

国内互联网异常流量白皮书

2020

GRIDSUM 国双

摘要



基于国双Gridsum Ad Dissector 全年监测的互联网广告数据，2020年全年异常流量均值为**29.4%**，GIVT占比均值为**2.7%**，较前几年异常流量总体呈现明显下降趋势。

01 背景介绍

国内外广告市场现状

根据WARC 《Gload Ad Trends: State of The Industy 2020/21》显示，2020年全球广告支出预计下降10.2%至5573亿美元，下降约634亿美元，网络广告也在2020年首次停止增长，2020年网络广告市场支出约占全球广告支出的54.4%，略微下降0.3%，总额为3033亿美元。2020年全球广告业支出与2019年相比减少11%（约为680亿美元）。

2020年在全球经济衰退的大背景下，国内媒体的广告收入下降6%，这是MAGNA广告支出记录以来的首次下降，国内广告支出总额将回落至5180亿人民币。国内的数字广告销售收入同比仍保持平稳，并推动广告收入的增长。其中，数字广告支出达3570亿人民币，占广告总支出的69%。



国内异常流量反作弊现状

2020年5月13日，最高人民法院举行新闻发布会，发布了人民法院大力弘扬社会主义核心价值观的十大典型民事案例，其中“暗刷流量”合同无效案入选。该案是全国首例涉及“暗刷流量”虚增网站点击量的案件。网络产品的真实流量能够反映出网络产品的受欢迎程度及质量优劣程度，流量成为网络用户选择网络产品的重要因素。而“暗刷流量”的行为违反商业道德，违背诚实信用原则，对行业正常经营秩序以及消费者的合法权益均构成侵害，有损社会公共利益。该案对“暗刷流量”交易行为的效力予以否定性评价，并给予妥当的制裁和惩戒，对治理互联网领域内的乱象有积极推动作用。

当下，行业内各方都在积极对抗异常流量，为净化互联网广告环境而努力。作为独立的第三方监测公司、国内广告反欺诈的先行者，国双有责任与义务为广告主和整个行业监测并识别出作弊行为和异常流量。同时，作为中国广告协会理事单位、中国广告协会互联网广告委员会常务委员单位，国双从2017年起就与行业协会、同行者一起推动异常流量监测的标准化和规范化，促进行业的健康发展。

此次，在中国广告协会的指导下，国双连续第四年制作并发布了《国内互联网广告异常流量白皮书》，旨在揭示国内的互联网广告流量现状，为广告主、代理公司等行业参与者提供直观的数据参考，同时展示国双多年来不断完善的异常流量监测方案。

DIF中广协异常流量工作组动态

Distributed Invalid Traffic Filter (分布式异常流量过滤器), 简称DIF联盟链, 是由中国广告协会统筹指导, 中国无线营销联盟 (MMA中国) 发起, 由各广告科技企业进行技术落地研发的开放源代码项目, 为解决互联网广告流量痛点开辟了全新的道路。成员可提交广告活动中, 依据中华人民共和国《互动广告: 投放监测及验证要求》国标中一般异常流量标准, 识别出来的IP和移动装置等数据。

成员亦可对自认为被误判的数据提出申诉, 申诉结果通过链上投票的方式进行处理。具有媒体属性的成员, 也可认领被误报的媒体服务器IP。统一的数据考察和应用期, 成员可利用自己的客户端从DIF区块链上下载最新版GIVT数据, 应用在广告系统的生产功能中。DIF将按照联盟规章制度, 透明、公正、自动地合并处理各成员提交的数据。

代码即法律, 利用区块链的技术特性实现共识。国双作为DIF核心成员, 通过真实广告投放环境异常流量监测, 洞察IP和移动装置等黑产业链数据, 并在每月积极上传DIF, 为行业提供有效反作弊保护。



互联网广告技术实验室动态



中国信通院与中国广告协会联合创建互联网广告技术实验室，致力于打击互联网广告数据造假和作弊行为，营造透明真实的广告经营环境，建设互联网广告绿色生态环境，促进互联网产业的科学发展。

2021年，互联网广告技术实验室将加大移动互联网广告标准体系研究，梳理国际互联网广告组织、标准化情况，提出国内互联网广告标准化体系规划，并推出互联网广告异常流量审计技术要求，全面净化国内广告环境。

国双作为中国广告协会核心成员，信通院标准小组成员，积极参与互联网广告技术实验室相关工作，积极推动中国异常流量识别，打击行业广告流量数据造假活动。



02 异常流量行业动态

TAG异常流量审计动态

TAG (The Trustworthy Accountability Group) 作为一家自律组织，于2014年由美国广告业组织、互动广告局 (IAB)、美国全国广告主协会 (ANA) 和美国广告代理协会(4A's)共同创办。参与企业共同创建消除流量作弊、反广告恶意软件等四项自律标准，每项标准都有一个市场交易印章，企业可以向合作伙伴、客户和竞争对手展示，以证明其坚持使用高标准来应对市场挑战。CAF，即消除流量作弊项目，致力于提供一种有效打击数字广告供应链中异常流量的方法来帮助企业打击异常流量犯罪。国双在积极准备MRC审计认证之余，已正式开展TAG消除流量作弊CAF项目认证，严格按照国内外行业审计标准，帮助广告行业减少因异常流量带来的损失。



02 异常流量行业动态

互联网设备标识动态及影响



随着全球用户隐私安全问题的不断发酵，2018年4月，谷歌发布DCM（DoubleClick Campaign Manager）不再提供日志级别带有用户ID的数据回传，仅提供广告曝光与点击数据报告。2020年6月，苹果官方宣布IOS14隐私新规，未来将不再回传IDFA用户标识。互联网设备标识已呈现紧缩趋势，未来是否更多的设备标识不再回传，则需看情况而定。设备标识的回传问题，将极大影响整个广告产业，尤其对于移动营销市场而言，这无疑是一个巨大的变化。鉴于当今的移动再营销，排除定位，受众群体细分，同龄受众和其他功能在很大程度上依赖于设备标识，因此设备标识的开放与否将对通过移动广告获利和用户获取带来巨大的负面影响。2021年，中广协&信通院旗下的CDA互联网广告技术实验室工作规划中提到，推出CAID技术开发与产业对接等工作，CAID将作为行业通用的互联网广告设备标识，用于解决未来设备标识不回传等问题。国双积极参与CAID的推广与应用，且与各头部媒体协定统一的用户设备标识，并与头部媒体共建Datahub来积极应对互联网设备标识不回传所带来的影响。

03 2020年异常流量分析

异常流量标准

- 总体标准

本报告中涉及的异常流量，参考了MRC、IAB、TAG等国际标准，国内权威组织（中国广告协会组织、MMA中国等）制定的国家标准和团体标准，以及中国无线营销联盟对异常流量的定义和分类。根据相关标准异常流量被分为两大类，即GIVT和SIVT。

- 一般异常流量

GIVT是General Invalid Traffic的缩写，即常规异常流量，是指能够通过应用多种名单或标准化参数等常规方式进行过滤的流量。

- 复杂异常流量

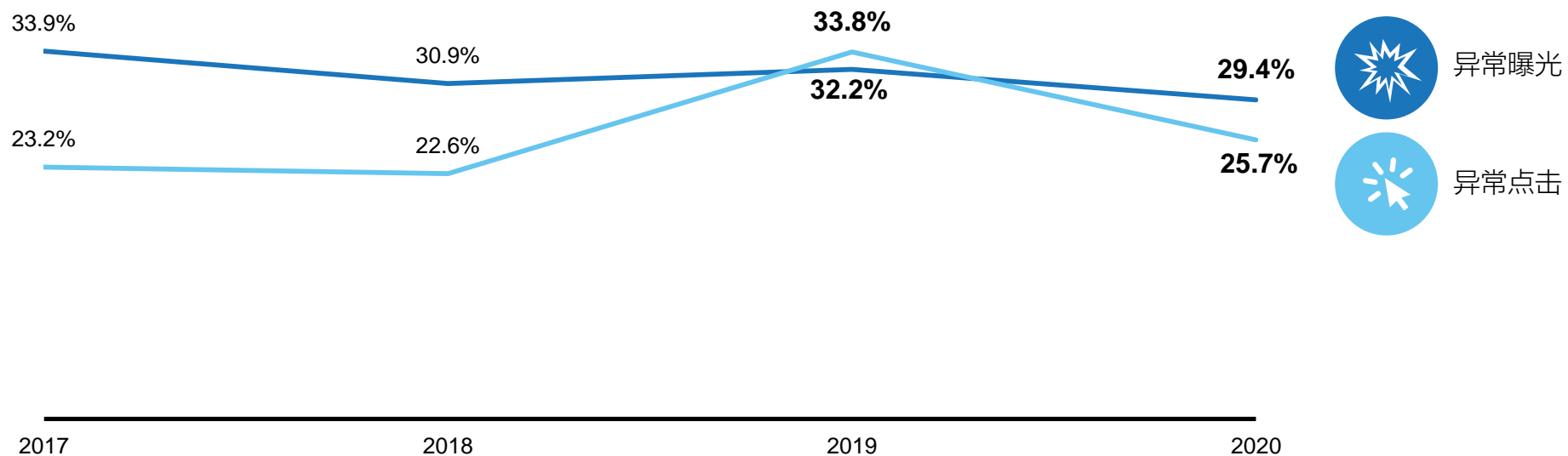
SIVT是Sophisticated Invalid Traffic的缩写，即复杂异常流量，该类型的流量无法通过简单的规则进行识别，需要通过一系列高级分析，甚至人工参与，经过多方合作与协调等方式，才能进行分析与识别。

03 2020年异常流量分析

2020年全年整体分析

2020年互联网广告异常曝光占比**29.4%**，较2019年**降低了2.8%**。异常点击占比降低明显，从**33.8%**降低至**25.7%**。这源于广告行业愈加规范，各利益方对于无效流量给予高度重视。国双认为，异常流量丰富的甄别规则，有效规范了广告行业，促进了行业健康发展。

2017-2020年异常流量对比分析

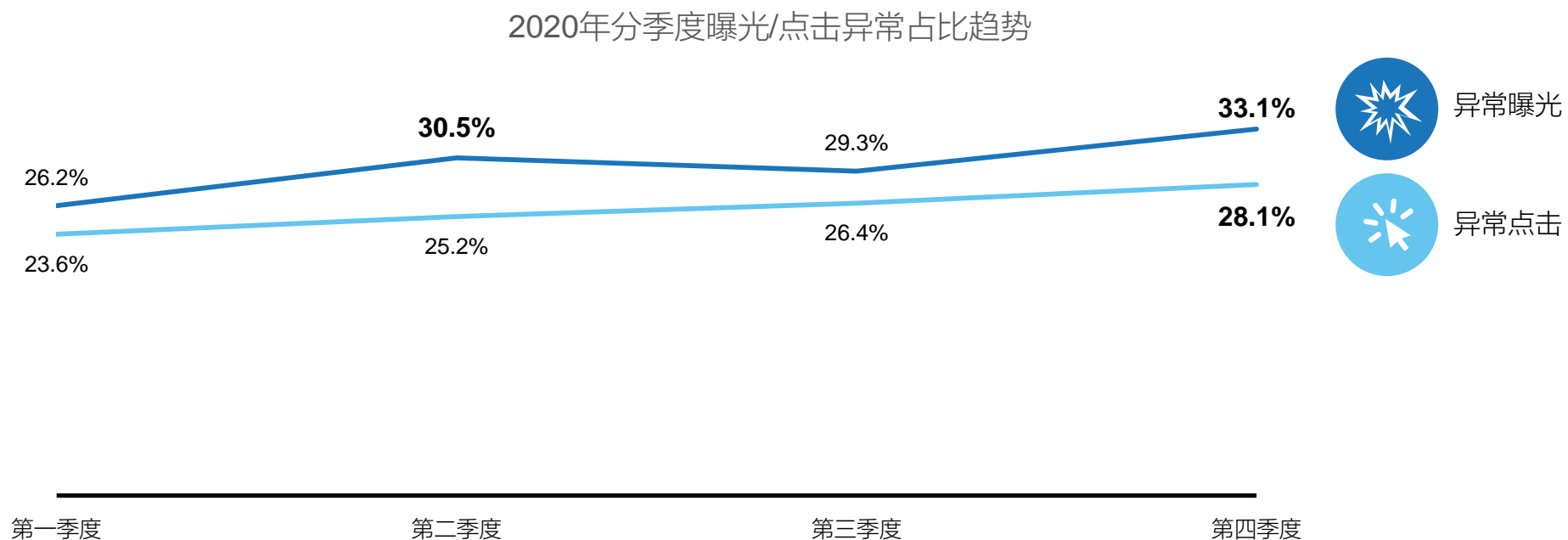


数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

2020年分季度趋势分析

异常曝光在**第二季度、第四季度**相对较高，而异常点击则呈现缓慢上升趋势，**第四季度**出现峰值。

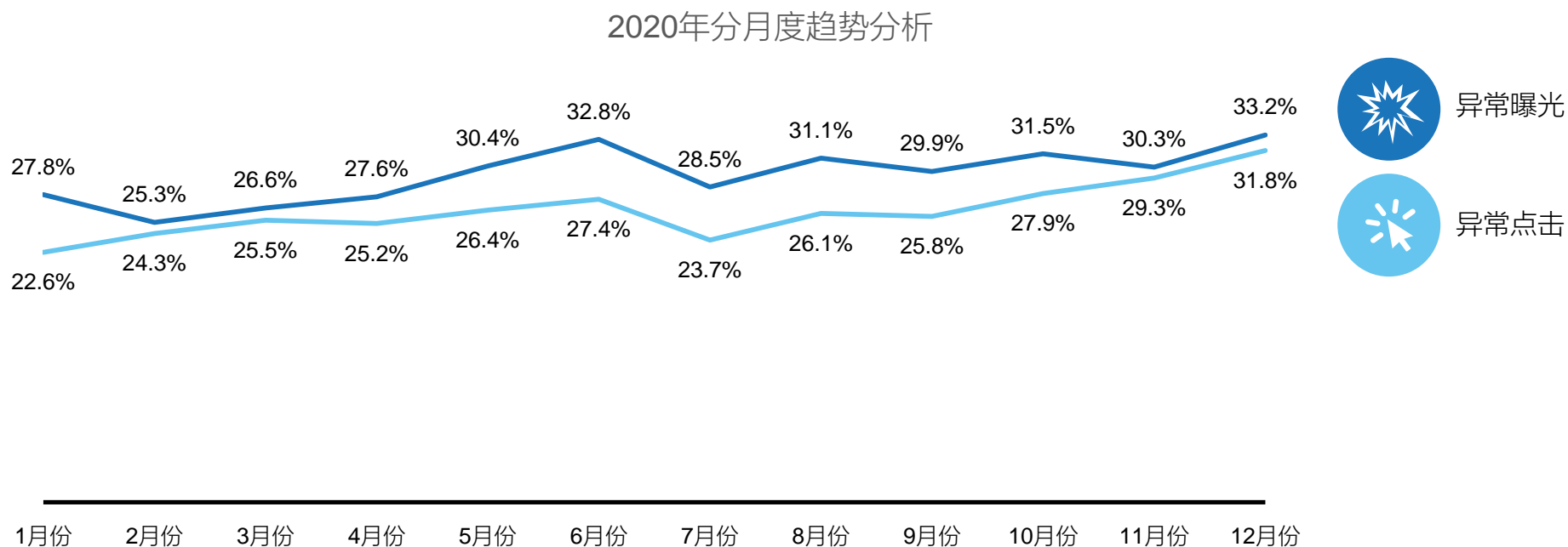


数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

2020年分月度趋势分析

2020年异常报告与异常点击波动较为平稳，整体趋势上，呈现递增，且在10月、11月和12月上升较为明显。国双认为，双11、双12的流量增多与这一趋势有较大的相关性。



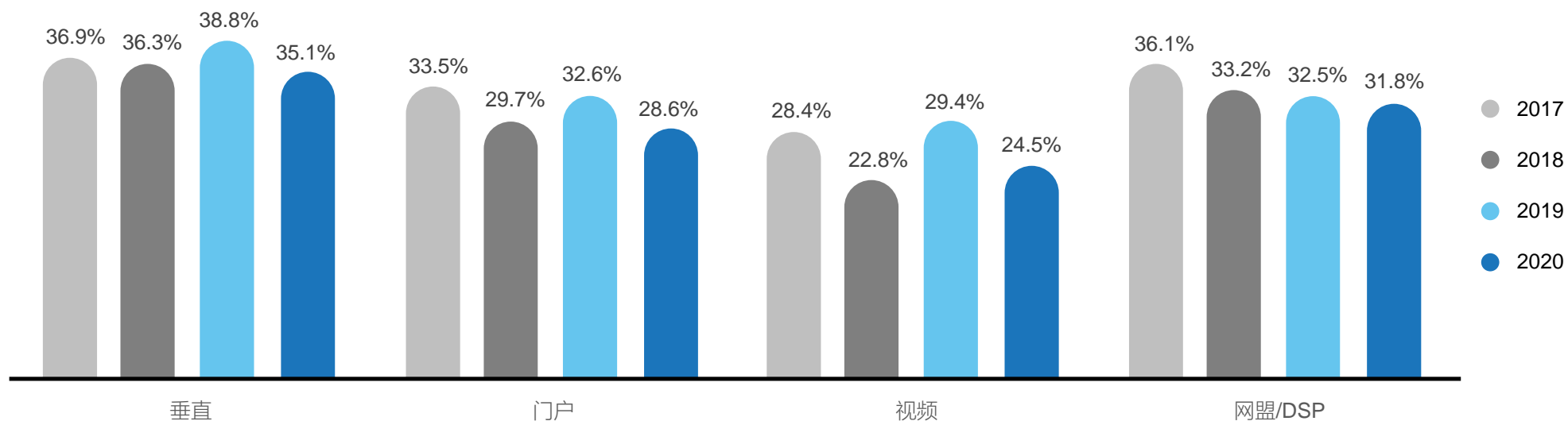
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 分媒体类型-异常曝光

四大媒体类型的异常曝光，整体均有降低，其中视频媒体波动较大。

2017-2020年分媒体类型异常曝光占比



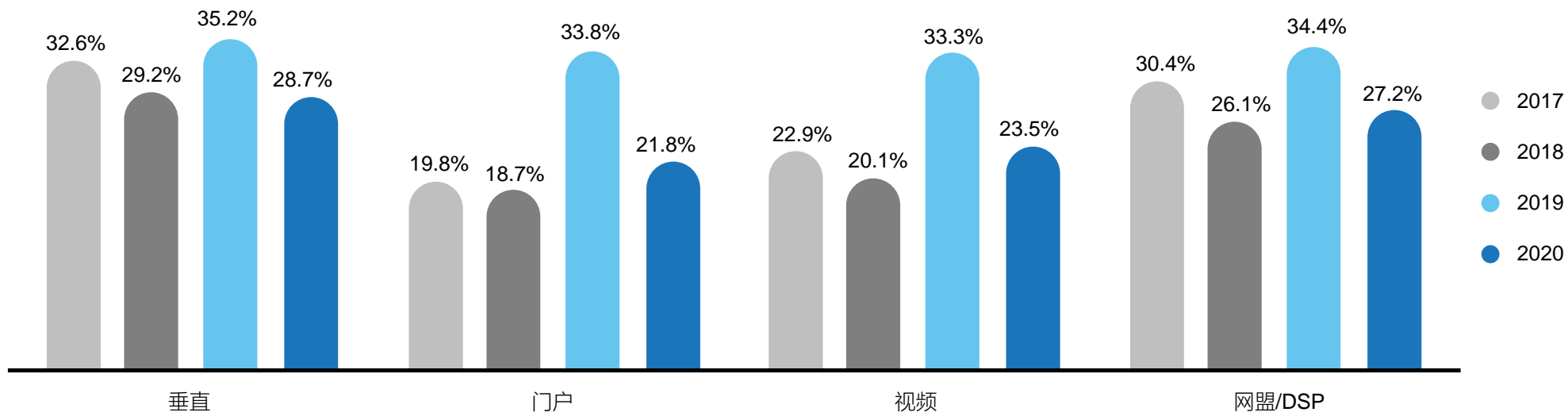
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 分媒体类型-异常点击

在异常点击上，2020年整体较前年均有所下降，且幅度不小，媒体呈现愈加规范的趋势。

2017-2020年分媒体类型异常点击占比



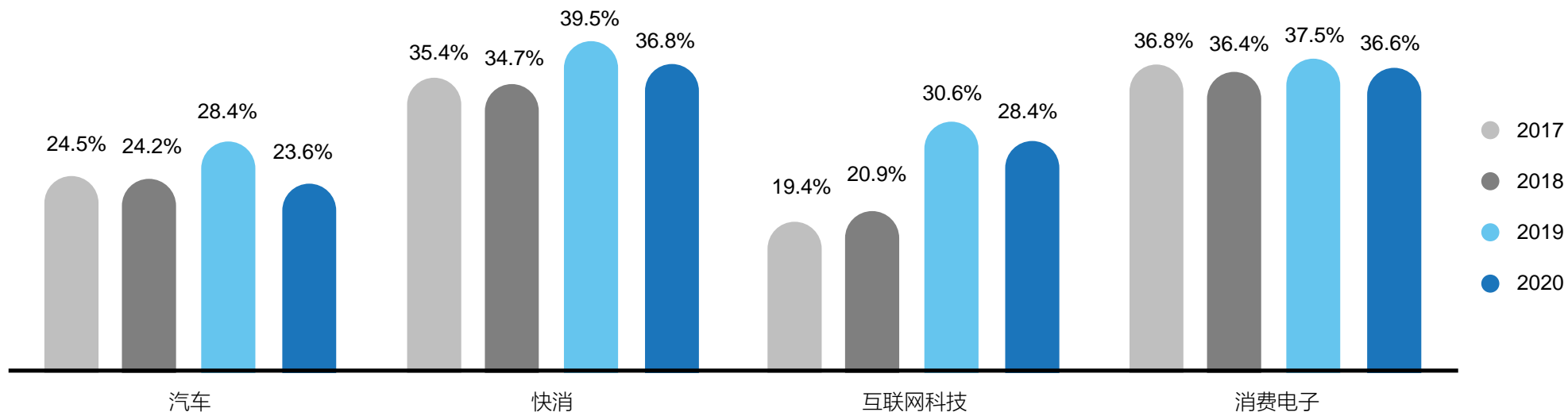
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 细分行业-异常曝光

在国双监测的行业数据中，以消费电子、汽车、快消、互联网科技为例，2020年各个行业的异常曝光数据均有所下降，其中汽车和快消下降幅度最大，这得益于行业组织以及国双在异常流量甄别上作出的努力，严格的异常流量排查能有效约束行业异常流量，规范行业各方行为，有效提升行业环境。

2017-2020年分行业异常曝光占比



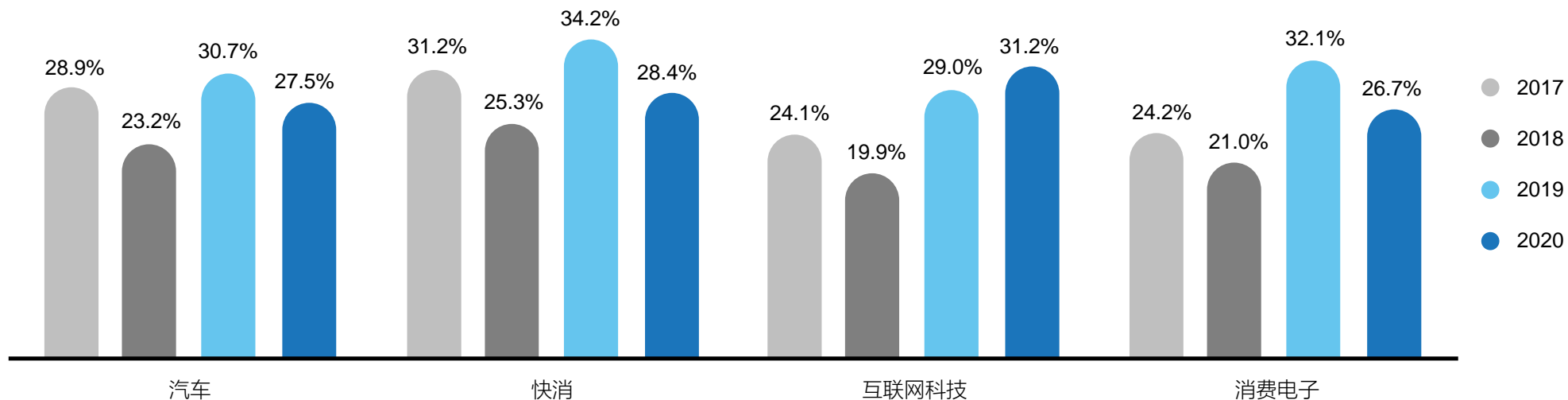
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 细分行业-异常点击

2020年汽车、快消与消费电子行业的异常点击占比明显下降，但互联网科技仍保持上升，国双参考行业统一规范，丰富异常流量识别规则，未来将针对性地来加强行业异常点击识别。

2017-2020年分行业异常点击占比

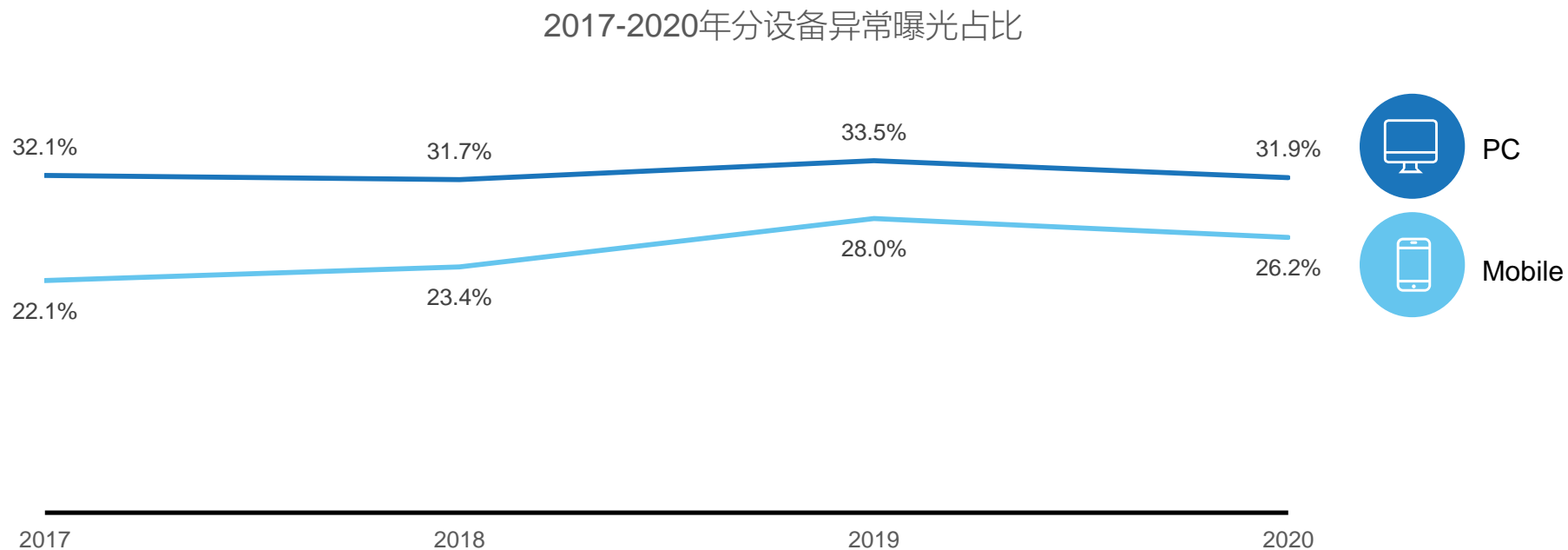


数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 分设备-异常曝光

PC端异常曝光占比高于移动端，趋势较为平稳，整体上呈现下降趋势。但是移动端异常曝光占比在持续三年增长后，首次出现下降趋势，这与国双持续加强异常识别能力有重要关系。

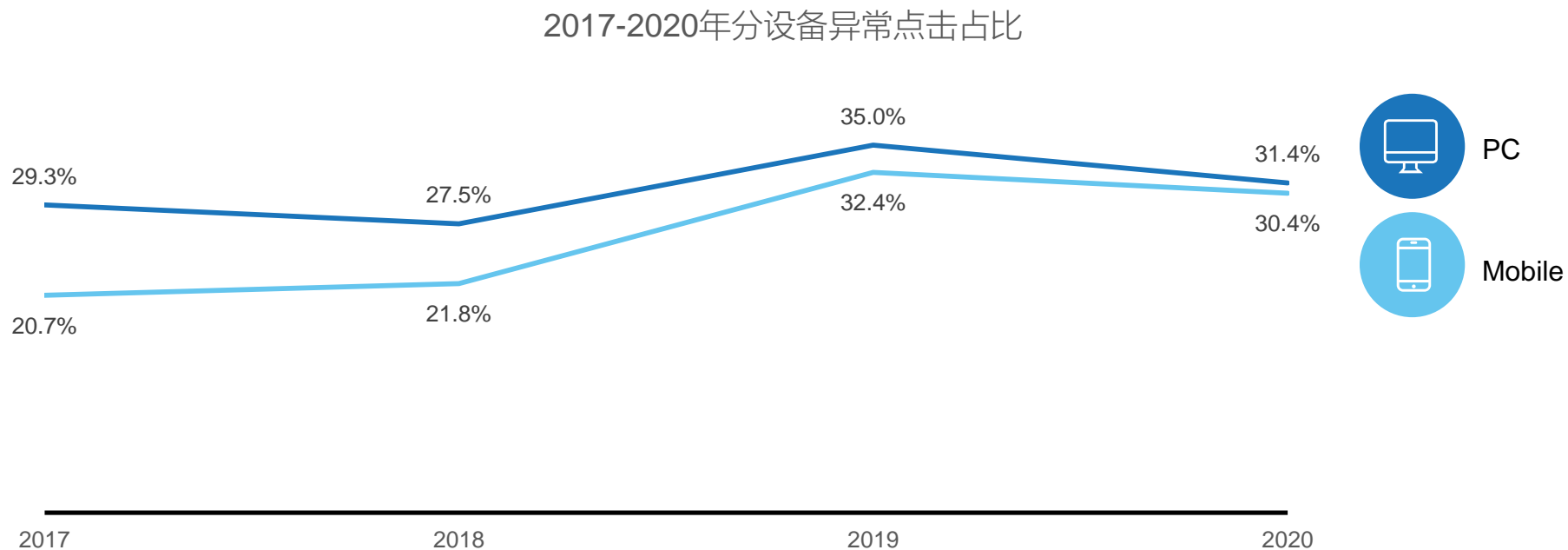


数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

异常流量数据分析 - 分设备-异常点击

异常点击上，PC端与移动端占比均有所减少，其中移动端为30.4%。



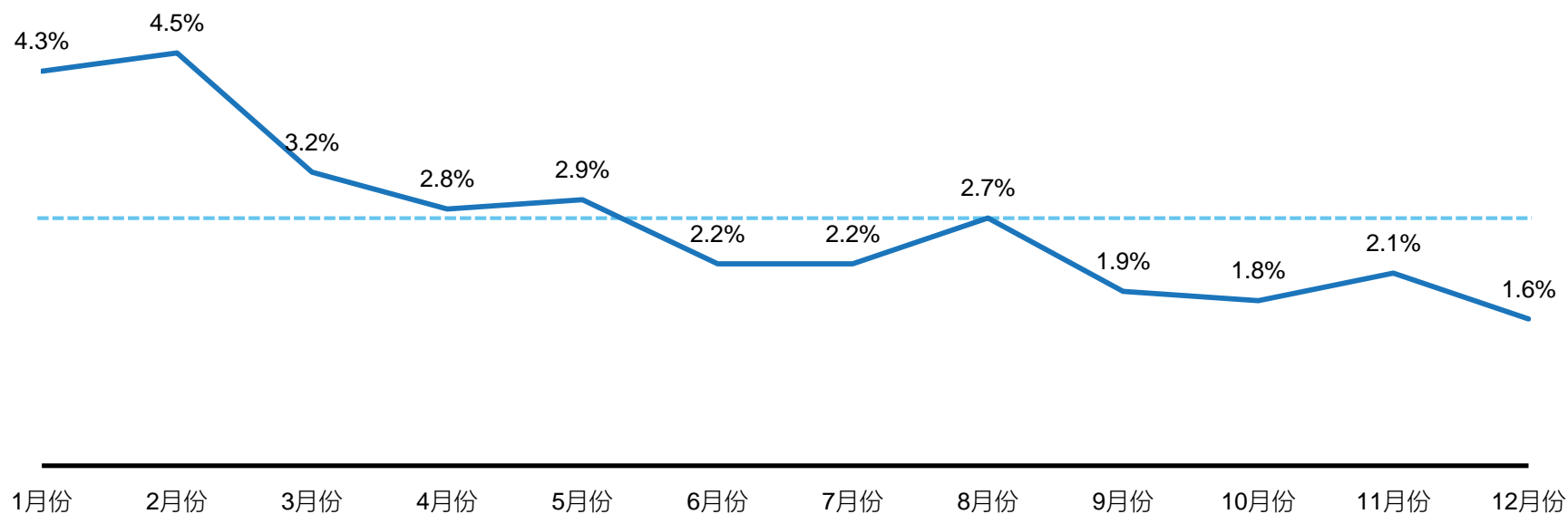
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

GIVT (常规无效流量) 数据分析 - GIVT整体分析

2020年国内互联网广告 GIVT (常规无效流量) 月平均占比为**2.7%**；与全年整体异常流量上升趋势不同，GIVT全年整体呈现下降趋势。

2020年GIVT占比分月趋势



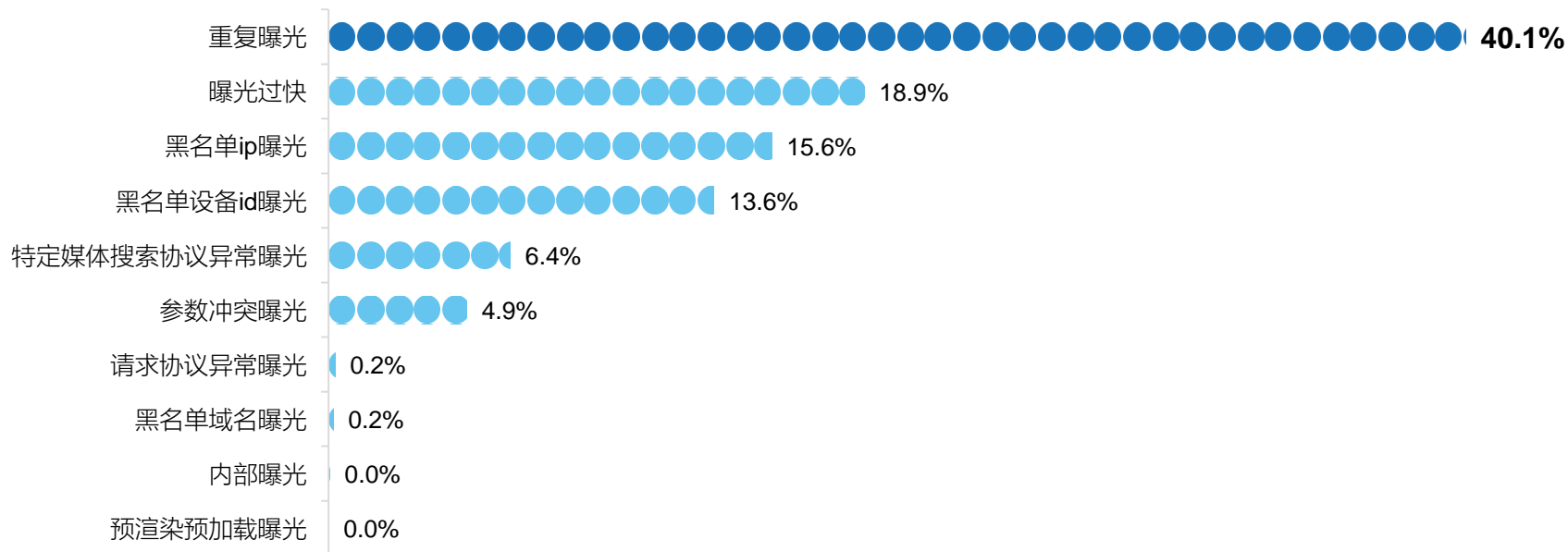
数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

03 2020年异常流量分析

GIVT (常规无效流量) 数据分析 - GIVT-分细项

2020年GIVT流量中，各细项之间占比最高的是重复曝光，占比达到40.1%。

2020年GIVT分细项占比



数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

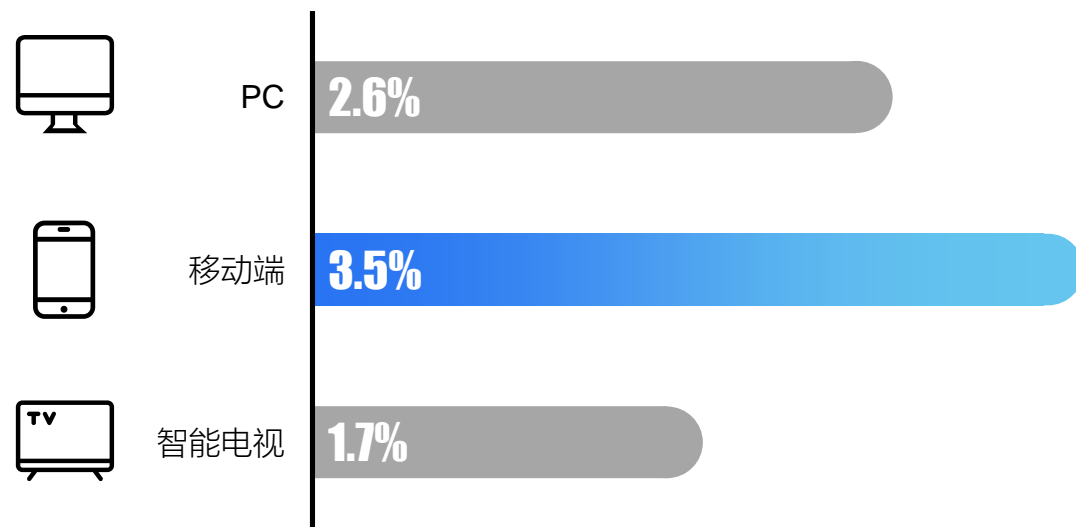
03 2020年异常流量分析

GIVT (常规无效流量) 数据分析 - GIVT-分设备分析

2020年，国双监测数据发现设备占比分别为PC端17%，移动端77%，OTT端6%；

其中国内数字广告GIVT (常规无效流量) 分设备占比PC端最高**3.5%**。

2020年GIVT分设备占比



数据来源：Gridsum Ad Dissector 数据时间：2020年1月1日-2020年12月31日

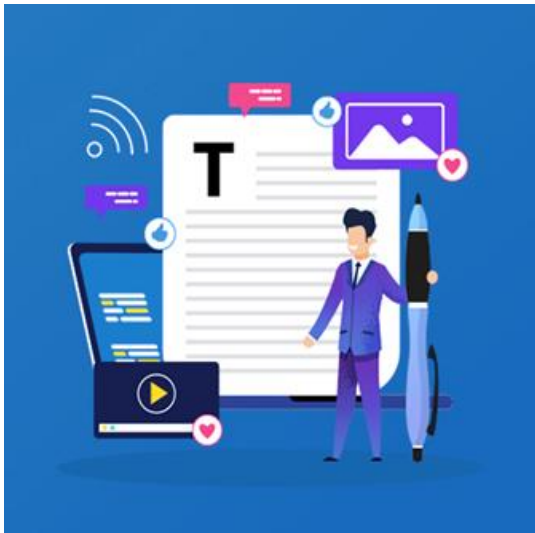
小结

通过对异常流量数据的洞察，在2020年疫情的笼罩下，国内外广告市场受到严重打击；同时，2020年异常流量总体有所下降，总体上广告行业异常流量较前两年有所好转；2020年，行业均值在29.4%，其中包含GIVT与SIVT两部分，GIVT的占比均值为2.7%；对于难以甄别出来的异常流量，需要通过更有效的技术手段来判定。疫情的大环境下，全球广告预算紧缩，广告主对于广告投放的ROI以及流量有效性提出了更高的要求。同时，行业组织在异常流量领域的不断深耕，有效地净化了互联网广告行业环境。

当然，随着黑灰产各类作弊手法会越来越隐蔽，常规或简单的甄别规则将愈加难以识别异常流量。国双在国内外行业组织的指导下，结合国双大数据、人工智能的基因，不断升级异常流量识别方案，用更有效的技术手段，来判定复杂异常流量；国双把保护客户流量作为自己的使命，持续为客户提供全方位的作弊防护。



04 国双软广行业洞察与监测



“ 内容营销是靠文字、图片、视频等不同形式，以让人易于接受的方式来宣传品牌，并结合当下热点，通过讲故事等方式植入品牌，潜移默化地将受众转化为粉丝，形成受众范围内的口碑传播，最终占据用户心智。

随着内容营销对社会和公众的影响不断增大，出现以网络炒作为业务的网络公关公司、推手、水军等。网络公关公司以营利为目的，为炒作某个话题、人物或产品，雇用大量的网络水军；在网络推手的组织下，短时期内在互联网的各大网络论坛上大量发布煽动性帖子，并通过跟帖、转帖及编辑评论等方式炮制热点事件，营造虚假民意，从而形成虚假网络舆情。网络公关公司、推手、水军等形成了灰色利益链。

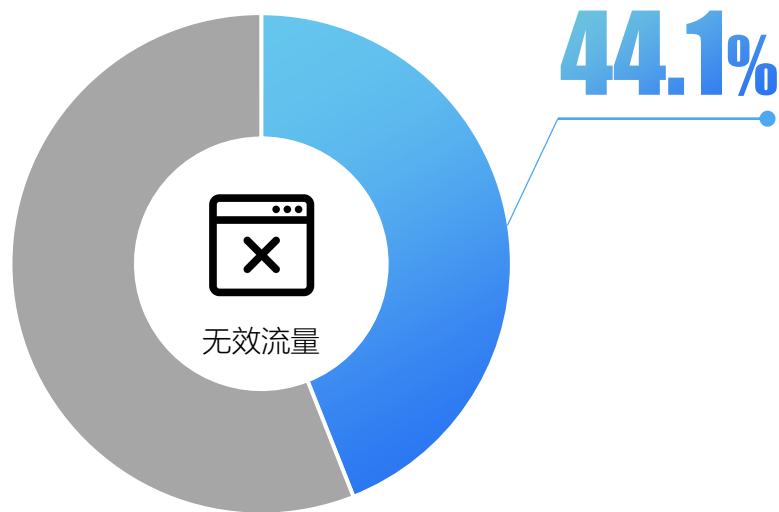
针对此类现状，广告主很难甄别内容营销活动的真实效果。为解决该问题，国双在合法范围内，通过多种技术手段实现对内容营销的效果跟踪，通过对来源网络信息进行实时监测和采集，利用自然语言处理、数据挖掘、计算语言学、机器学习、人工智能等技术对数据进行解析与处理，有效识别虚假热帖、网络水军，为广告主提供内容营销活动真实效果的数据支持。

04 国双软广行业洞察与监测

2020年内容营销监测-无效流量整体分析

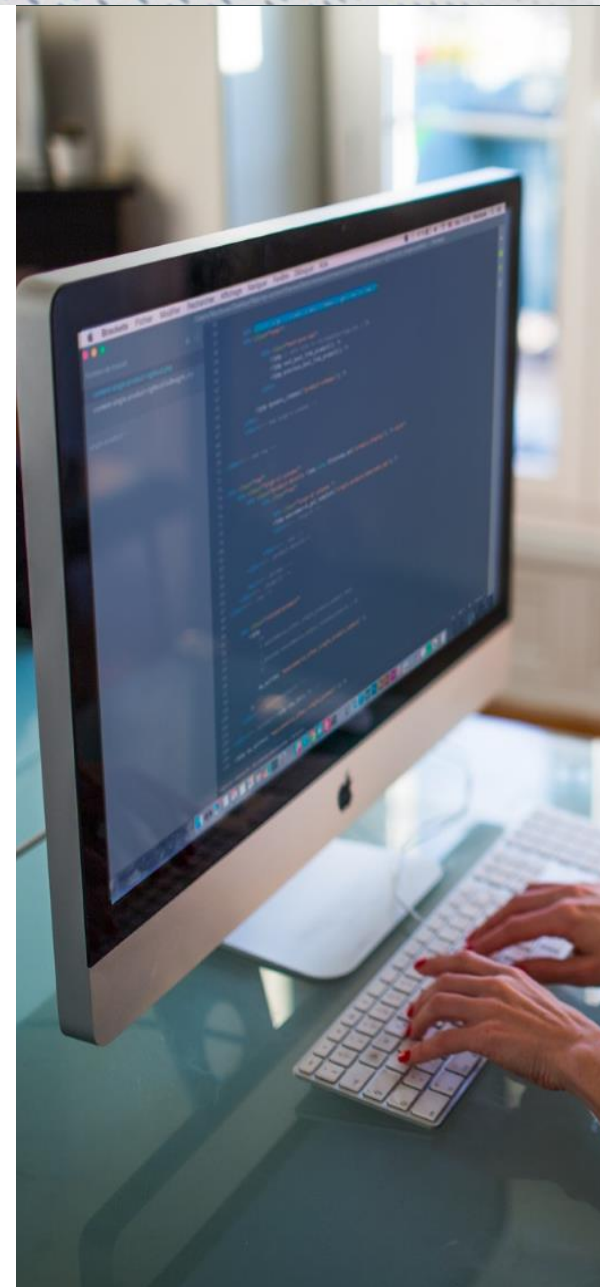
2020年社交平台无效流量为**44.1%**，国双监测的社交平台已涵盖了当下绝大部分社交媒体。

2020年社交平台无效流量占比



数据来源：国双AD系统社交数据

指标说明：社交平台无效流量指机器人刷量行为生产的内容和账号，以及人为密集参与商业贴评论的行为，也就是水军刷量行为。

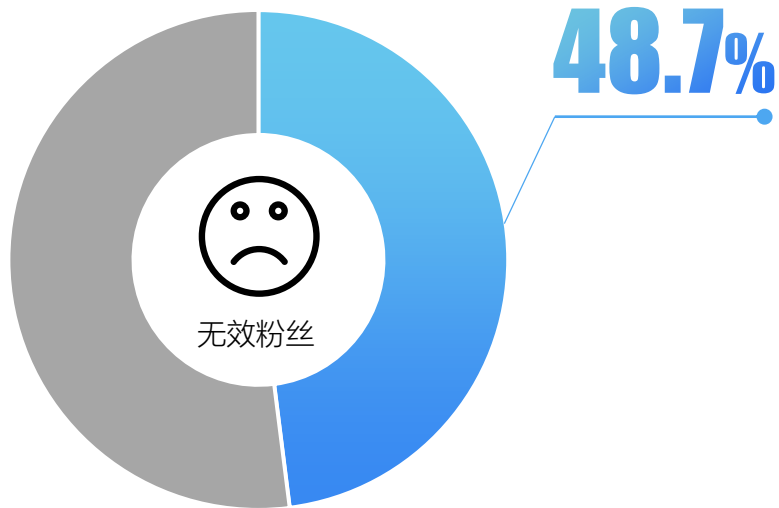


04 国双软广行业洞察与监测

2020年内容营销监测-KOL无效粉丝流量整体分析

2020年社交平台KOL无效粉丝数高达**48.7%**，国双认为无效粉丝将对品牌植入策略带来极大影响。

2020年社交平台无效粉丝占比



数据来源：国双AD系统社交数据

指标说明：社交平台KOL无效粉丝指通过国双算法识别出的异常行为关注者和评论者

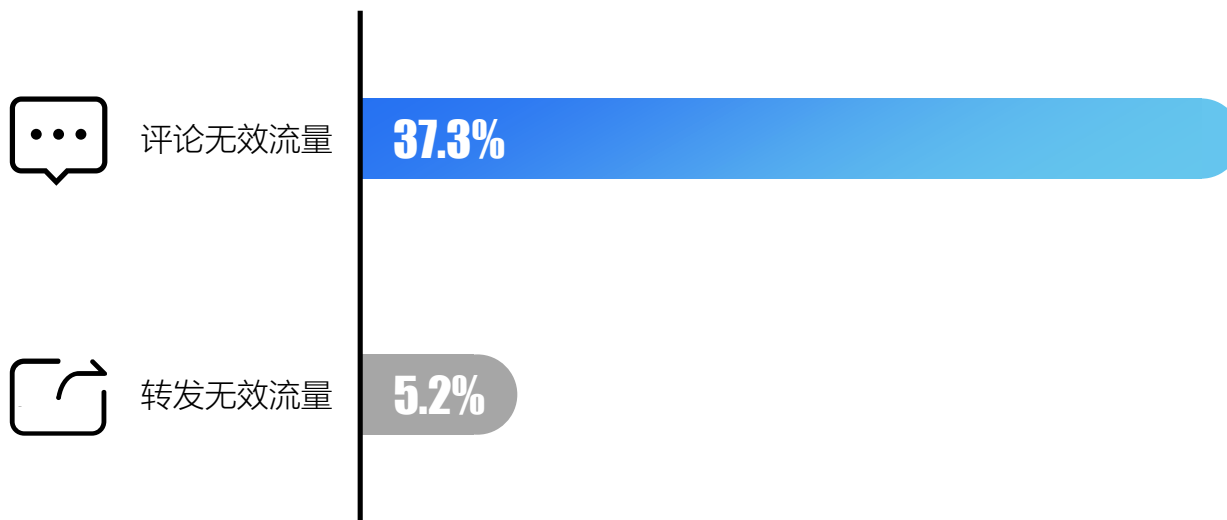


04 国双软广行业洞察与监测

2020年内容营销监测-评论无效流量与转发无效流量分析

2020年水军评论占比明显比水军转发的占比要高，高达**37.3%**，拉升了全年内容营销网络水军的流量占比。

2020年评论无效流量与转发无效流量占比



数据来源：国双AD系统社交数据

指标说明：评论无效流量：无效评论数/曝光量；转发无效流量：无效转发数/曝光量

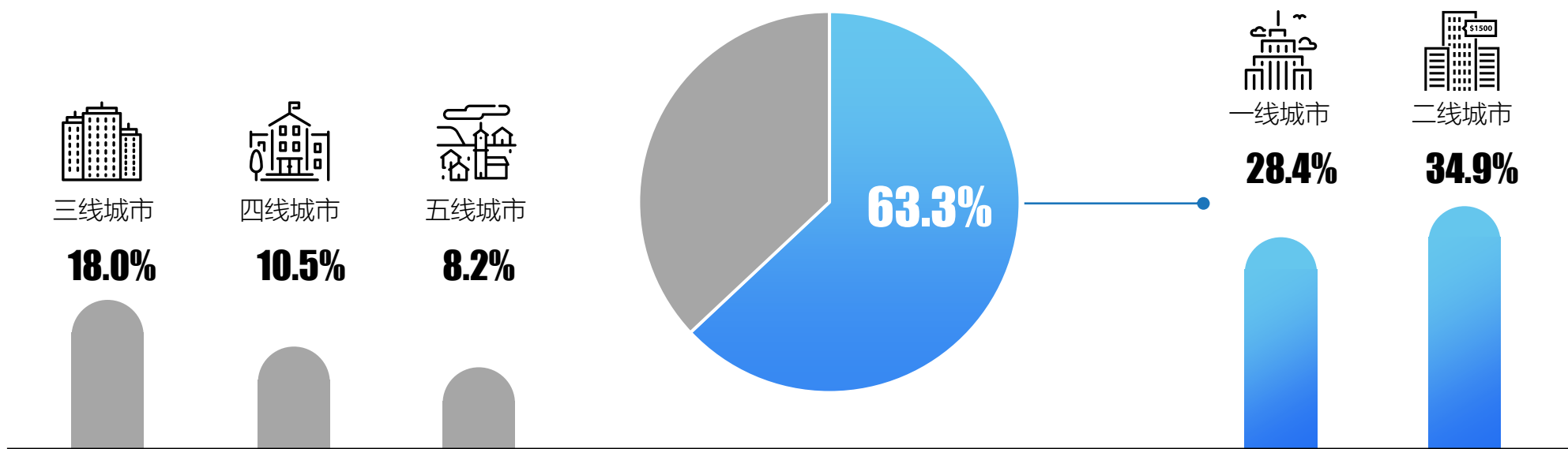
曝光量：曝光量的计算指多个维度的指标组成向量的总和。包括阅读量、转发量、评论量、点赞量、收藏数五个维度。

04 国双软广行业洞察与监测

2020年内容营销监测-无效流量地域来源分析

2020年无效流量地域来源分析发现，网络水军主要来自于一、二线城市，占比超过总体的一半，高达**63.3%**

2020年软广监测分地域来源分析



数据来源：国双AD系统社交数据

指标说明：无效流量占比：该城市级别的无效流量/无效流量

05 国双监测方案

多年以来，国双深耕广告监测领域，解决了客户各种广告监测与验证问题，积累了丰富的实践经验。在国双服务客户的过程中逐渐发现，客户越来越注重广告的效果与质量，对第三方监测公司的异常流量识别与过滤提出了更高的要求。

为了能够更好地服务客户，国双进一步升级了已有的异常流量甄别技术，参考国内外行业标准，研究出了更多的异常流量识别方法，帮助广告主守护好每一次广告的展示。

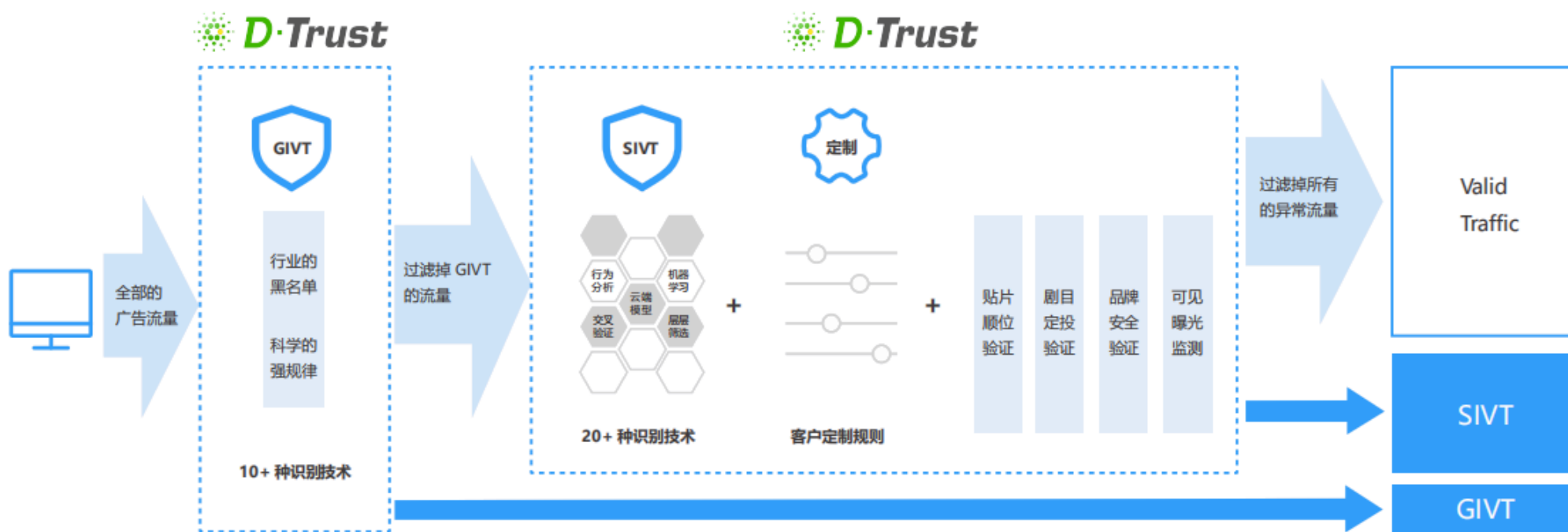
作弊手段层出不穷，不断迭代，国双监测方案也并非一成不变，而是会根据市场的变化，不断进行自我升级，通过更加先进的识别技术，在保证识别准确的前提下，更加及时地甄别出潜在的异常流量。



05 国双监测方案

国双流量反欺诈数据验证流程

1. 通过将检测到的IP、Device ID与权威的行业GIVT List中的IP、Device ID 进行比对，结合经过验证的合理规则，来识别并过滤GIVT
2. 通过设备与广告相关的各种参数、传感器信息和设备信息等，并通过数据模型、机器学习等方式识别SIVT，让复杂作弊流量无处遁形
3. 支持客户定制个性化异常规则，由国双通过技术手段实现，来将更加细分场景的异常流量及时甄别
4. 对于广告品牌安全和可见曝光等问题，国双也有针对性的验证方案



多重组合识别数据异常



90+ 维度异常判定技术



基于前后端流量打通，分析每一次访问后的像素热力图、点击热力图，从中发现异常流量



监播实录技术有效减少视频类广告滥竽充数现象



实时的数据更新和一键式报告能够迅速发现并排查流量问题

05

国双监测方案

热力图识别异常流量

广告来源	访问量	综合浏览量	平均停留时间	跳出率	平均页面访问数	Sales leads完成数	Sales leads转化数
媒体B	774,950	1,745,938	00:02:26	56.60%	2.25	1.741	0.22%

媒体B各项基础数据及转化数据看似正常，但查看不同渠道在页面上的热力图点击分布，能发现媒体B存在异常点击分布，符合程序点击发送数据包特征。

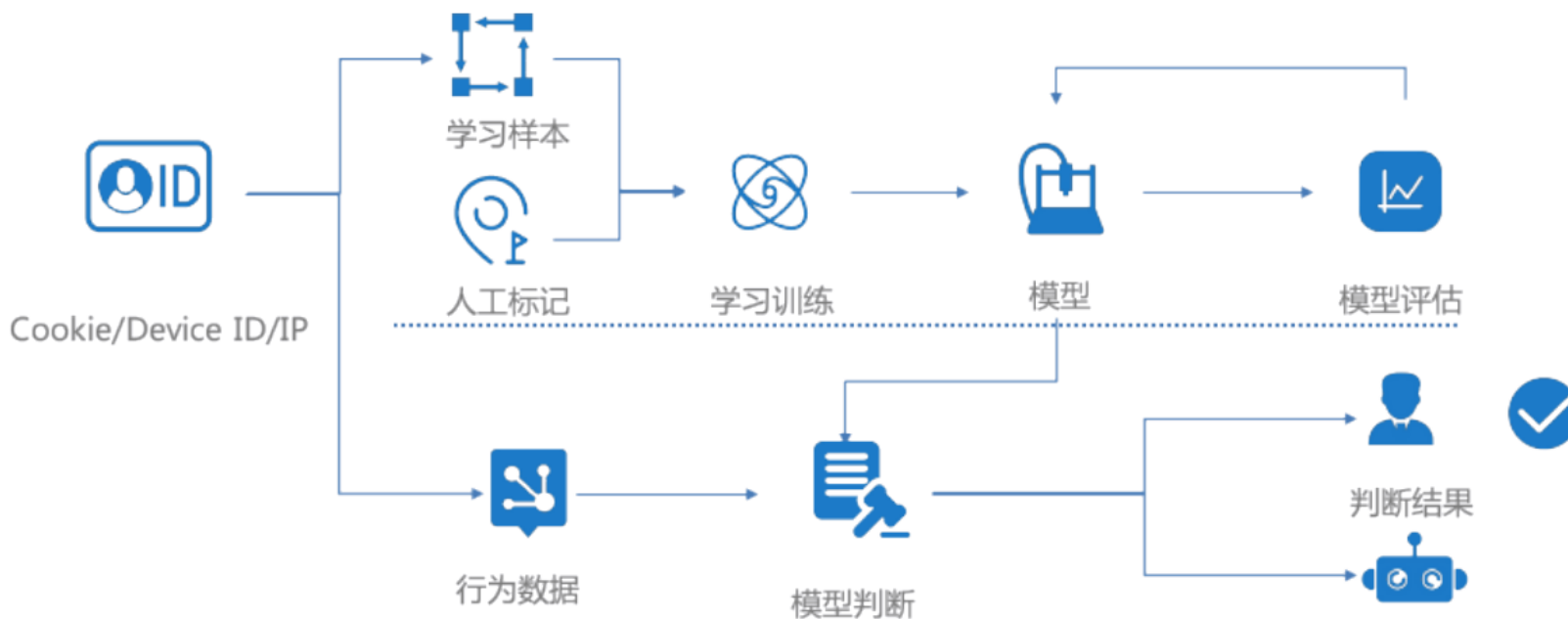


更多异常流量判断方法	异常分析排查部分原则子
<ul style="list-style-type: none"> • 连续曝光 • Stable 人群占比 • 时差异常 • 周期性波动异常 • 定向人群异常 • 稳定率异常 • 频次分布异常 • 曝光不可见 • 	<p>前端</p> <ul style="list-style-type: none"> ➤ 曝光或点击的频次 (ID/IP) 超过设定的阈值 ➤ 在较长的时间段内, 曝光或者点击数据呈现出剧烈的、周期性的上下波动 ➤ 曝光时间分布不符合网民活跃规律 ➤ 地域分布非定向, 但与行业均值、历史分布相差非常大, 表现为某一城市数据非常突出。 ➤ 同一广告曝光 / 点击来源于不同媒体, 或者是相同媒体的不同页面上, 一般是媒体在用不同页面的流量增加数据量数据非常突出。
	<p>后端</p> <ul style="list-style-type: none"> ✓ 热力图 (点击热力图、像素热力图) ✓ 停留时间分布、跳出率、二跳率、硬件厂商、机型、操作系统等维度指标 ✓ 行为路径相同、销售线索重复、用户行为特征 ✓

05 国双监测方案

基于AI自动发现复杂作弊特征

基于固定规则基本能够帮助广告主过滤掉大部分的异常流量，剩下的部分是由于作弊技术不断升级，导致固定规则无法识别。对于这些异常流量，国双采取基于机器学习的方式进行反制；通过AI技术自动发现作弊规律，不断学习并改进模型，帮助广告主识别复杂异常流量。



06 关于国双

国双是中国领先的企业级大数据和人工智能解决方案提供商。基于国双大数据平台独有的分布式数据架构和先进的实时、多维度关联性分析技术，国双的解决方案能够使客户充分洞察数据间的复杂关系，获得全新的商业洞察，以利于企业和政府客户作出更好的业务决策。“国双（Gridsum）”其名，是分布式计算（Grid）与分析（Sum）的结合。作为数字智能化的先行者，国双致力于帮助企业 and 政府客户以新颖有效的方式使用数据，提高生产力。

更多详情，请访问官网：www.gridsum.com

欢迎关注国双官方微博（@国双Gridsum）、微信公众号（ID：gridsumtech）



国双官微



国双营销研究院

GRIDSUM 国双

www.gridsum.com



Scan to know more about
Gridsum

Gridsum is the legal owner of Gridsum trademark and its products and services (including but not limited to products and services mentioned in this PowerPoint), and the holder of the intellectual property right. Gridsum does not delegate to others enjoy alone or enjoy jointly with Gridsum all the above-mentioned ownership and/or intellectual property rights. The unauthorized use or infringement of Gridsum trademark and (or) legal rights of Gridsum products and services is prohibited.