

Research on Recommendation Algorithm and Communication Strategy of Science and Technology Headlines

Liangwei Dong, Jinluan Ren*, Rui Huang, Lihua Liu, Lin Cai

School of Economics and Management, Communication University of China, Beijing
Email: liangweidong94@163.com, *jinluan_ren@163.com

Received: Feb. 18th, 2019; accepted: Mar. 4th, 2019; published: Mar. 12th, 2019

Abstract

The mobile Internet and the new media based on it provide a new model for the dissemination of scientific and technological achievements. Today's headlines make it unique in new media platforms by providing customized news to specific groups based on recommendation algorithms. This paper analyzes the running status and existing problems of science and technology headlines in today's headline platform, and systematically studies the text analysis, user tags and personalized recommendation in today's headline recommendation algorithm, then constructs the operation strategy model of science and technology headlines. Finally, suggestions are made on how to improve the dissemination effect of scientific and technological achievements through headlines.

Keywords

Science and Technology Communication, Headline, Recommendation Algorithm, Communication Strategy

科技类头条号推荐算法与传播策略研究

董良威, 任锦鸾*, 黄锐, 刘丽华, 蔡霖

中国传媒大学经济与管理学院, 北京
Email: liangweidong94@163.com, *jinluan_ren@163.com

收稿日期: 2019年2月18日; 录用日期: 2019年3月4日; 发布日期: 2019年3月12日

摘要

移动互联网以及基于此发展起来的新媒体为科技成果的传播提供了新的模式。“今日头条”以推荐算法为基础*通讯作者。

向特定人群提供定制化新闻资讯的形式使得其在新媒体平台中独具一格。本文分析了“今日头条”平台中科技类头条号运行现状和存在问题,对“今日头条”推荐算法中的文本分析、用户标签、个性化推荐三个方面进行了系统研究,构建了科技类头条号运营策略模型,最后为如何通过头条号提高科技成果的传播效果提出了建议。

关键词

科技传播, 今日头条, 推荐算法, 传播策略

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着移动通信技术和数字技术的发展,传播模式已经迈入了全媒体时代。全媒体以其巨大的信息量、丰富的前沿知识、多样的表现形式、强大的互动性等优势在科技传播中所发挥的作用越来越大,如何利用全媒体时代的传播工具和平台进行科技成果传播已经成为相关机构所面临的问题。

全媒体时代越来越多的网民从 PC 端向移动端迁徙,中国互联网协会发布的《中国互联网发展报告 2018》显示,截至 2017 年年底,中国网民规模达 7.72 亿,其中手机网民规模达 7.53 亿,上网人群占比由 2016 年的 95.1% 提升至 97.5% [1]。与 PC 时代相比,移动互联网时代降低了网民获取新闻资讯的门槛,移动新闻资讯 APP 已经成为网民获取信息的主要渠道。2017 年 6 月,移动端新闻资讯服务月独立设备数达到 6.2 亿,增长了 34.8%,在移动网民中的渗透率由 10.5% 上升至 55.5% [2]。

诞生于移动互联网时代的新兴新闻资讯平台如今日头条、天天快报等,依靠其天生适应小屏、碎片化阅读的优势,通过精准的推荐算法,吸引了大量忠实用户。“今日头条”凭借其在推荐算法上的优势,在新闻移动综合资讯平台领域占有大量且稳定的市场,成为在全媒体时代传播科技成果的重要平台。陈昌凤等(2017)以 BuzzFeed、今日头条为代表,分析了国内外新闻聚合类媒体的发展现状与运作逻辑,探讨新闻聚合类媒体如何影响新闻生产流程、分发渠道以及受众的内容消费方式[3]。王熙元等(2017)从内容资源来源、内容资源组织和内容资源盈利三个角度来分析今日头条[4]。郭全中等(2016)以今日头条为例,提出了一种智能传播平台的构建模式[5]。

科技信息传播已经受到了人们的关注,但是如何更好地促进科技成果在移动端上的传播还缺少一套具体可行的策划方案;将科技传播与移动端结合起来的研究还比较少。尽管有一些相关研究者对“今日头条”的推荐算法进行分析,但单独针对科技成果如何在“今日头条”进行更好地传播还未有研究。本文首先通过“今日头条”科技类头条号发展现状,找到目前发展中存在的问题,然后结合“今日头条”目前的推荐算法给出解决方案,最后提出了科技类头条号的运营策略。

2. 科技类头条号运行现状与面临挑战分析

今日头条将自媒体头条号分成 22 类,其中数码科技类专门用来收集了科技类文章,因此本文将作为研究对象。在今日头条媒体实验室媒体榜单采集了 2018 年 7 月份榜单排名前一百的头条号数据,以此作为研究基础。

1、科技类头条号媒体指数总体处于中游偏下水平

在今日头条,系统依据媒体指数¹进行头条号排名。图 1 给出了 2018 年 7 月份今日头条各类型头条

¹媒体指数是根据头条号文章的阅读量、发文章量、粉丝数、点赞数、评论数、转发数等指标计算得出,通过该指标反映头条号受关注的程度。

号前一百名的平均媒体指数，不难发现除拥有强大资本推动的 MCN (Multi-Channel Network, 多频道网络)和特殊形式的视频类型外，其它头条号媒体指数大多集中于 100~400 之间，而科技类头条号处于中游偏下位置，与娱乐、体育、游戏这些更加贴近生活的头条号相比，其受关注的程度有一定差距。

2018年7月份各类型头条号榜单前一百平均媒体指数

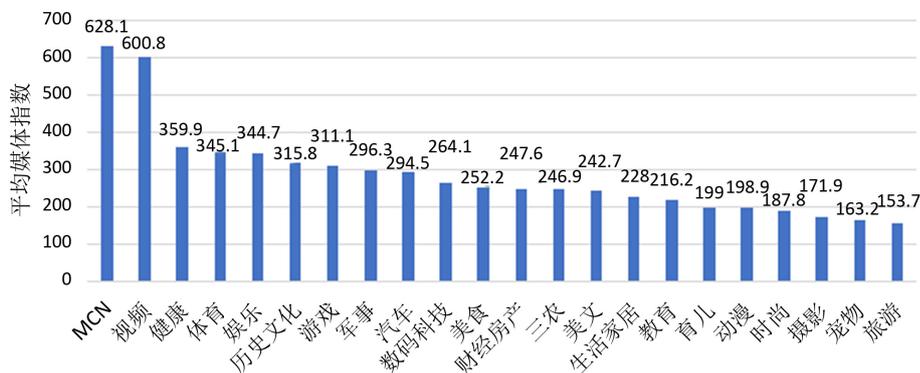


Figure 1. Top 100 average media index in July 2018

图 1. 2018 年 7 月份各类型头条号榜单前一百平均媒体指数

科技领域总体公众关注度不高。通过对比媒体指数发现，科技类头条号的媒体指数与榜单排名在前的头条号还是有一定差距，这一现象与公众对科技类头条号的兴趣和关注度较低有一定关系。公众普遍更多愿意关注贴近生活的文章和头条号，所以对科技领域关注并不多。

2、科技类头条号日均发文量较高

图 2 汇总了各类型头条号 2018 年 7 月份前一百名的日均发文量，相对来说科技类头条号日均发文量较高，保持 24.7 篇日均发文量，这一数据超过了游戏、娱乐等类型的头条号，可见科技类内容还是相对丰富。针对 7 月科技头条号榜单分析，发现科技类头条号日均发文最高可达 763.3 篇，最低为 0.45 篇，其他头条号基本都保持在日均发文 1~10 篇之间，其中保持较高发文量的头条号，其媒体指数排名也较高，当然也有特例，“盖世汽车”虽然日均发文量高达 101.3 篇，但是其媒体指数仅有 200.3，可见在高频发文的情况下，也不一定能够大幅度提升媒体指数。

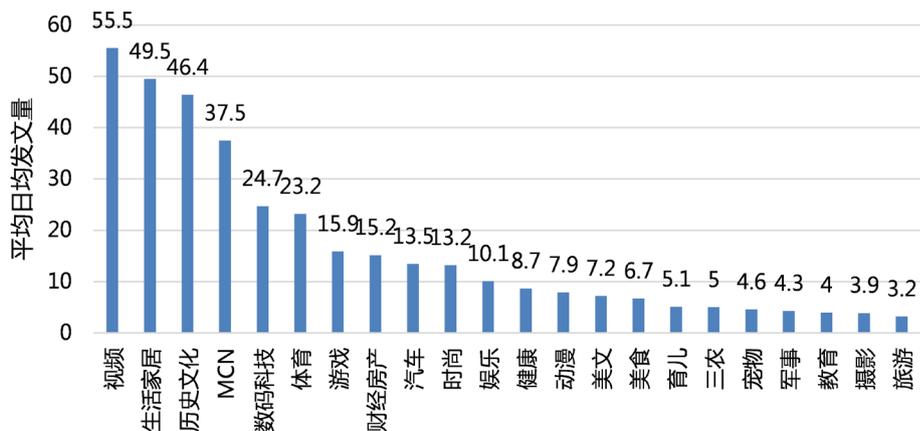


Figure 2. Average daily circulation of the top 100 papers in July 2018

图 2. 2018 年 7 月份各类型头条号榜单前一百平均日均发文量

科技类头条号文章量足质不高。目前，科技类头条号的发文量还是有一定保证的，但是其媒体指数

并没有与发文量成正比,除了公众对科技领域关注不高外,科技类头条号的发文质量也是有一定影响的。目前科技类文章存在较多“标题党”,并没有发出实际有用文章,导致其文章阅读量、分享量较低。在科技类头条号第1名与第100名对比也能够看出,两者在其他数据方面并没有成倍的差距,然而在粉丝数方面却存在十倍的差距,这与头条号发文质量有很大关系,发文内容质量高,会吸引更多粉丝关注。

3、科技类头条号媒体指数水平集中, 优劣差距大

从图3可以看出,媒体指数大于400的头条号仅有6个,通过查询2018年2月~7月榜单发现,创业家、36氪、数码周刊等少数几个头条号能一直保持较高媒体指数。除此之外其它头条号媒体指数多集中于200~300之间,可以看出科技类头条号总体差距不大,真正做出较好传播效果的科技头条号并不多。

数码科技类头条号媒体指数分布

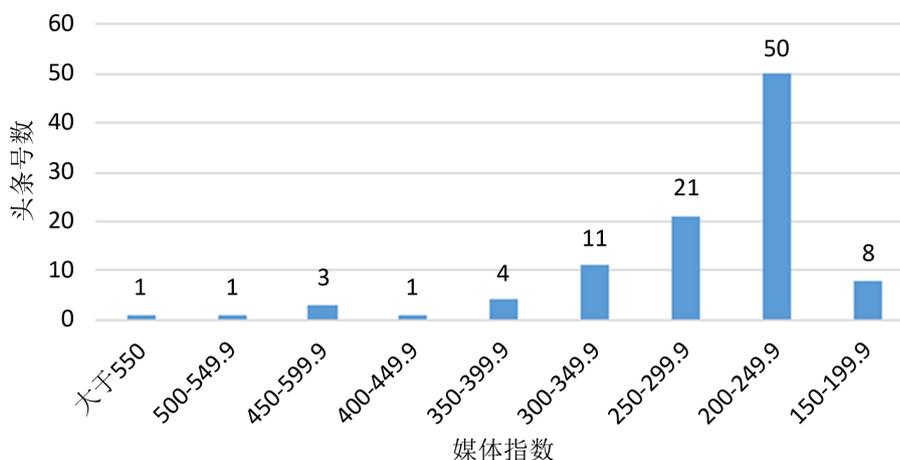


Figure 3. Index distribution of top media in digital science and technology
图3. 数码科技类头条号媒体指数分布

优质科技类头条号较少。经过对数码科技类头条号媒体榜单的分析,不难发现其中媒体指数能够超过400的仅有6个头条号,然而目前科技类头条号数量并不少,大多头条号处在相同境地,如何能突破重围,是科技类头条号需要考虑的问题。对比榜单第1与第100名的数据,其中粉丝数差距非常明显,头条号粉丝往往能在文章阅读量、评论量、转发量等方面做出巨大贡献,如何拥有大量粉丝也是科技类头条号需要考虑的问题。

3. 头条号推荐算法分析

经过上文对科技类头条号现状的研究,本章希望通过对今日头条推荐算法的研究分析,找到一套可行的运营策略,帮助科技类头条号提高影响力,提高其文章的宣传能力,达到良好的传播效果。

今日头条推荐算法在向用户推荐文章时主要从文章内容、用户标签和环境标签上进行分析和匹配,下面就推荐算法的内容分析、用户标签分析和环境分析进行探讨,根据探讨结果为科技类头条号提出运营策略的建议。

3.1. 头条号内容分析

内容分析包括文本分析,图片分析和视频分析。大多头条号以文字咨询为主,所以本文主要梳理今日头条的文本分析。通过对内容进行文本分析,给每个文本内容打上相应的标签,通过文本标签与用户标签进行匹配,可以更好地提高推荐精准度。

图 4 梳理了今日头条对内容分析的流程。一篇文章由头条号创作者产生后，发布到平台，平台会对其进行文本分析。文章的标签分为两类，一类是显式语义特征，这类特征标签是由人为提前设定好一个标签体系，通过对文章的分析，将文章分配到不同的标签体系下。另一种是隐式语义特征，这类标签是根据文章自身特点所生成的标签，包括主题词和关键词。主题词是指文章中出现次数较高的词，是对词频率分布的描述，关键词是对文章进行分析后，对文章特征进行的概括性描述。

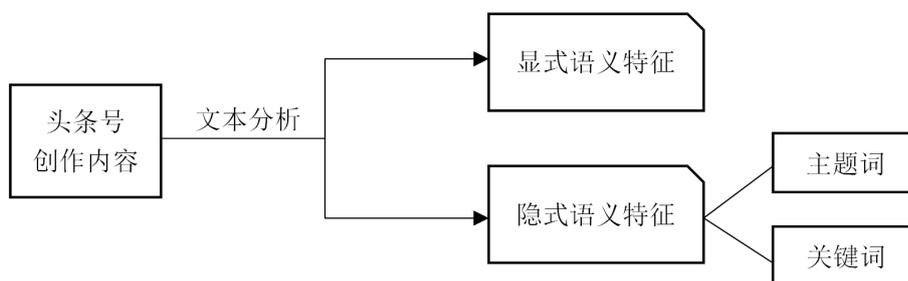


Figure 4. Headline document analysis flow chart
图 4. 头条号文本分析流程图

通过对今日头条文本分析的流程梳理，不难发现对于一篇文章，往往其高频词能够决定文章的标签，这也是文本分析应用的主要逻辑。所以对于头条号文章撰写者而言，完全可以通过在写作中的人工控制，使得推荐算法对文章的分析结果与预期一样。

3.2. 头条号用户分析

每个用户进入今日头条后，系统会对其进行分析，打上属于每个用户自己的独特标签，进而进行个性化推荐。内容分析和用户特征是推荐系统的两大基石，其中用户标签的工程量更为巨大。

用户在刚进入“今日头条”APP时，系统会要求用户登录，用户往往会选用微信、QQ等第三方社交账号登录，在用户登录后，后台会自动获取用户在其它社交账号的一些基本信息，如性别、年龄等，这是用户的基本身份特征，通过这一方式，可以很好化解新用户的冷启动问题，对未在APP产生浏览行为的新用户推荐信息时提供了基础的特征。在用户开始使用APP后，后台会不断获取用户的行为，进而产生用户兴趣特征，如图5所示。

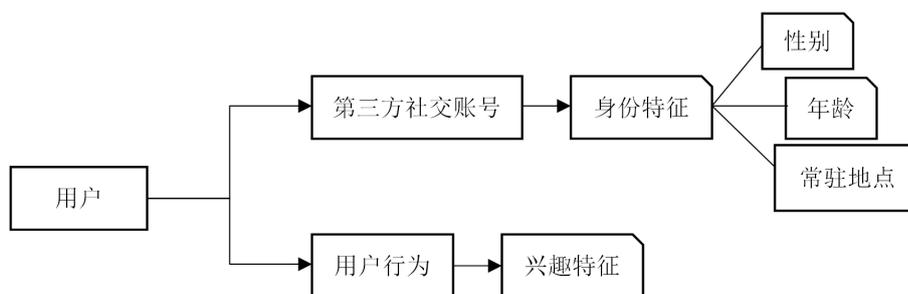


Figure 5. Today's headline user label analysis flow chart
图 5. 今日头条用户标签分析流程图

其中，对用户行为进行分析是用户标签的关键，今日头条主要采用了四种策略保证分析的可靠性。第一是过滤噪声，系统会根据用户停留时间长短，自动过滤掉用户停留时间短的文章；第二是惩罚热点，对于热门事件的相关文章，系统对用户的行为做适当的降权处理，保证能够客观地获取用户兴趣特征；第三是时间衰减，用户的兴趣不是一直保持不变的，所以系统会根据用户的新动作不断更新用户兴趣标签；第

四是惩罚展现,对于系统推荐给用户的文章,用户却没有点击,那么这篇文章对应的特征标签都会被惩罚。

3.3. 个性化推荐过程

今日头条头条号的个性化推荐系统一共有三个维度,第一个维度是内容,通过对文本内容的分析,为每篇文章打上标签。第二个维度是用户,每个用户都是一个独立个体,通过对其基本信息和行为的分析,构建其专有的兴趣图谱。第三个标签是环境特征,系统会根据用户的环境、场景进行分析,为其推荐符合环境的信息。今日头条个性化推荐过程如图6所示。

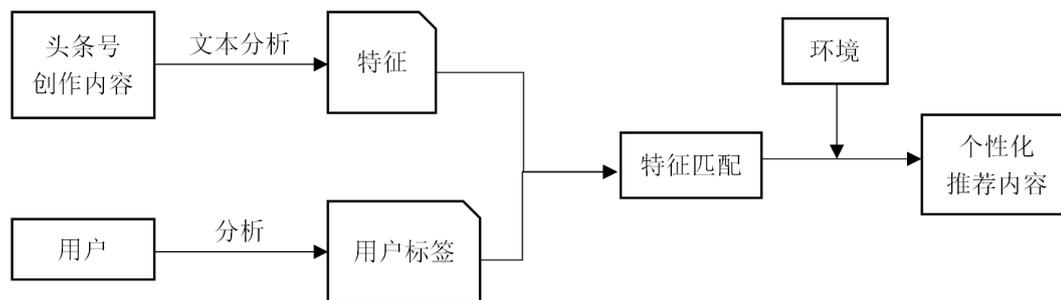


Figure 6. Headline personalized recommendation process

图6. 头条号个性化推荐流程

今日头条推荐系统采取了四种比较典型的推荐特征。第一种是相关性特征,将内容标签与用户标签进行匹配,评估内容是否与用户匹配。第二种是环境特征,这是一种基本特征,根据用户的环境、时间等进行推荐。第三种是热度特征,在推荐时,系统会考虑热门事件,将其推荐给用户,尤其是新用户。第四种是协同特征,系统不仅仅限于考量一个用户特征,而是通过寻找相似用户,根据相似用户的浏览历史,为用户推荐内容。

今日头条个性化推荐实际上就是参考用户所处环境,将文章特征与用户标签进行匹配,如果匹配度较高,就将文章推荐给用户。对于一个公众号的运营者来说,了解了今日头条的推荐流程,能够对其后续工作产生非常大的帮助。

4. 科技类头条号运营策略建议

4.1. 运营策略

通过对“今日头条”推荐算法的分析不难看出,其规则就是通过文章特征与用户特征相匹配,参考环境特征进行推荐,对于头条号运营者来说能够掌控的只有文章的内容,发文内容是科技类文章传播效果的关键。本节提出科技类头条号运营策略,如图7所示供头条号运营者参考。

在准备时期,头条号一方面需要找准自己定位,合理找寻目标用户。用户是头条号的关键,在准备阶段,应该对头条号的目标用户进行分析,获取他们的基本信息,如年龄、性别、职业等,通过基本信息进一步了解目标用户,给出目标用户特征,构建兴趣图谱。另一方面头条号需要在海量的科技咨询中寻找用户感兴趣的,有价值的科技内容。在信息爆炸的时代,存在海量的科技类信息,头条号需要在繁多的信息中,紧跟时代热点,找到能够吸引用户,并且能够创造出有深度文章的科技资讯。

撰写时期是整个头条号运营的关键时期,在这个阶段科技类头条号需要将前期准备的科技资讯通过各种方式产生出深度好文,同时也能匹配目标用户特征。在撰写内容的时候需要注意特征化、趣味化、科普化、碎片化四个方面。特征化是指撰写的内容应该能够拥有明显主题、关键词等特征,并且这些特征还应该符合目标用户特征。第二个趣味化是指科技类文章往往语言枯燥,需要注意在文章撰写过程中,

改变行文风格，用更加贴近生活的表现方式表达科技成果。科普化是指文章拒绝“标题党”，撰写的文章应该拥有充实的内容，真正达到科普的效果。碎片化是指可能一个科技成果需要大幅文字内容，需要将其分解，保证文章篇幅不过大，迎合用户阅读习惯。

在反馈时期，是从粉丝画像和文章数据两个方面进行分析。粉丝画像是今日头条后台给出的头条号粉丝数据，不仅仅包含粉丝基础特征，还有一部分粉丝喜好特征，科技类头条号可运用这些数据，完善自身用户特征体系。文章数据是指文章发表后的相关数据，包括被推荐次数、阅读次数等相关数据以及系统给文章的关键词，头条号分析这些数据，匹配关键词是否与预期特征相同，找到在文章撰写过程中的问题与不足以便进一步完善。

在整个科技类头条号运营中，需要不断更新用户特征值，改进文章撰写方式，提高科技成果传播效果。

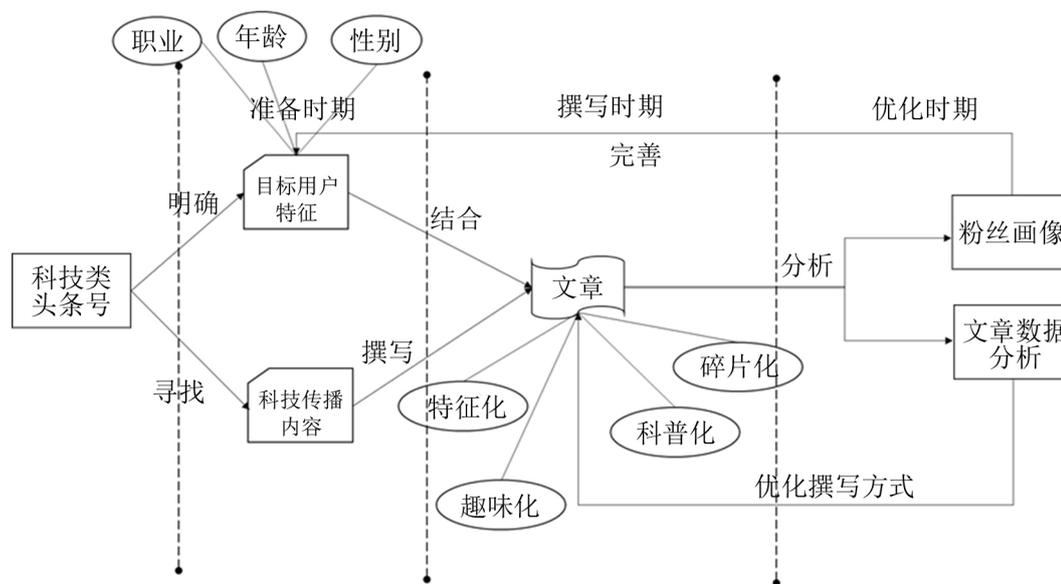


Figure 7. Technological top number operating strategies development process
图 7. 科技类头条号运营策略制定流程

4.2. 发文内容建议

本文在对科技类头条号进行现状分析时，找到科技类头条号面临的问题，结合“今日头条”推荐算法，提出了一套运营策略作为参考，本节结合上述分析，对科技类头条号文章的撰写提出一些建议。

1、了解信息目标用户，构建用户兴趣图谱

今日头条推荐机制是通过文章特征与用户特征进行匹配进行推荐，所写文章如果与用户特征匹配，其被推荐率可能会大大增加。每条科技成果和科技宣传信息针对的目标用户是不同的，在每次进行内容创作前，应该充分了解目标用户群，构建目标用户的兴趣图谱，通过对用户特征进行分析，创作出符合目标用户特征的内容，可在文章中适当增加符合用户特征的关键词，保证文章通过推荐算法分析后，文章特征与用户特征高度匹配。

2、明确文章主题，突出文章内容

今日头条对头条号文章的分析及特征匹配，大多以词频为关键因素。在撰写头条号文章时，作者需要将文章定位、特征考虑清楚，结合目标用户特征，确定文章主题，撰写文章时紧密围绕文章目标主题撰写，文章内容多出现期望主题词以及与期望主题词相关词语，这样能够保证文章的特征词与自己的预期相同，更容易借助推荐算法将内容传播到用户。

3、内容贴近生活，主题吸引眼球

从各个类型头条号的媒体指数可以发现，类似于游戏、娱乐这类贴近生活，且内容有趣的头条号，往往会有很高的媒体指数，从众多头条号中脱颖而出。根据用户习惯而言，大多数用户往往更热衷于热点、有趣的文章，对于足够吸引眼球的题目往往有浏览欲望，因此，在内容创作上需要足够吸引眼球。在创作内容时，可以尽量将无趣的科技报道通过平实语言表达出来，用通俗语言代替专业术语，不仅能够保证内容得到更多的传播，也能够保证读者明白理解，提高了传播效果。通过有趣的文字表达形式，也能够更加吸引用户关注科技领域，提高用户对科技领域的总体关注。

4、高频发文下保证文章质量，利用内容增加粉丝数

今日头条推荐机制对于用户行为很重视，这意味着如果一个头条号创作的文章被用户点开，那么这个头条号的其他文章被推荐给用户的概率也会增大，所以，一个头条号需要保证一定的发文量，能够增加用户粘性，并且，文章题目在制定时，非常重要，好的题目往往能够吸引用户眼球。当然，在保证发文量的同时，如果太多低质量文章，系统会因用户停留时间较短，降低权重，以至于高发文量也无法得到很好的阅读量。通过榜单第一名与最后一名对比也不难看出，拥有高质量的文章往往会增加用户对头条号的关注与喜爱，更容易吸引粉丝，而粉丝数往往也决定了头条号文章的被阅读数。

基金项目

北京市科学技术委员会科技专项(Z181100000718009): 设计领域知识产权保护工作研究。2018.8~2019.3;
中国传媒大学研究生培育项目(Z181100000718009): 基于“今日头条”传播平台的科技成果全媒体宣传策略研究。2018.6~2019.1。

参考文献

- [1] 中国互联网协会. 中国互联网发展报告 2018 [R]. 北京: 2018 中国互联网大会, 2018.
- [2] 艾瑞咨询. 2017 年中国移动端新闻资讯行业报告[R].
- [3] 陈昌凤, 王宇琦. 新闻聚合语境下新闻生产、分发渠道与内容消费的变革[J]. 中国出版, 2017(12): 3-7.
- [4] 王熙元, 包韞慧. 今日头条新闻客户端内容资源分析[J]. 出版广角, 2017(13): 46-48.
- [5] 郭全中, 胡洁. 智能传播平台的构建——以今日头条为例[J]. 新闻爱好者, 2016(6): 4-8.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7908, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: ssem@hanspub.org