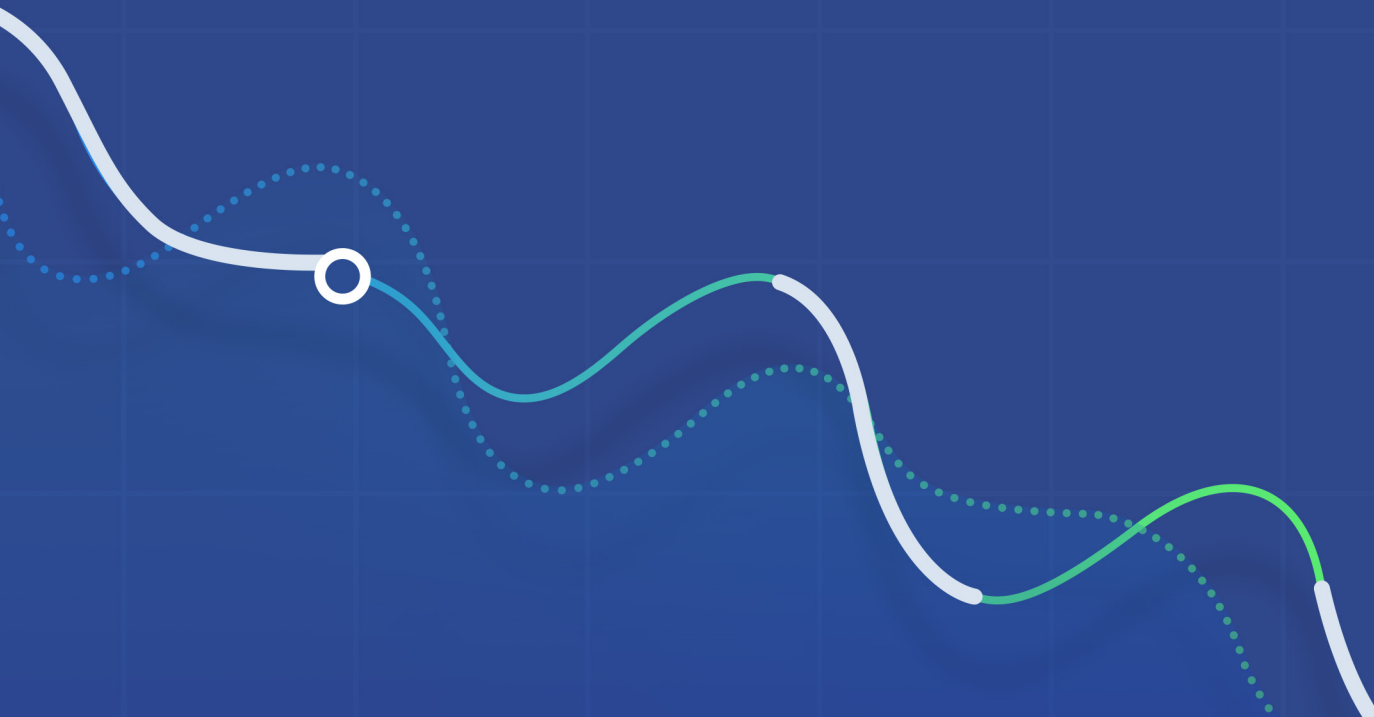


国内互联网广告无效 流量白皮书

GRIDSUM 国双

2017



CONTENT



背景介绍



2017年无效流量分析



国双的监测方案



总结

PART ONE



背景介绍



2017年无效流量分析



国双的监测方案



总结

互联网广告作为重要的营销手段，近年来一直在高速发展，相对于传统的广告形式，互联网广告由于形式多样、互动性强、易于评估等特点，更容易得到市场营销人士的青睐。据知名市场研究公司eMarketer的趋势预测，国内数字营销广告费用将从2017年的542亿美元上升到2018年的693亿美元，平均增速达到了30%。

虽然市场容量巨大，发展迅猛，但是互联网广告也面临着自己的问题。

- 2016年，通过制造机器人流量，一群位于俄罗斯的犯罪黑客在全球范围内一天能够攫取300-500万美元的不法广告收入，这是迄今为止发现的影响最大的机器人流量造假。——White Ops
- 2017年，美国在数字营销领域的投入将超过830亿美元，其中，无效流量带来的经济损失将会达到65亿美元。——eMarketer&White Ops

从上面的新闻数据可以明显看到，无效流量带来的损失是很巨大的。究其原因，虽然互联网广告比较受欢迎，但行业参与者众多，且在缺乏透明性的情况下，各方都有可能在利益的驱动下，做出有损行业健康的行为。

显然这些不法行为产生的不良影响，对广告主而言是不可容忍的。怀着对广告真实效果的不懈追求和对品牌安全的担忧，越来越多的广告主意识到不能坐以待毙，必须与行业内权威的组织、媒体以及第三方技术公司合作，一起推动互联网广告行业的透明性和安全性。国内的情况也是如此，在权威组织——中国广告协会、MMA中国召开的多次行业会议上，宝洁、可口可乐等大广告主积极参与，并一直呼吁在行业内建立统一的监测和验证标准，不断规范国内互联网广告的投放和效果评估。

国双作为行业第三方监测和验证公司，有责任甄别这些作弊行为，找出异常的流量，还原广告的真实效果，从而挽回广告主的损失。

时值3·15国际消费者权益日之际，国双通过无效流量监测体系，分析所收集到的2017年互联网广告监测数据，发布了《国内互联网广告无效流量白皮书》。凭借雄厚的技术实力和多年深耕大数据分析的经验，国双希望在展示国内互联网广告无效流量现状的基础上，为广告主、代理公司等行业相关者提供直观的数据参考，促进互联网广告行业健康可持续发展。

PART TWO



背景介绍



2017年无效流量分析



国双的监测方案



总结

无效流量标准



本报告针对的无效流量，是参照国内权威组织中国广告协会对于无效流量的定义和分类。这套标准是在欧美成熟市场的无效流量标准基础上，结合了国内实际的互联网技术环境，对部分标准做了针对性的调整之后形成的。

根据中国广告协会的标准，无效流量被分为两类——GIVT和SIVT：

GIVT是General Invalid Traffic的缩写，即一般无效流量，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。

SIVT是Sophisticated Invalid Traffic的缩写，即复杂无效流量，这种类型的流量无法通过简单的规则识别出来，一般需要通过高级分析，多方合作与协调，乃至人工干预等方法以及广告投放活动以外更大范围的数据信号才能分析和识别。

无效流量标准

中国广告协会对于无效流量的完整分类，具体如下：

GIVT/一般无效流量	SIVT/复杂无效流量
机器人和爬虫或其它伪装成合法用户的流量数据以及非浏览器用户代头或其它形式的未知浏览器带来的流量	劫持设备以及设备中的会话
超出频次、时间间隔等目标设定的流量数据	非法劫持广告创意和操纵流量
通过隐藏 / 堆叠 / 覆盖或其它方式导致用户无机会看到正常广告内容的流量	内容盗用、伪造、虚假展示
已知的来自数据中心的流量（指明显具有非人类访问广告所在的特定网络 IP 或 IP 段所产生的流量来源）	恶意修改、插入或删除 cookie 内容以改变用户访问记录
预获取或浏览器预览的广告流量	操纵或伪造位置数据以及相关属性
已知的来自高危或者作弊来源流量	无效代理流量（即来自中间代理设备的无效流量，包括通过代理设备操纵流量计数、创建 / 传输非人类流量或无法通过协议验证的流量
基本信息缺失或不一致的流量（基本信息至少应包含事件类型、广告系列 ID、时间戳、IP、请求方式、用户代理 UA 字段）	

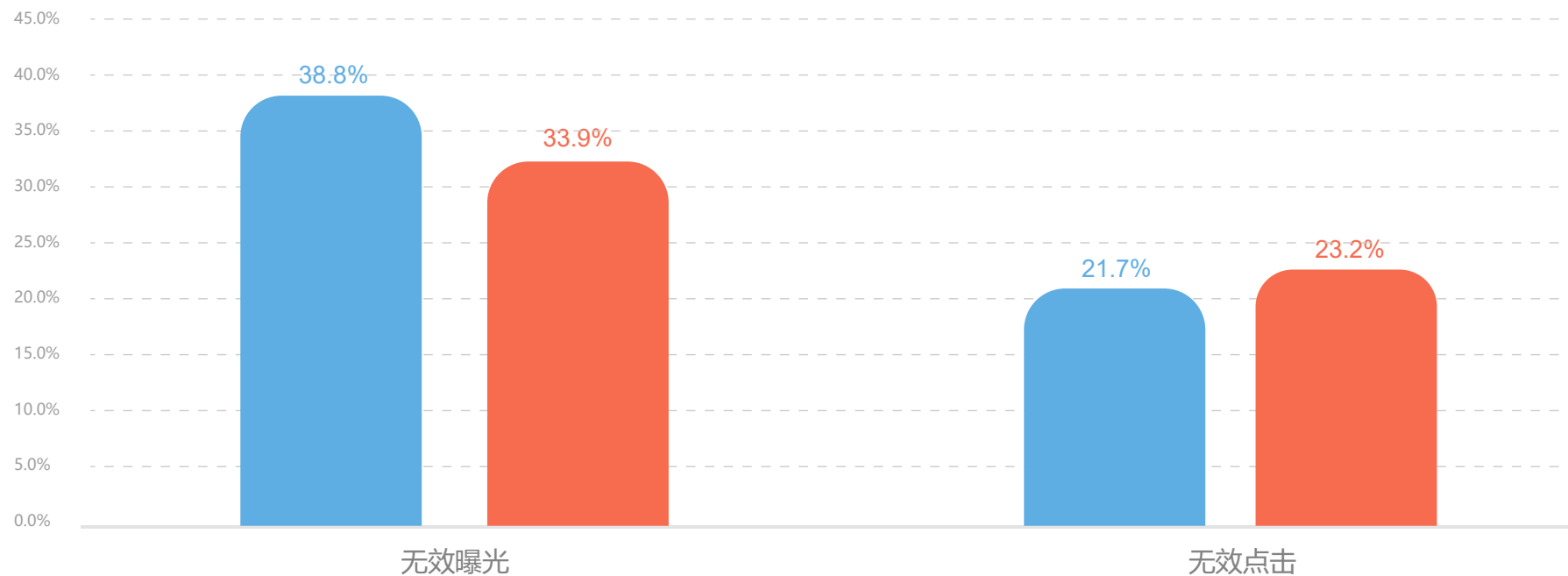
无效流量数据分析

通过对2016—2017年的数据进行全面地监测和分析后，我们发现：

2017年全年无效曝光占比达到了33.9%，较2016年有小幅下降；而2017年的无效点击比例却较2016年提升了1.5个百分点。

2016-2017年无效流量对比分析

◆ 2016年 ◆ 2017年



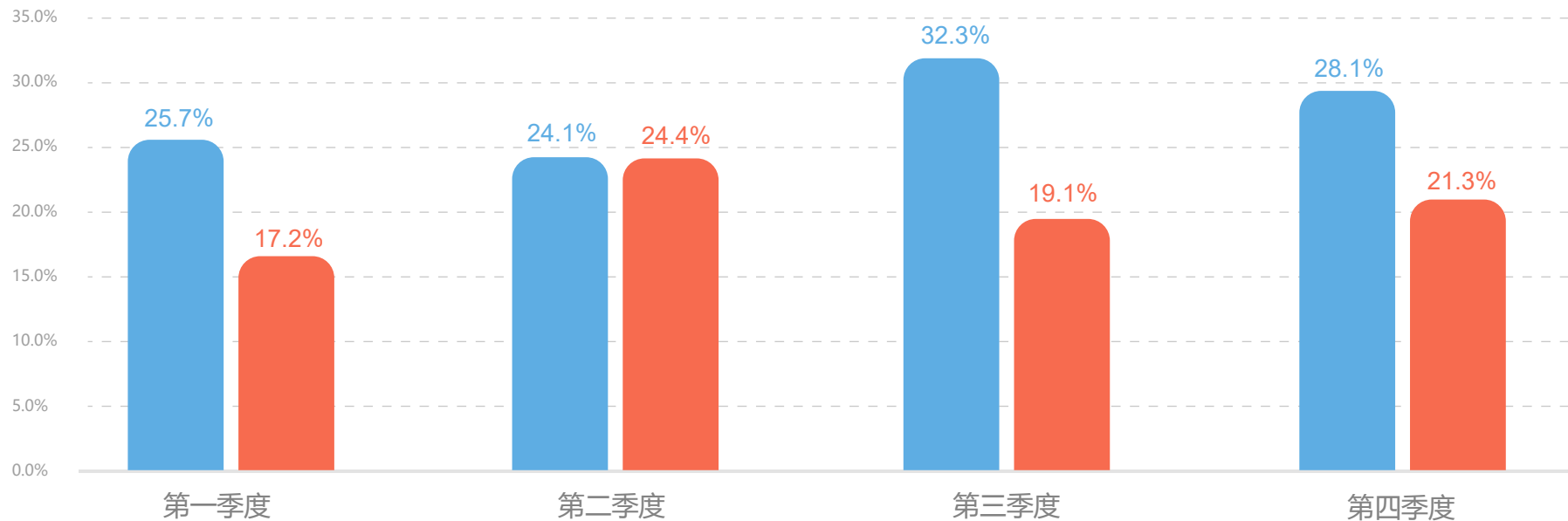
数据来源：2016年1月1日至2017年12月31日之间，国双监测的广告数据

2017年分季度趋势分析：

我们发现无效曝光在第三季度出现了一个较为明显的上升，拉升了全年的无效曝光比例。而点击方面，全年高峰出现在第二季度，达到了24.4%。

2017年分季度无效曝光/点击趋势分析

◆ 无效曝光 ◆ 无效点击

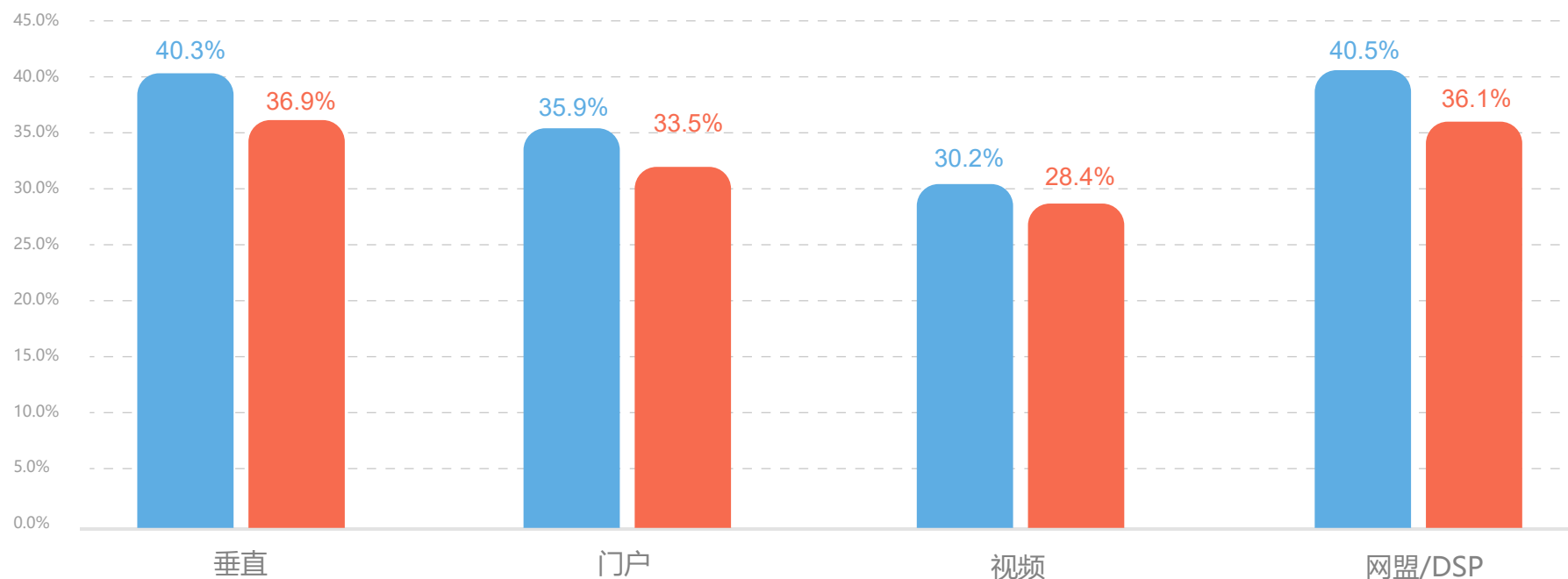


分媒体类型——无效曝光：

我们发现垂直类和网盟/DSP类的媒体，无效曝光占比较高，均接近40%。
另外，全部四类主要媒体类型无效曝光都呈现出下降的趋势。

2016-2017年分媒体类型无效曝光对比

◆ 2016年 ◆ 2017年



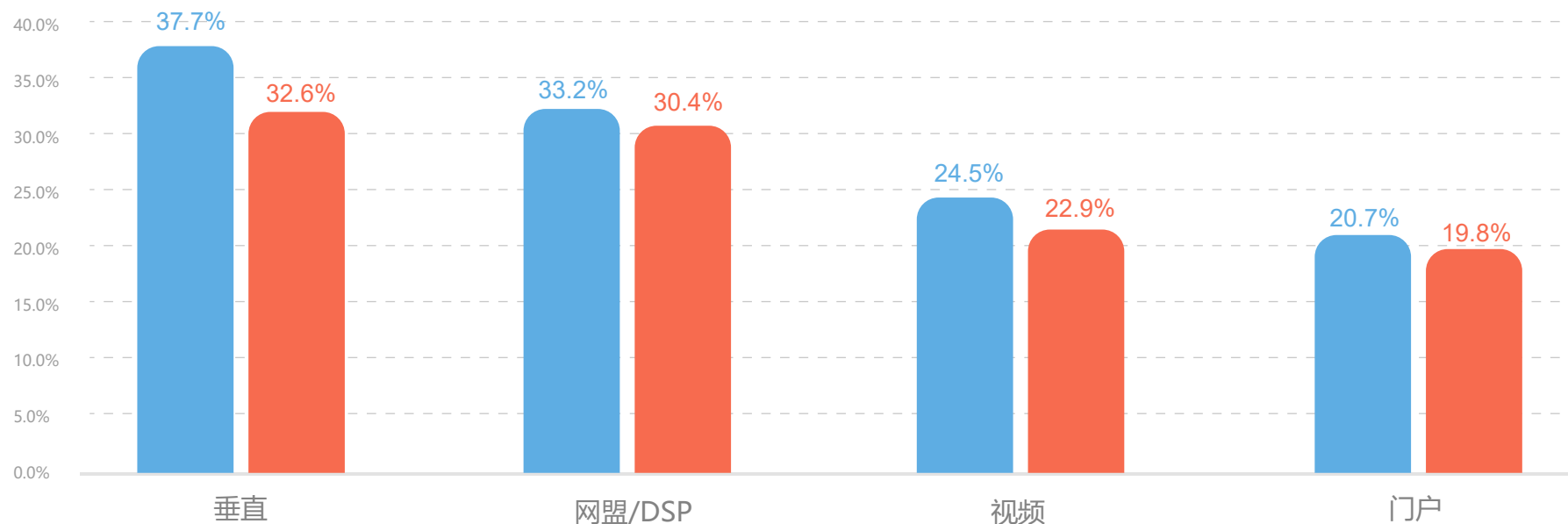
分媒体类型——无效点击：

相较于2016年，2017年主要的4个行业类型在无效点击方面都有小幅下降；其中，垂直行业的下降幅度甚至超过了5个百分点。

2016-2017年分媒体类型无效点击对比

◆ 2016年

◆ 2017年



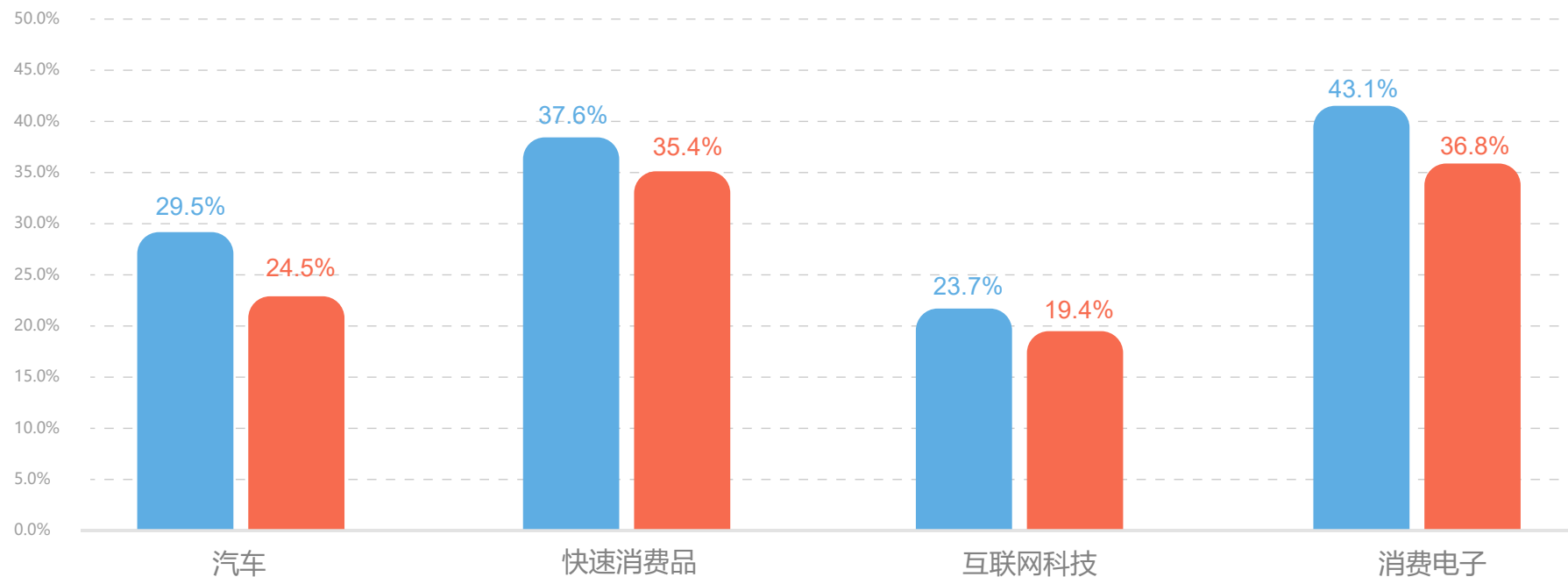
细分行业——无效曝光：

通过对多个行业进行对比后，我们发现，不同行业的无效曝光差异比较明显，消费电子行业的占比去年曾超过40%；而汽车和互联网科技行业的无效曝光占比均维持在30%以内。

2016-2017年分行业无效曝光对比

◆ 2016年

◆ 2017年



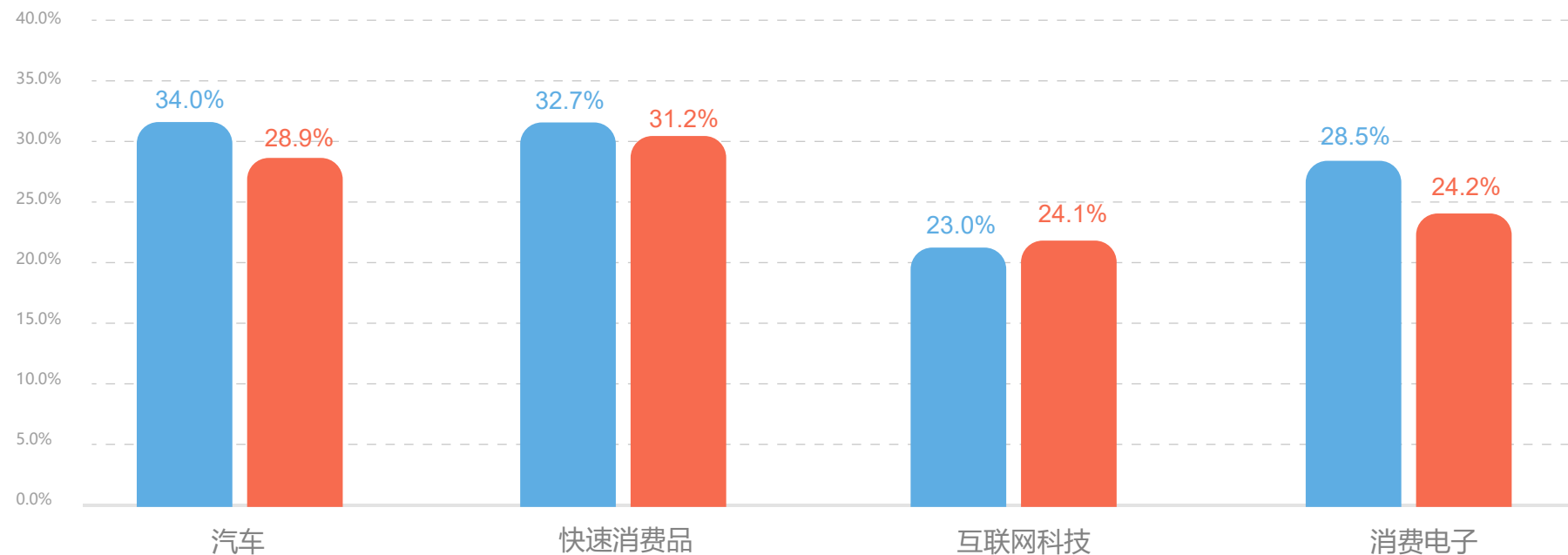
细分行业——无效点击：

2016-2017年无效点击情况整体呈下降趋势，但是互联网科技行业却反其道而行，无效点击情况较上年度增加了1.1个百分点。

2016-2017年分行业无效点击对比

◆ 2016年

◆ 2017年

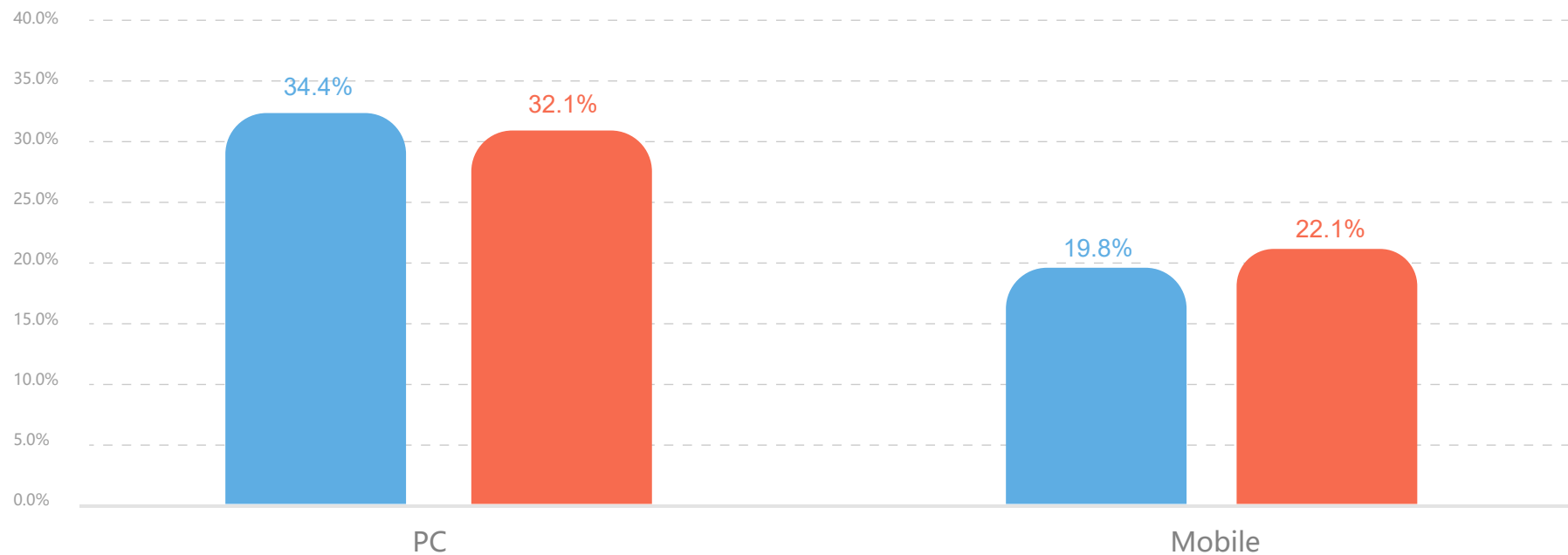


分设备——无效曝光：

随着流量在不断向移动端迁移，移动端广告的问题也日益凸显，通过对2016-2017年的分设备流量进行分析后，我们发现，虽然目前移动端的无效曝光占比较低，但是却呈现出了上升的趋势。

2016-2017年分设备无效曝光对比

◆ 2016年 ◆ 2017年

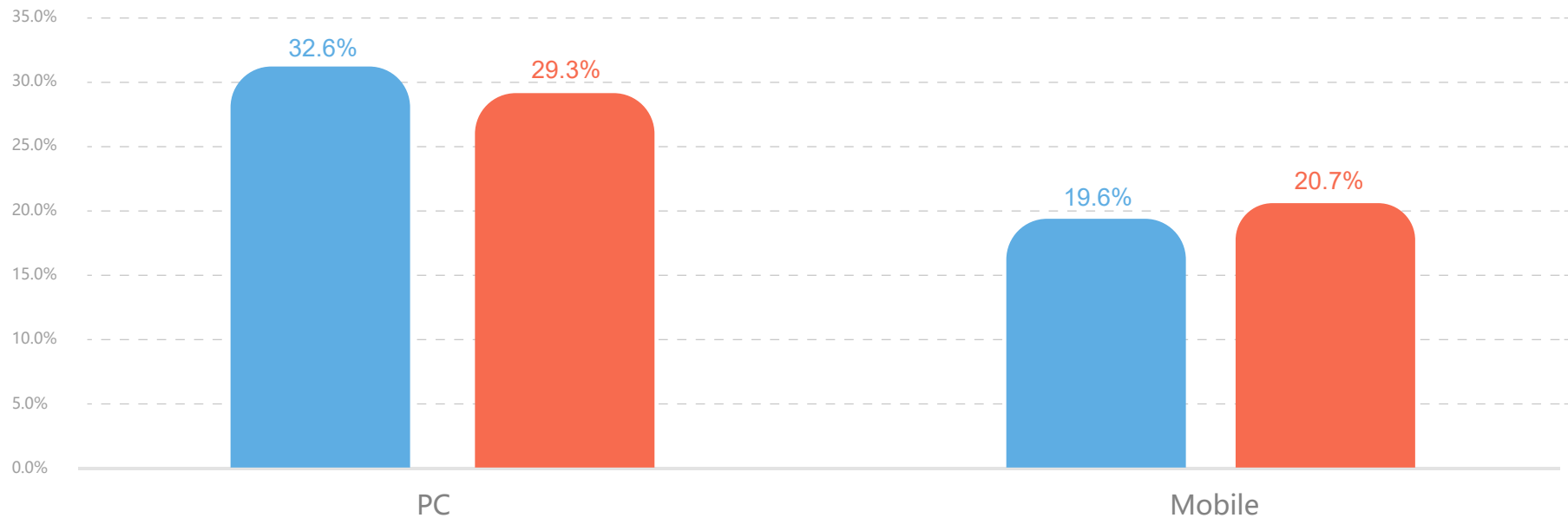


分设备——无效点击：

与无效曝光类似，无效点击也呈现出了类似的趋势，PC端占比虽高，但较上一年有小幅下降。而移动端却有了微幅增长。

2016-2017年分设备无效点击对比

◆ 2016年 ◆ 2017年



PART THREE



背景介绍

2017年无效流量分析

国双的监测方案

总结

如何取得信任？



为了能够更科学、更准确地评估互联网广告的投放效果，无效流量的识别和剔除是必不可少的，而这往往也是广告效果评估的第一步。

作为中国领先的企业级大数据和人工智能解决方案提供商，国双提供的技术方案和数据结果必须具有足够高的可信度，才能被广告主和媒体所认可。因此，国双很早就开始研究并参考欧美成熟市场的权威组织（主要有MRC、IAB、MMA和TAG）对无效流量的定义和分类，来指导自己的无效流量识别；同时，考虑到国内市场环境的特殊性，国双也积极加入了中国广告协会和MMA中国，并以核心会员的身份，推动相关无效流量标准在国内市场的落地。

国双的监测方案



在国内权威的行业标准指导下，国双逐步研究并开发出了自己的一套无效流量监测体系，不仅包含前端广告数据的检验（如：黑名单、频次、定向投放异常等），还结合了广告后端数据的异常监测（如：热力图、多个维度构建的用户行为特征等），前后端广告数据打通，串联起广告流量的完整路径，有效识别异常情况。最后，我们还利用AI的机器学习，自动发现复杂的作弊特征，来甄别一般手段无法发现的无效流量。

作弊的方式在不断增多，作弊技术手段一直在升级，而国双也在不断强化自己的监测体系，以达到对新型的异常场景，进行及时有效地识别和剔除；同时国双也在积极寻求外部合作，利用外部的庞大数据资源，更好地为客户服务。

国双的无效流量监测体系架构图：

前端数据

【曝光点击分析】

【来源分析】

【地域分析】

【频次分析】

【.....】

+

后端数据

【转化分析】

【热力图分析】

【用户行为分析】

【多维度指标】

【.....】

+

【AI 自动建模分析】

国双特色监测能力——热力图

作为国双的特色监测能力，网站热力图在检验实际项目后端异常方面发挥了重要的作用。

原理：通过热力图工具绘制出的鼠标点击分布，并与正常的分布做对比，即可发现常规手段无法发现的异常。

案例分析：下方左图的点击区域与正常的点击情况，有着明显的区别，甚至一些没有任何按钮或链接的区域，却产生了大量的点击。

可疑流量



正常流量



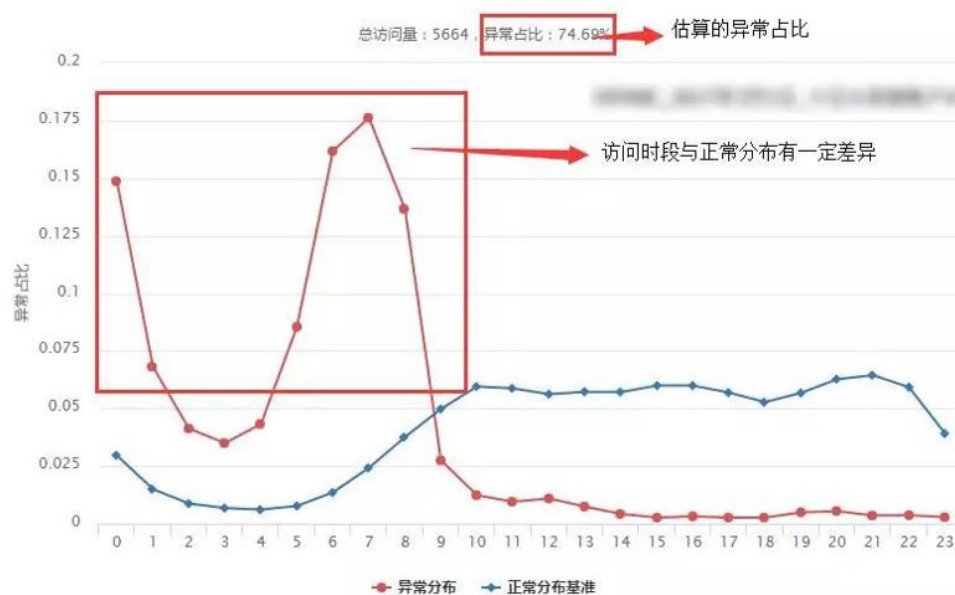
(热力图颜色的深浅代表着点击的聚集程度，颜色越深，点击量越大，点击也越聚集)

国双特色监测能力——多维度自动监测

国双的另一项特色武器是机器学习算法，它可以从不同维度（如24小时分布、城市、IP分布、页面停留时间、屏幕分辨率、操作系统版本、设备品牌、浏览器、设备名称等）实现异常流量的自动甄别，不放过每个流量。



上图流量出现的很多如六安市、抚州市等城市，与正常基准不一致，所以这部分流量就很可疑了。



上图流量中很大比例发生在凌晨，不符合正常网民的作息规律，在没有特意加大此时段投放量的前提下，该渠道流量就可算作异常了。

PART FOUR



背景介绍

2017年无效流量分析

国双的监测方案

总结

总结



与互联网广告中无效流量的斗争是个长期且艰巨的工程，无法单靠某一股力量实现，所以需要全行业各个参与者联合起来，通过大家都认可的规则 and 标准，不断减少无效流量的占比。

国双一直积极与行业内权威的组织和企业合作，一起呼吁更多的参与者加入抵制无效流量的队伍中来，不断净化互联网广告环境，推动广告数据的透明性，最终构建健康的数字营销生态。

法律声明



【版权声明】

本报告为国双制作，报告中文字、图片、表格均受有关商标和著作权的法律保护，部分文字和数据采集于公开信息，所有权为原著者所有。使用该报告相关内容时请注明出处。

【免责条款】

本报告的各项内容和数据仅用于研究和参考，任何第三方对于本报告各项内容和数据的使用或者引用所导致的任何结果，本公司以及国双数据中心不承担任何的法律责任，请任何第三方在接受该免责声明的前提下，在法律允许的条件下，合理使用本报告。