与用户使用气象服务变化之间的相关性关系。对于用户画像，将在后期采用更多有指导意义的用户信息进行融合，做更加精细的画像，从而能够进一步根据天气的变化对不同人群对气象服务使用的行为进行预测，在交通、物流、旅游、保险、餐饮外卖等民生气象服务领域推送以需求为导向的精细化气象服务产品。此外，还将引入其他大数据平台进行数据量的扩展和质的提升，实现多元化大数据与气象数据的碰撞，提升专业气象服务水平，进一步增强信息化对提升气象服务业务水平的支撑、引领和带动效用。

深入阅读

- 曹梅, 李海龙, 肖然, 2018. 深圳市重大天气过程互联网数据变化与气象服务. 广东气象, 40(5): 51-53.
- 达月珍, 2019. 浅析气象部门运作天气类app和气象微信公众号的优势. 气象科技进展, 9(3): 60-64.
- 梁文昌, 2017. 数据平台的设计与实现. 北京: 北京交通大学.
- 沈文海, 2016. 再析气象大数据及其应用. 中国信息化, (1): 85-96.
- 唐延婧, 彭芳, 罗喜平, 等, 2017. 大数据在贵州专业气象服务的应用及展望. 气象科技进展, 7(2): 54-59.
- 田兰, 熊伟, 赵秋云, 等, 2019. 面向服务的气象大数据集成与共享技术研究. 中国科技成果, 20(13): 78-79.
- 王海成, 2016. 大数据时代政府管理的应对策略. 哈尔滨学院学报, 37(6): 41-43.
- 于庚康, 申双和, 罗艳, 等, 2012. 基于江苏省公众气象服务效益的分析与研究. 气象, 38(12): 1546-1553.
- 喻迎春, 吴楠, 黄震宇, 等, 2018. 基于大数据的公共气象服务需求分析与效益评估. 气象与减灾研究, 41(3): 68-72.

(作者单位: 上海市气象服务中心)