

# 国内互联网广告异常流量 白皮书

2019

GRIDSUM 国双



# Content

---



背景介绍



2019年异常流量分析



国双监测方案

# 01 背景介绍



## 背景介绍

据 eMarketer 数据显示, 2019 年, 全球数字广告支出增长 17.6%, 达到 3332.5 亿美元。这意味着, 数字广告将首次占据全球广告市场的一半左右。2019 年中国全部的媒体广告支出上涨 14.6%, 达到 1147.7 亿美元。数字广告一路高歌, 但伴随着迅猛增长的, 还有一直以来被诟病的数字广告异常流量问题。

2018 年末, Integral Ad Science Threat Lab (IAS) 监测到一个被称为“404bot”的机器人程序, 它可绕过 ads.txt 保护进行域名欺骗, 制造异常流量, 使广告主蒙受巨大损失。仅 2019 年 4 月到 9 月, 就有超过 15 亿个广告受到影响, 损失 1500 万美元甚至更多。

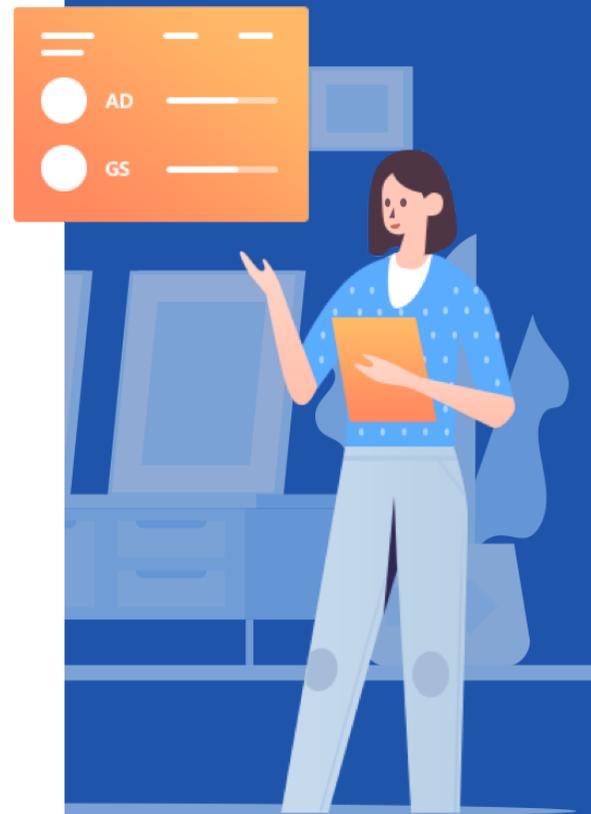
2019 年 2 月, 甲骨文揭露了一个名为“DrainerBot”新广告欺诈行动, 它会在使用者不知情的状况下播放视频广告, 数百款 Android 应用受到影响, 这些应用下载量超 1000 万次。这些视频广告不会在屏幕上出现, 也不会被用户看到, 白白浪费广告主营销预算。

以上只是在 2019 年全球发生的因异常流量造成的 420 亿美元损失中的冰山一角。

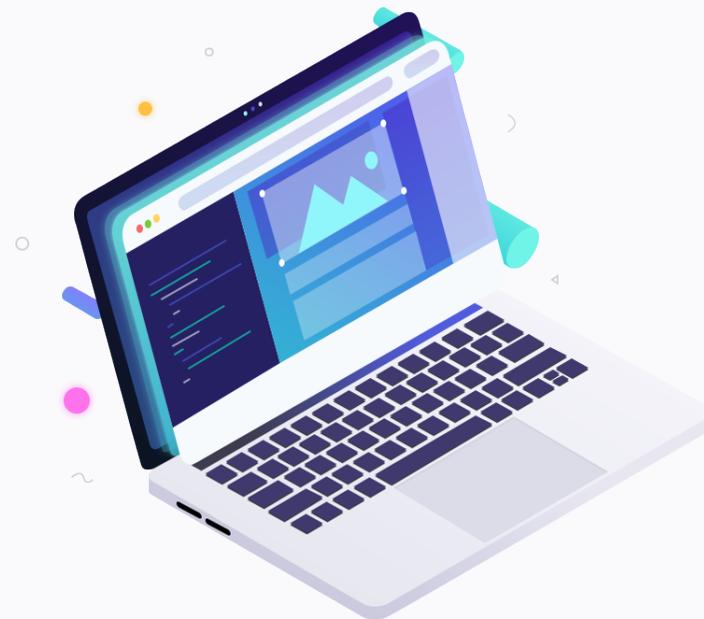
与全球相比，国内形势同样不容乐观。当前，行业内各方都在积极对抗异常流量，为获得真实数据而努力。作为独立的第三方监测公司、国内广告反欺诈的先行者，国双有责任与义务为广告主和整个行业监测并识别出作弊行为和异常流量。同时，作为中国广告协会理事单位、中国广告协会互联网广告委员会常务委员单位，国双从 2017 年起就与行业协会、同行者一起推动异常流量监测的标准化和规范化，促进行业的健康发展。

此次，在中国广告协会的指导下，国双连续第三年制作并发布了《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总结并完善的监测方案。

指导单位： 中国广告协会  
CHINA ADVERTISING ASSOCIATION



# 02 2019年异常流量分析



## 异常流量标准

本报告针对的异常流量，是参照了 MRC、IAB 的国际行业标准，国内权威组织中国广告协会组织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。根据标准异常流量被分为两类：GIVT 和 SIVT。

GIVT 是 General Invalid Traffic 的缩写，即常规无效流量，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。

## 异常流量数据分析

### 2019 年全年整体分析：

2019 年互联网广告异常曝光占比 32.2%，较 2018 年增长 4.2%。异常点击占比有明显增加，占比达 33.8%，并首次超过异常曝光占比。这源于广告主愈加重视并加大对于效果类广告的投放，以及国双根据市场趋势，针对新的异常形式加强了异常点击的识别。

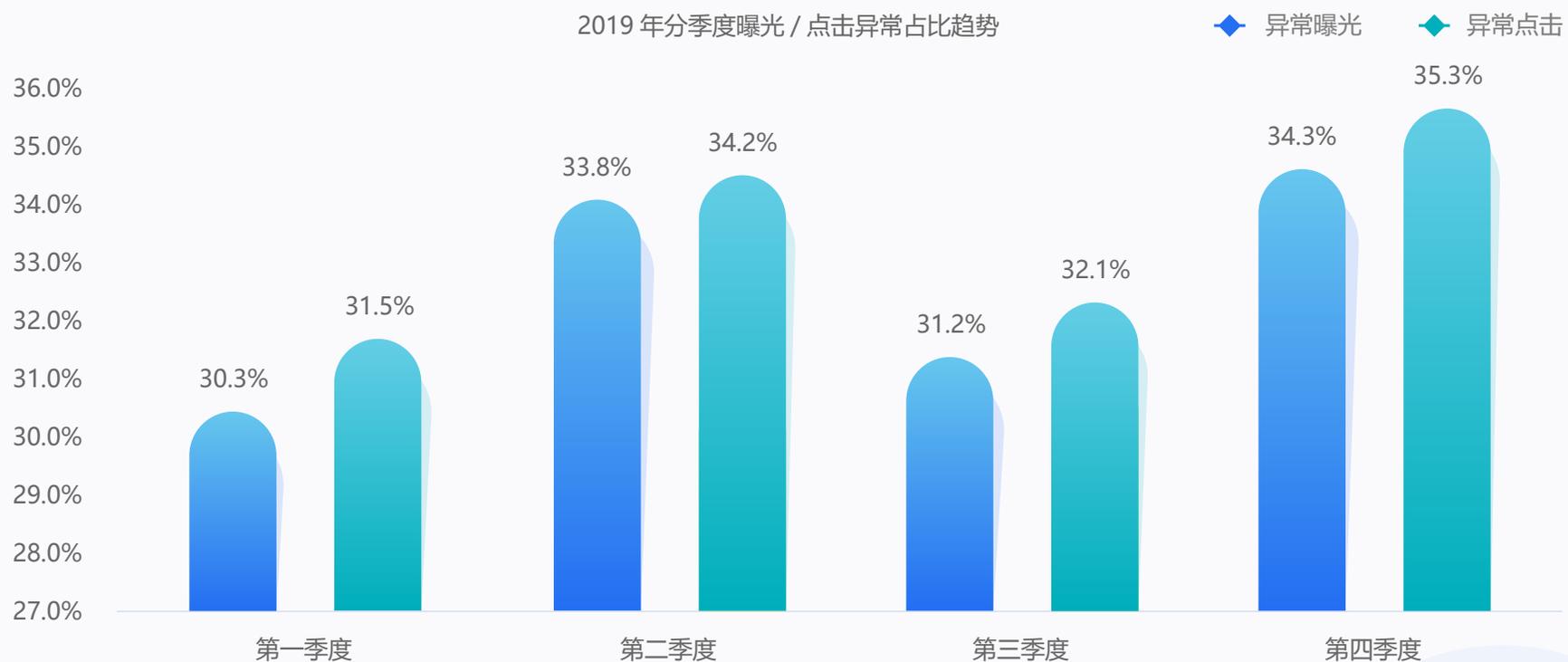


数据来源：Gridsum Ad Dissector 数据时间：2019 年 1 月 1 日 -2019 年 12 月 31 日

## 异常流量数据分析

### 2019 年分季度趋势分析：

异常曝光与异常点击占比均在第二、第四季度相对较高，第四季度高于第二季度，这可能与二、四季度有 618，双十一活动有关。



数据来源：Gridsum Ad Dissector 数据时间：2019 年 1 月 1 日 -2019 年 12 月 31 日

## 异常流量数据分析

### 2019 年分月度趋势分析：

2019 年，异常曝光和异常点击占比波动总体较为平稳，下半年占比略高于上半年。



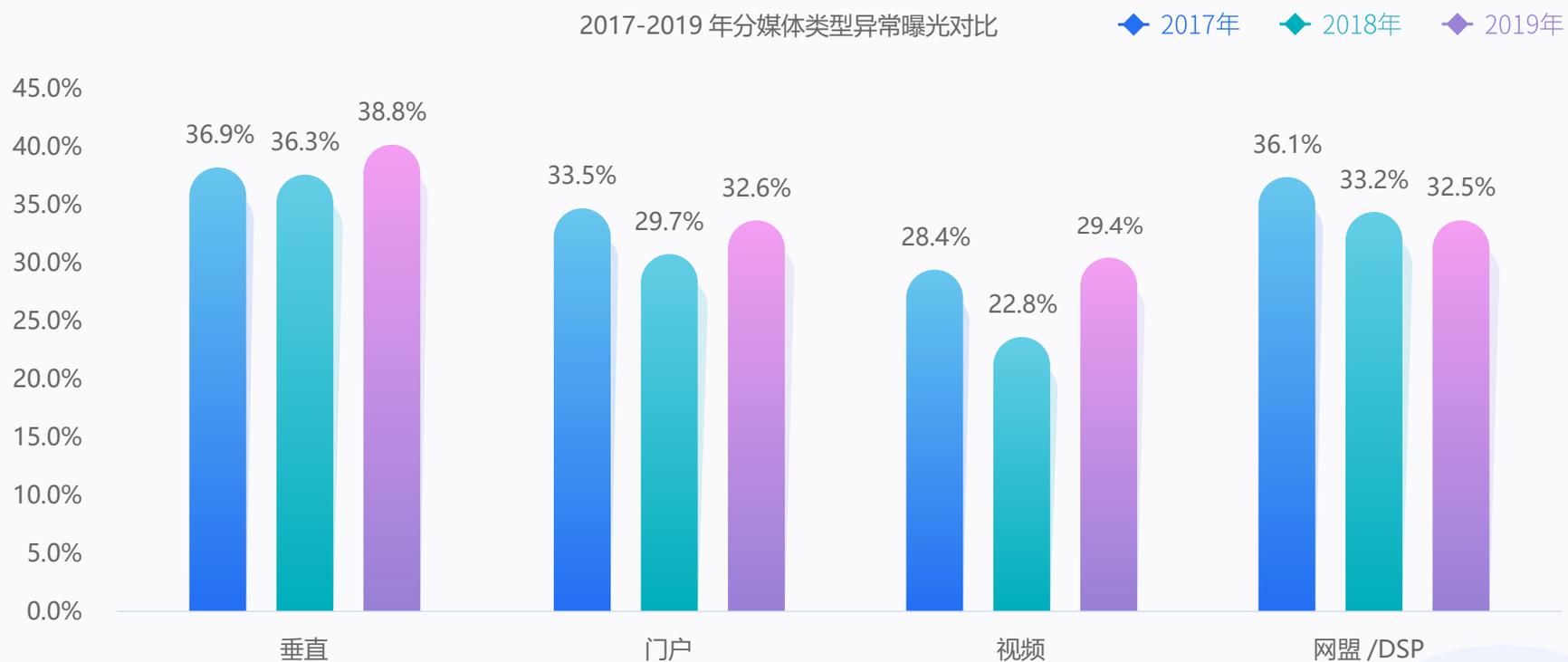
数据来源：Gridsum Ad Dissector 数据时间：2019 年 1 月 1 日 -2019 年 12 月 31 日

## 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占比，除网盟 /DSP 占比稍有降低外，其他各媒体均有增高。其中垂直媒体异常曝光占比近 4 成。

视频媒体异常曝光占比增速明显，同比 2018 年增长 28.9%。



数据来源：Gridsum Ad Dissector 数据时间：2019 年 1 月 1 日 -2019 年 12 月 31 日

## 异常流量数据分析

### 分媒体类型——异常点击：

在异常点击占比方面，各媒体也均明显增高，门户、视频最为突出。

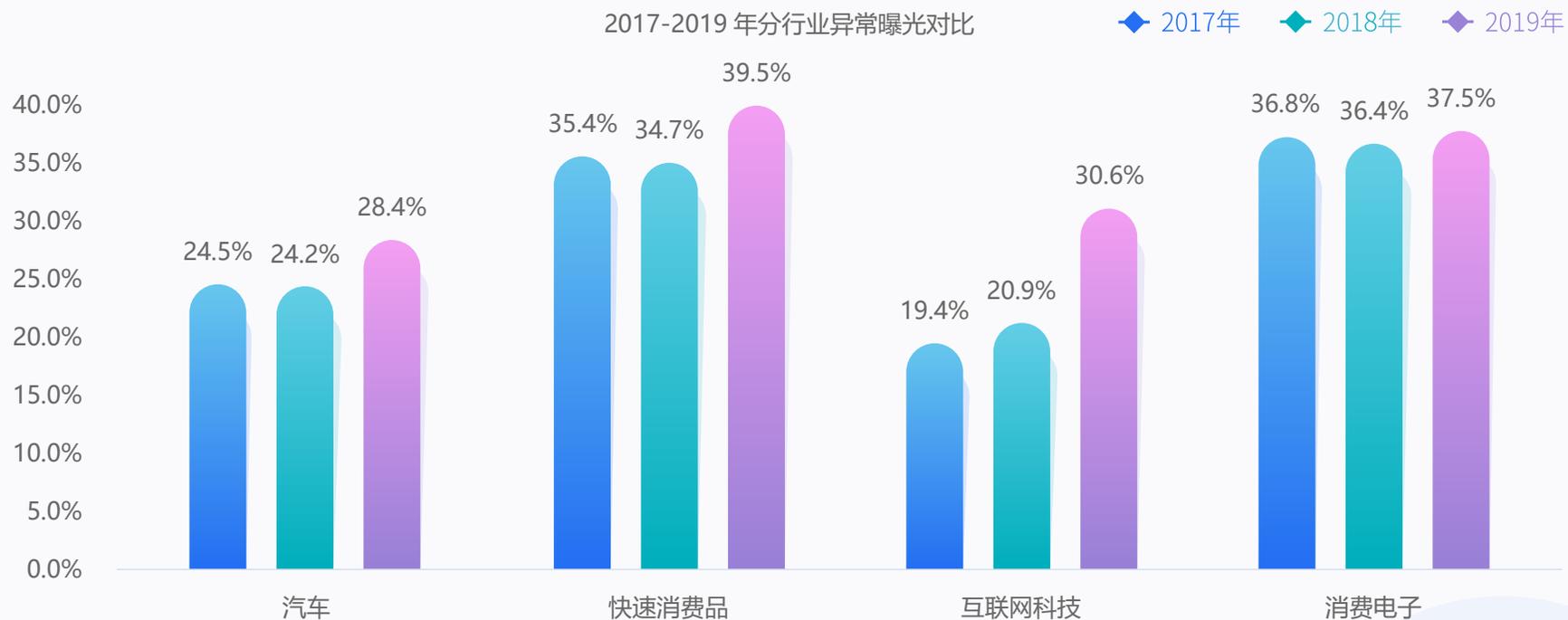


数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## 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，以汽车、快速消费品、互联网科技、消费电子行业为例，2019年各行业异常曝光占比均有所增长，其中快消行业异常曝光占比最高达39.5%。汽车行业异常曝光占比28.4%，低于行业平均水平，这得益于国双在汽车行业异常流量的突出研究。

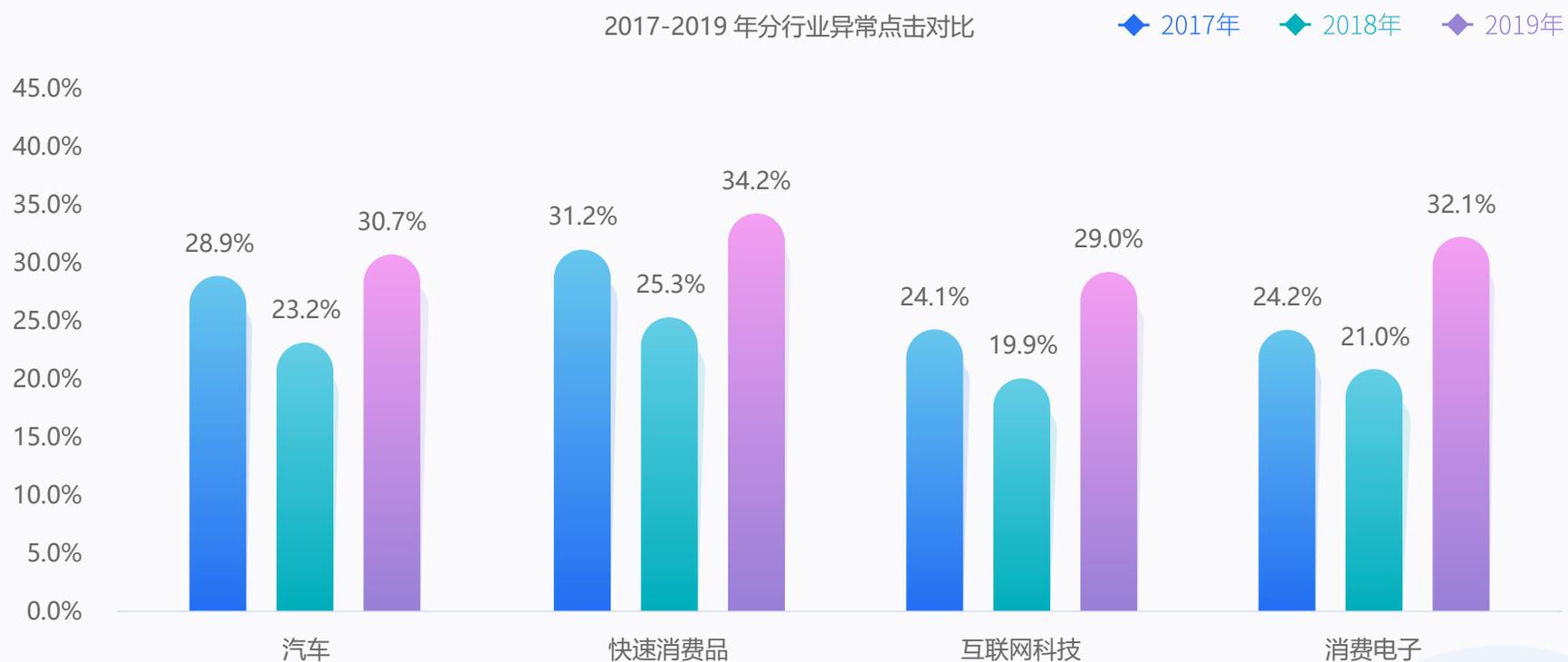


数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## 异常流量数据分析

### 细分行业——异常点击：

同时，2019年汽车、快速消费品、互联网科技、消费电子行业的异常点击占比也有明显增加。其中，快消行业的占比较高，达34.2%，对抗异常流量，任务依旧艰巨。



数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## 异常流量数据分析

### 分设备——异常曝光：

PC 端异常曝光占比持续高于移动端，不过在涨幅上移动端更为明显。2019 年 PC 端较 2018 年增长 5.7%，而移动端增长 19.7%。移动端作为目前数字营销领域各品牌主重点投入的领域，反作弊的局势也日趋严峻。



数据来源：Gridsum Ad Dissector 数据时间：2019 年 1 月 1 日 -2019 年 12 月 31 日

## 异常流量数据分析

### 分设备——异常点击：

在异常点击方面,PC 端和移动端占比均有大幅度增高,分别达到 35.0%、32.4%。其中移动端涨幅达 48.6%,需要着重关注。



数据来源: Gridsum Ad Dissector 数据时间: 2019 年 1 月 1 日 -2019 年 12 月 31 日

## OTT广告规模趋势

据国内 OTT 大数据行业权威第三方大数据公司勾正数据报告显示, OTT 用户数在 2019 年已达到 2.19 亿户, 渗透率达 38%, OTT 广告将进入高速发展期。

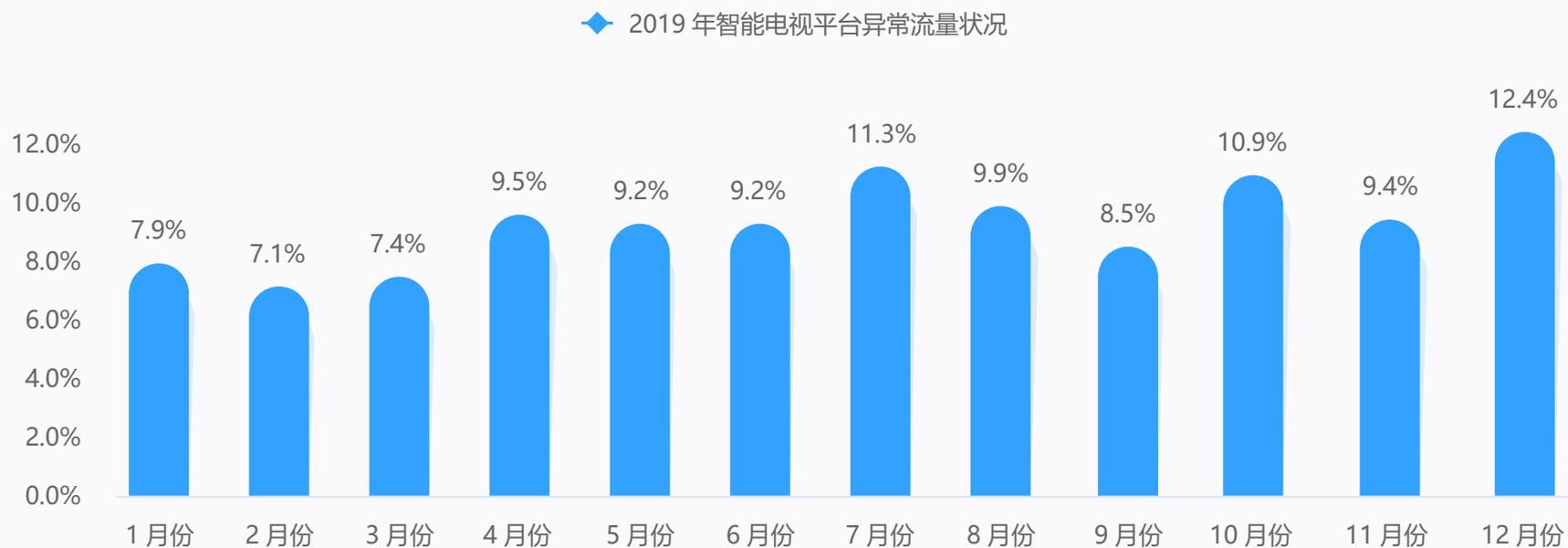


数据来源: 勾正数据推算, OTT 广告规模不包含 OTV 电视端广告 时间周期: 2016-2022 年

## OTT异常流量 (C2S)

2018 年，由国双牵头起草的《MMA 中国无线营销联盟 OTT 广告监测标准》正式发布。越来越多的厂商和媒体愿意开放支持第三方监测，愿意接受 C2S API 方式的监测。以下我们仅针对 C2S 监测模式的流量进行异常流量的分析。

2019 年 OTT 市场 C2S 流量中，异常流量占比 9.4%，较 2018 年增长 7.9%。分月占比趋势如下



数据来源: Gridsum Ad Dissector 数据时间: 2019 年 1 月 1 日 -2019 年 12 月 31 日

## OTT异常流量 (C2S)

### OTT 异常流量——按广告类型：

在 OTT 的异常流量中，最常见的广告类型为暂停广告、信息流广告、贴片广告、开关机广告，其中暂停广告的异常流量占比最高达 12.6%。

2019年OTT异常流量常见广告类型占比



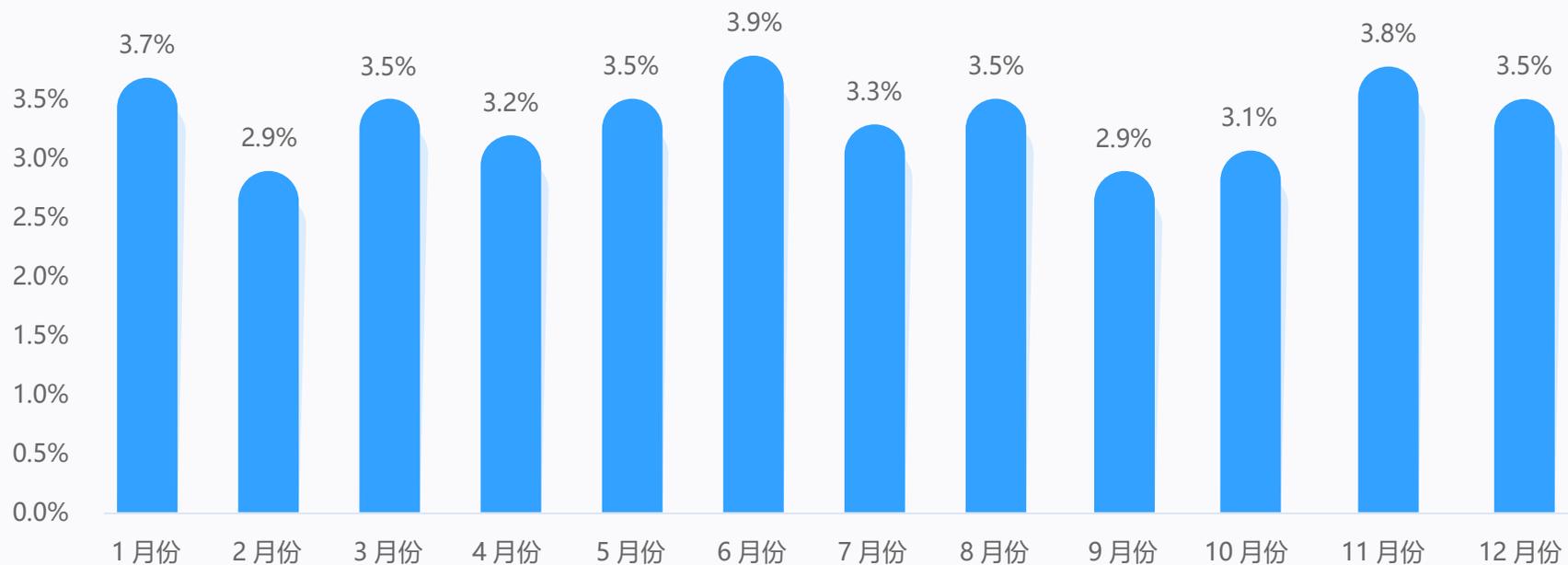
数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## GIVT（常规无效流量）数据分析

### GIVT:

2019 年国内互联网广告 GIVT（常规无效流量）平均占比为 3.4%，年中 6 月份 GIVT 占比最高，达 3.9%。

2019 年 GIVT 占比分月趋势

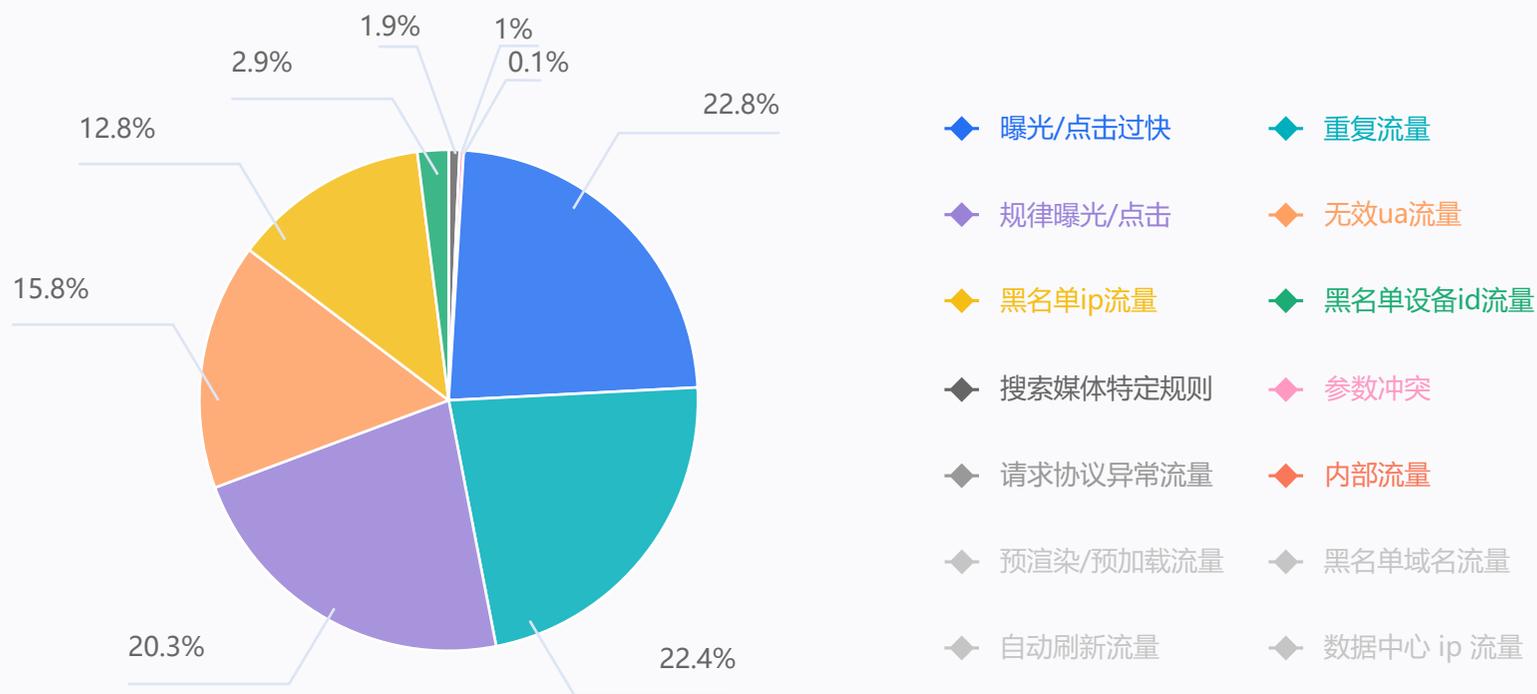


数据来源: Gridsum Ad Dissector 数据时间: 2019 年 1 月 1 日 - 2019 年 12 月 31 日

## GIVT（常规无效流量）数据分析

### GIVT——分细项：

2019年GIVT流量中，各细项之间占比最高的是曝光/点击过快 22.8%和重复流量 22.4%，规律曝光/点击、无效UA流量、黑名单IP流量也是常见的GIVT。



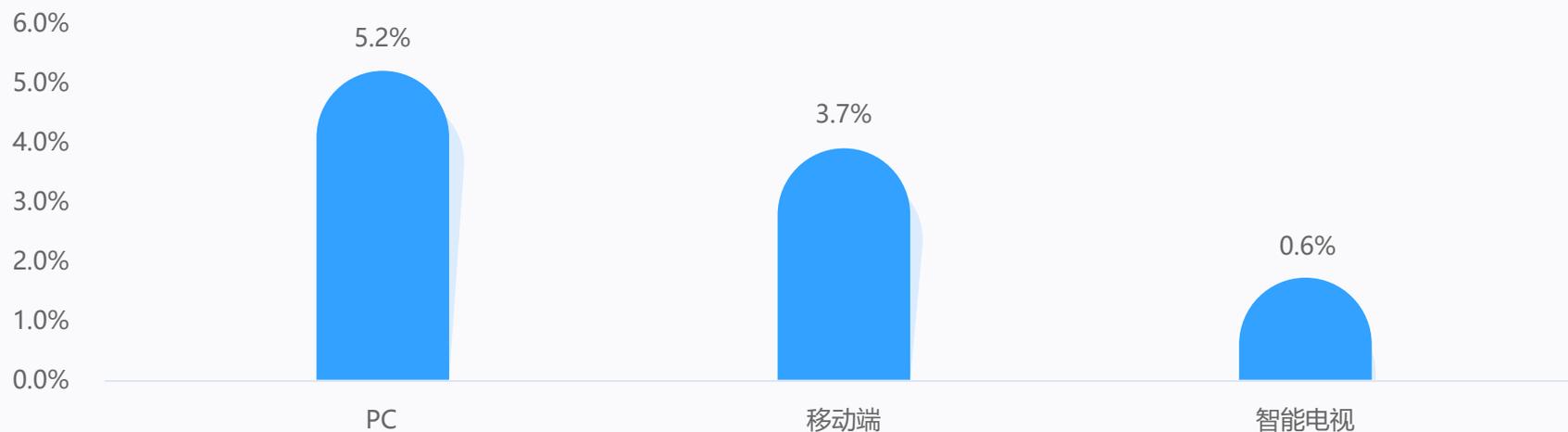
数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## GIVT（常规无效流量）数据分析

### GIVT——分设备分析：

2019年，国内数字广告 GIVT（常规无效流量）分设备占比 PC 端占比最高 5.2%，移动端 3.7%，智能电视端 0.6%。

常规异常流量分设备占比

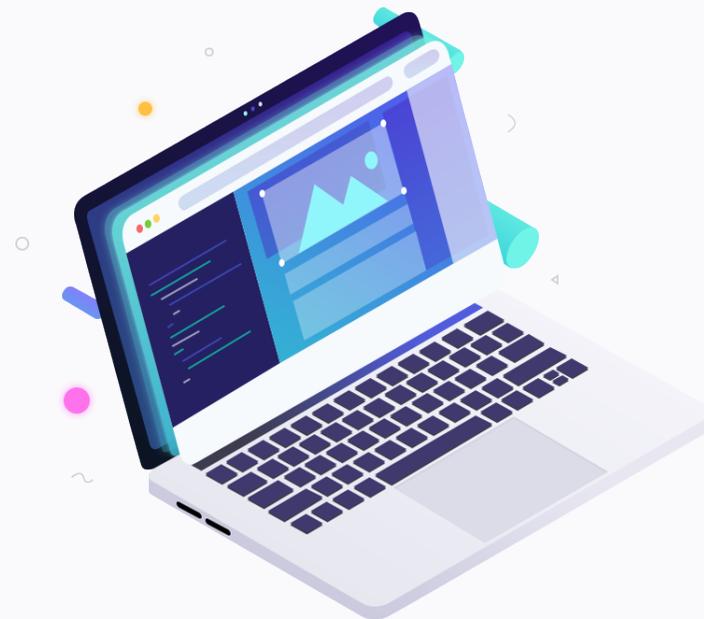


数据来源：Gridsum Ad Dissector 数据时间：2019年1月1日-2019年12月31日

## 小结

通过对数据的洞察发现，2019 年异常流量总体均有明显提升，可见广告行业异常流量的情况仍然非常恶劣。2019 年，行业均值在 32%，主要包含 GIVT 和 SIVT 两个部分，而 GIVT 的占比均值在 3.4%，对难以甄别出的异常流量，我们需要更智能的技术手段有效判定出复杂异常流量。一方面是因为当前行业的黑名单量级还不足，另一方面也是因为黑产各类作弊的手段越来越复杂越来越高明，简单常规的规则已经难以甄别出异常流量，我们需要更智能的技术手段有效判定出复杂异常流量。

# 03 国双监测方案



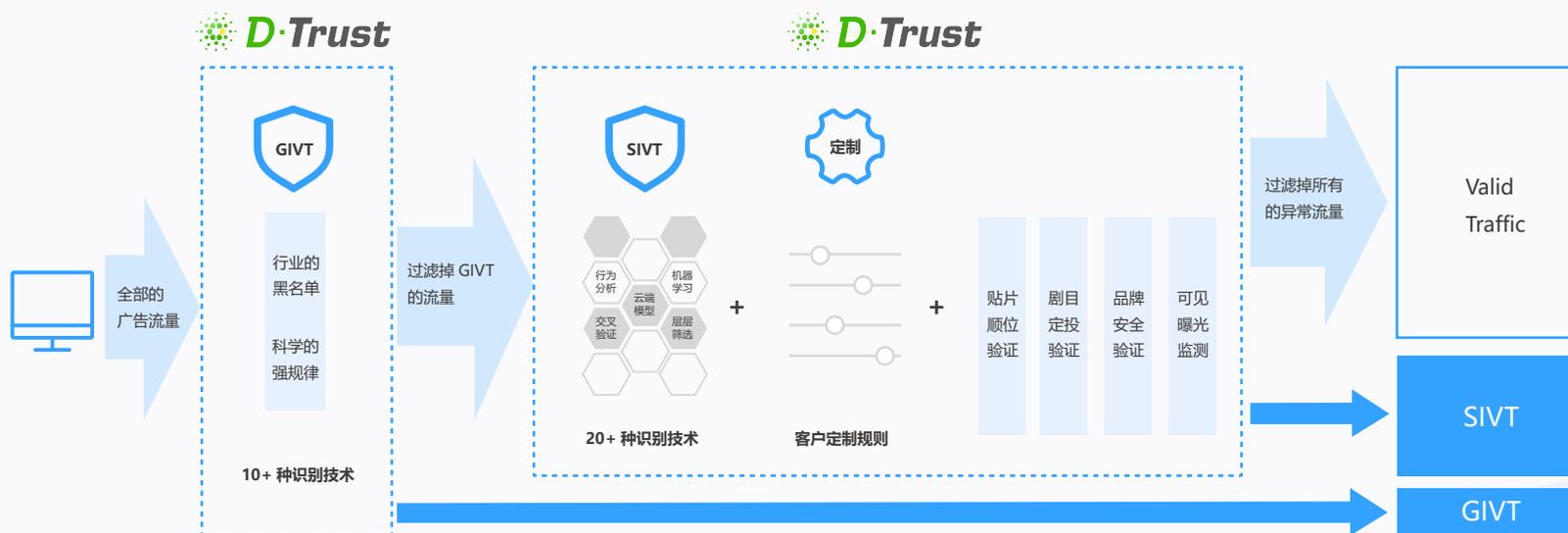
## 国双监测方案

多年以来，国双深耕广告监测领域，解决了客户多种多样的广告监测和验证问题，积攒了丰富的实践经验。在服务的过程中，我们也逐渐发现，客户越来越注重广告效果和数据质量，对于第三方监测公司识别和过滤异常流量也提出了更高的要求。为了更好地服务客户，国双进一步整合并升级了已有的异常流量甄别技术，2019 年国双正式推出全新异常流量识别解决方案 D-Trust。其中 D 代表 Digital Data，因此 D-Trust 即显示了我们想为数字营销带来更多透明度的使命，也表达了我們想为客户带来更多信任的责任。D-Trust 对一般无效流量 GIVT 和复杂无效流量 SIVT 分别提供了针对性的方案解决不同类别的异常流量。

作弊手段总是不断翻新、不断升级的，所以融合了多项黑科技的 D-Trust 并不是一个静止的体系，而是会顺应市场的变化，而不断进行自我升级，淘汰无效的识别手段，增加新式的识别技术。一方面保证识别的准确率，另一方面也能及时发现潜在的异常流量。

## 国双广告反欺诈数据验证流程

1. 通过将监测到的 IP、DeviceID 与 权威的行业 GIVT List 中的 IP、Device ID 进行比对，及经过验证的合理规则，来识别并过滤 GIVT。
2. 通过收集与广告相关的各种参数、传感器信息和设备信息等，并依赖数学模型、机器学习算法等共同构建识别 SIVT 的技术手段，让高阶作弊流量无处可藏。
3. 客户定制属于自己的异常规则，由国双通过技术手段实现，从而实现更细分场景的异常识别；
4. 对于视频广告投放错误、品牌安全和可见曝光问题，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

前后端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

监播实录技术有效规避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问题

专业的代码监测工具及时发现人为错误

完善的历史数据库做参考

## 热力图强力识别验证

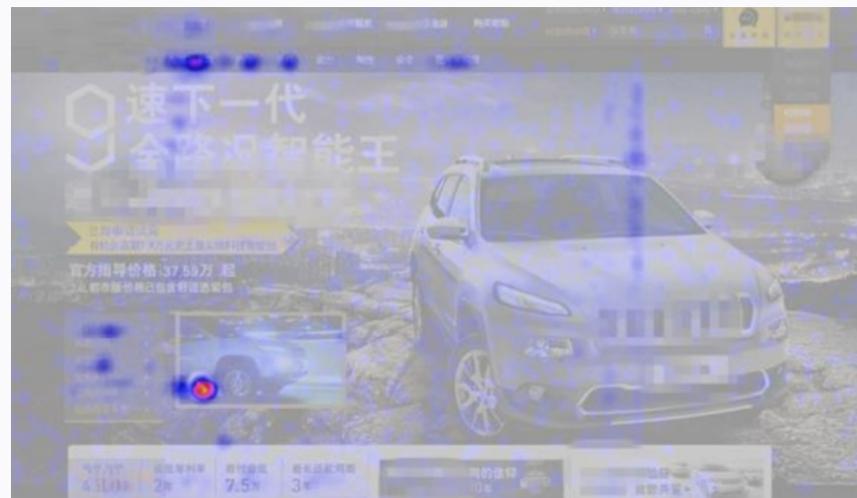
广告来源	访问量	综合浏览量	平均停留时间	跳出率	平均页面访问数	Sales leads完成数	Sales leads转化数
媒体B	774,950	1,745,938	00:02:26	56.60%	2.25	1.741	0.22%

媒体 B 的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标点击分布，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。

媒体 B 媒体 B 的实际效果需重新评估



媒体 A

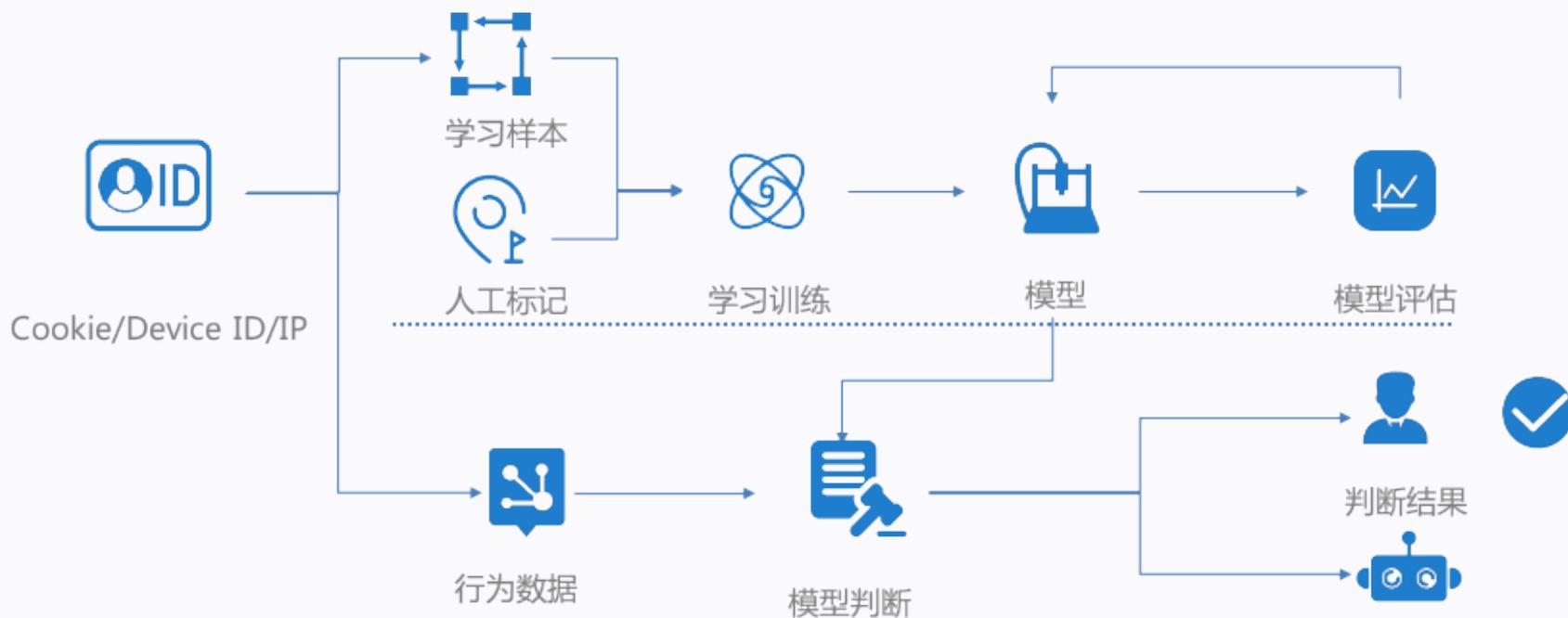


## 异常流量形式及排查原则

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"> <li>• 连续曝光</li> <li>• Stable 人群占比</li> <li>• 时差异常</li> <li>• 周期性波动异常</li> <li>• 定向人群异常</li> <li>• 稳定率异常</li> <li>• 频次分布异常</li> <li>• 曝光不可见</li> <li>• .....</li> </ul>	<h3>前端</h3> <ul style="list-style-type: none"> <li>➤ 曝光或点击的频次 (ID/IP) 超过设定的阈值</li> <li>➤ 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</li> <li>➤ 曝光时间分布不符合网民活跃规律</li> <li>➤ 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。</li> <li>➤ 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。</li> </ul>
	<h3>后端</h3> <ul style="list-style-type: none"> <li>✓ 热力图 (点击热力图、像素热力图)</li> <li>✓ 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</li> <li>✓ 行为路径相同、销售线索重复、用户行为特征</li> <li>✓ .....</li> </ul>

## 基于AI自动发现复杂作弊特征

基于固定规则的方法基本能够帮助我们过滤掉 60%-70% 的异常流量，剩下的部分是由于刷量技术不断升级绕开了规则。对于这些可以绕开规则的异常流量，国双采取基于机器学习的方式来反制。通过应用 AI 技术去自动发现作弊规律，不断学习改进模型，帮助我们识别那些不太容易被甄别出来的作弊。



# 关于国双

国双 (Nasdaq: GSUM) 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悉数据间的复杂关系，获得全新的商业洞察，以利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，请访问官方网站：[www.gridsum.com](http://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