移动数据分析

用户指南

为了无法计算的价值 | 阿里云

用户指南

对接个性化推荐

1. 简介

移动数据分析(Mobile Analytics)提供的小时级别日志数据源可以由App开发者加工成规范的行为数据后对接到推荐引擎，进而快速实现个性化推荐服务。

2. 使用方法

2.1 开通日志分析

只有获取到日志数据后，App开发者才能进一步加工为推荐引擎所需的数据格式，开通步骤参见自定义数据分析

2.2 了解推荐引擎

产品介绍

快速入门

数据格式规范

其中数据格式规范这一页与移动数据分析密切相关，其中的行为表（user_behavior）的数据源可以很方便从移动数据分析采集的数据获取到。下面就来介绍推荐引擎需要的字段，移动数据分析的怎么埋点？在哪里取埋点内容？

字段名	*类型*	*注释*	*Nullable*	*SDK埋点*	*SDK日志字段*
user_id	string	用户ID	否	取设备ID是无需埋点自动采集，取会员ID是需要调用updateUserAccount()来传入会员ID	取设备ID可使用utdid或imei字段，取会员id用user_id字段

item_id	string	物品ID	否	页面调用updatePageProperties()或自定义事件调用setProperty()以kv方式埋入物品ID值	取日志args字段埋入的kv值，一般用ODPS的keyvalue()可以很方便解析出来
bhv_type	string	行为类型：;view：物品曝光;click：用户点击物品;collect：用户收藏了某个物品;uncollect：用户取消收藏某个物品;search_click：用户点击搜索结果中的物品;comment：用户对物品的评论;share: 分享;like：点赞;dislike：点衰;grade：评分;consume：消费;use：观看视频/听音乐/阅读;行为表记录的用户行为用于用户偏好建模	否	这一类行为事件一般埋自定义事件，具体参见SDK手册	在日志数据中用event_id=19999筛选出自定义事件后的arg1即为自定义的行为埋点
bhv_amt	double	用户对物品的评分、消费、观看时长等。	否	页面调用updatePageProperties()或自定义事件调用setProperty()以kv方式埋入物品ID值	取日志args字段埋入的kv值，一般用ODPS的keyvalue()可以很方便解析出来
bhv_cnt	double	行为次数，默认为1，消费可以埋购买件数	否	页面调用updatePageProperties()或自定义事件调用setProperty()以kv方式埋入物品ID值	取日志args字段埋入的kv值，一般用ODPS的keyvalue()可以很方便解析出来

bhv_datetime	datetime	行为发生的时间，UTC格式。	否	SDK自动采集	取日志的localtime字段
content	string	用户对物品的评价文本	是	自定义事件调用setProperty()以kv方式埋入物品ID值	取日志args字段埋入的kv值，一般用ODPS的keyvalue()可以很方便解析出来
media_type	string	如果bhv_type=share，该字段记录分享到目标媒体。短信：sms，邮件：email，微博：sina_wb，微信好友：wechat_friend，微信朋友圈：wechat_circle，QQ空间：qq_zone，来往好友：laiwang_friend，来往动态：laiwang_circle	是	自定义事件调用setProperty()以kv方式埋入物品ID值	取日志args字段埋入的kv值，一般用ODPS的keyvalue()可以很方便解析出来
pos_type	string	行为发生的位置及类型	是	SDK没有采集经纬度	默认pos_type 取‘poi’
position	string	行为发生的位置	是	SDK没有采集经纬度	可在日志中取province或city字段
env	string	JSON String：IP、network、device等其他自定义环境变量	是	SDK自动采集	日志中取ip、network_type、brand、device_model等字段
trace_id	string	返回的推荐列表用于跟踪效果。如果对item_id的行为不是	是	该ID可以埋入行为的属性kv值里面	分析效果时候取日志args字段埋入的kv值

		来自推荐引导，则为NULL			

*注：不同行为bhv_type要埋点不同的参数，具体参见帮助文档，如物品点击click需埋点商品金额bhv_amt参数，用品评论comment需埋点content参数

2.3 加工行为数据

示例：各个城市的用户把商品加入心愿单的行为那首先需要用ODPS DDL创建数据表，如下：

```
CREATE TABLE user_behavior (

user_id STRING,
item_id STRING,
bhv_type STRING,
bhv_amt DOUBLE,
bhv_datetime DATETIME,
pos_type string,
position string
)
PARTITIONED BY (ds STRING);
```

接着，按上述的取埋点内容方法，将内容对应更新到相应字段：

```
   INSERT OVERWRITE TABLE user_behavior PARTITION (ds='20160101')

select
utdid AS user_id
, keyvalue(args, '\;', ':', 'itemid的key值') AS item_id
, 'favorite' AS bhv_type
, keyvalue(args, '\;', ':', 'item金额的key值') AS bhv_amt
, local_time AS bhv_datetime
,'poi' as pos_type
, city as position

FROM SDK日志表名
```

RAM子账号访问

使用RAM子账号访问移动数据分析

第一步：主账号登录RAM控制台，点击左侧“策略管理”菜单，添加自定义授权策略：

EMAS-MAN-FullAccess 管理移动数据分析(MAN)权限

```
{
"Version": "1",
"Statement": [
{
"Action": "man:*",
"Resource": "*",
"Effect": "Allow"
}
]
}
```

第二步：然后点击左侧“用户管理”菜单，授权某个子账号该权限并同时给用户授权系统策略

AliyunMHubFullAccess

第三步：确保在子账号登录状态下，访问移动数据分析控制台，即可。

注：目前移动数据分析不支持针对App粒度的权限控制，因此，以上授权方法，子账号能查看到主账号拥有的所有App数据，请谨慎使用。

指标说明

用户及行为类

活跃用户

简要描述：指定时间段内，使用过应用的独立用户数

计算口径：
android/aliyunos：当天有页面(2001,2002)行为

ios/wp：全部行为(包含所有事件)

新增用户

简要描述：指定时间段内，第一次使用应用的独立用户数

计算口径：
用户是活跃用户，并且这些用户的imeisi在APP上第一次出现

登录会员

简要描述：指定时间段内，登录过应用的独立会员数，会员ID是通过开发者调用接口updateUserAccount()传入的nick

计算口径：
用户是活跃用户，并且长登陆nick不为空

新注册用户

简要描述：指定时间段内，通过应用新注册的独立会员数,会员ID是通过开发者调用接口userRegister()传入的nick

升级用户

简要描述：指定时间段内，有版本升级行为的独立设备数；按版本计算时，即升级到该版本的独立设备数

计算口径：
当天升级到最高版本的用户数(如果当天用户只有一个版本，还需要跟昨天的最高版本比较)

新增登陆会员

简要描述：指定时间段内，首次登录应用的独立会员ID数在指定时间内的总和

计算口径：
用户是活跃用户,并且这些用户的长登陆nick在APP上第一次出现

使用时长

简要描述：指定时间段内，应用总使用时长，单位：秒

衍生指标：次均使用时长 = 使用时长 / 启动次数

计算口径：
通过USDK采集的应用启动事件来计算，启动时长 > 0的才会计入。

启动次数

简要描述：指定时间段内，用户启动应用的总次数

计算口径：
通过SDK采集的应用启动事件来计算，启动时长 > 0的才会计入。

页面访问次数

简要描述：指定时间段内，页面的被访问次数（pv）

统计口径：
SDK采集的页面事件日志，每条页面事件日志被记为一次访问次数

页面停留时长

简要描述：指定时间段内，用户访问当前页面的停留时间。

页面退出率

简要描述：在该页面关闭或退出客户端的次数 / 该页面的访问次数实际汇总用减法完成——根节点页面的访问次数-跳转到其他页面的次数 / 根节点页面访问次数。

计算口径：
页面进入次数 / 页面进入次数 - 跳转次数

活跃度

简要描述：指定时间段内，应用指定版本的活跃用户数 / 指定时间内的累计用户数

计算口径：
日活跃度：一天内应用指定版本的活跃用户数 / 指定时间内的累计用户数
周活跃度：一周内应用指定版本的活跃用户数 / 指定时间内的累计用户数
月活跃度：一月内应用指定版本的活跃用户数 / 指定时间内的累计用户数

留存率

周期留存率

计算口径：
活跃 / 新增用户在+N周期内的留存情况

每日留存率

计算口径：
活跃 / 新增用户在+N日后的留存情况

新增留存率

计算口径：
新增用户在 +N 日后的留存情况

活跃留存率

计算口径：
活跃用户在 +N 日后的留存情况

性能类

crash次数

简要描述：系统指定版本异常退出的次数。

crash率

简要描述：系统指定版本异常退出的次数在该版本中所有启动次数比例

计算口径：
crash次数 / 启动次数

Crash用户数

简要描述：系统指定版本异常退出影响的独立用户数

次均TCP建连时间

简要描述：网络请求中TCP建立连接的平均耗时，单位毫秒

次均首字节到达时间

简要描述：网络请求中首字节到达的平均耗时，单位毫秒

次均请求完成时间

简要描述：网络请求完成的平均耗时，单位毫秒

次均请求资源大小

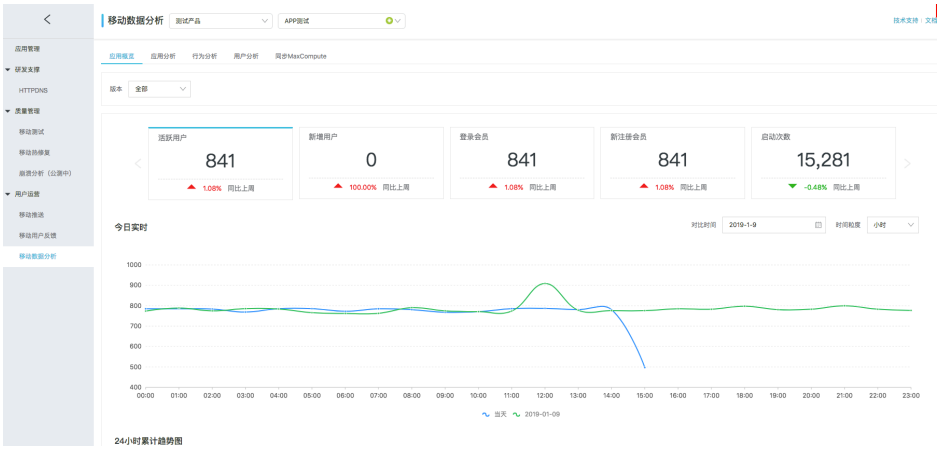
简要描述：网络请求完成的平均耗时，单位毫秒

异常数量

简要描述：发生异常的网络请求数量

应用概览

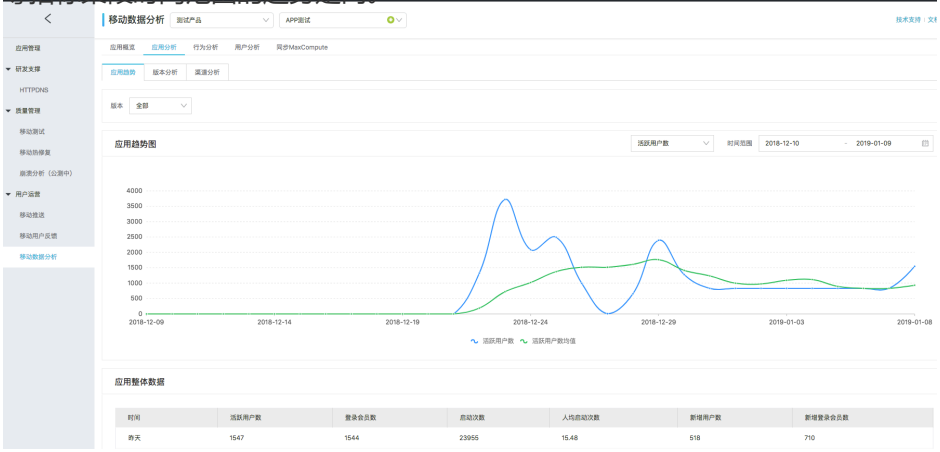
可以查看秒级实时的活跃用户、新增用户、登录会员、新注册会员、启动次数等信息。同时还可以看到24小时的累计信息以及小时粒度的明细数据。



应用分析

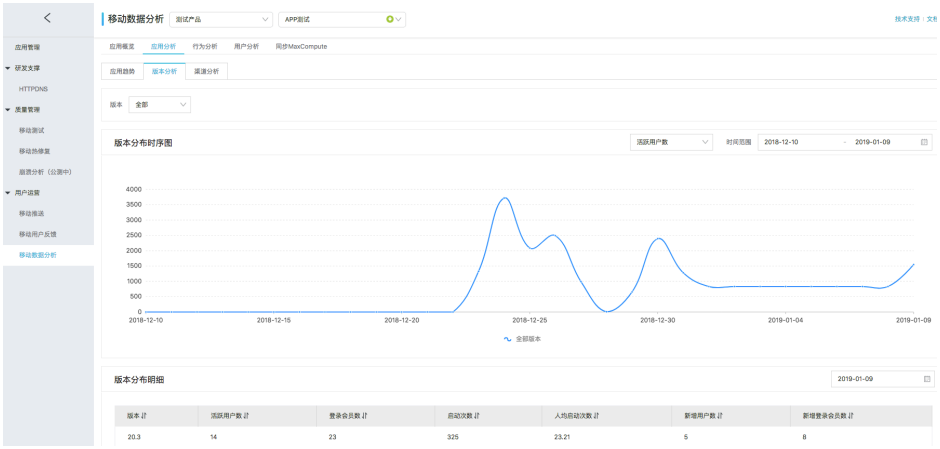
应用趋势

通过应用趋势可以看到活跃用户、新增用户、登录会员、新增登录会员、新注册会员、人均启动次数、活跃度等指标某段时间范围的趋势走向。



版本分析

版本分析可以查看按版本维度、每天的活跃用户、新增用户、登录会员、新增登录会员、新注册会员、人均启动次数、活跃度等指标趋势。默认展示top5版本(按活跃用户数)，可以选择对比的其他版本。



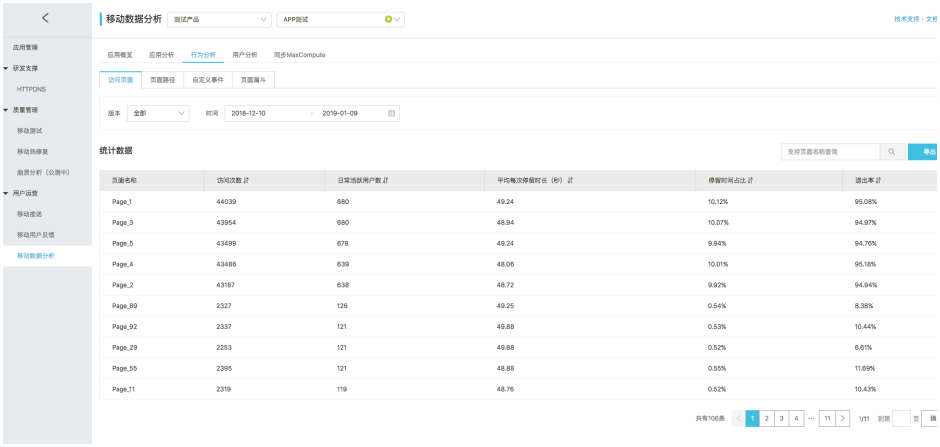
渠道分析

渠道分析可以查看不同版本Top5、Top10渠道天粒度的活跃用户数、登录会员数、启动次数、人均启动次数、新增用户数、新增登录会员数的分布时序图和渠道的明细数据展示

行为分析

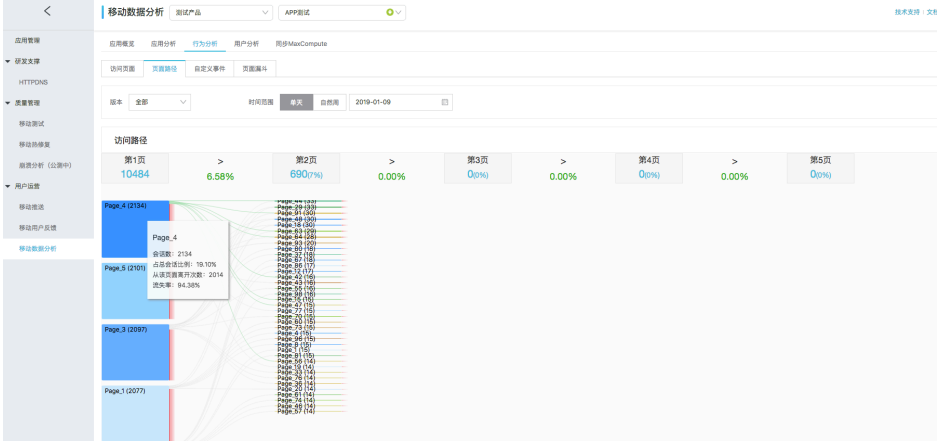
访问页面

展示了用户在APP页面访问情况：访问次数、活跃用户数、平均每次停留时长、停留时间占比、退出率等。一般地，页面采集在SDK初始化后是自动打开的。如果同时采集的H5页面，在这里页面名称将显示为url路径。



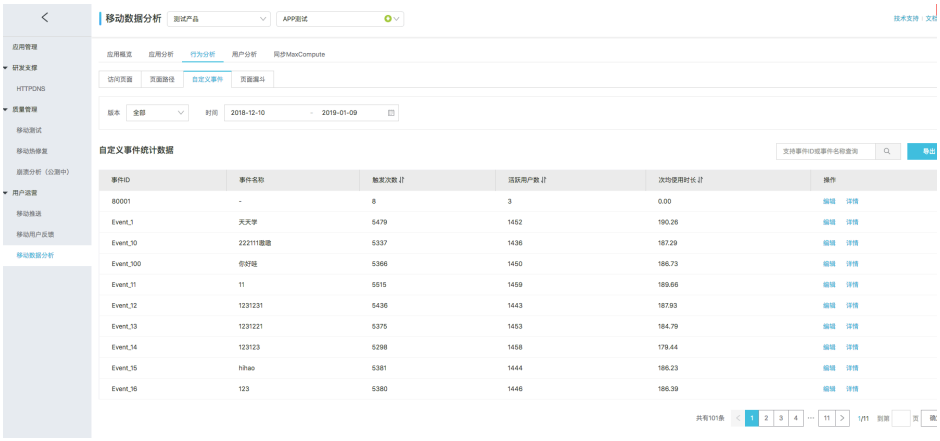
页面路径

展示了用户访问的Top50链路中的页面操作路径，并对页面的会话数、会话占比、页面离开次数、流失率等。清晰展示当前APP用户的流量分布和流转情况。



自定义事件

自定义事件是为了能更方便的统计应用的一些个性化的事件，如应用的点击事件、后台事件、音乐播放时长、游戏play时长、读书阅读时长等。自定义事件无需开发者进行注册，即可自动运算调用SDK自定义事件接口埋点的自定义事件数据，建议埋点事件id尽量使用字母、数字及下划线组合，你可以在“自定义事件-编辑”给自定义事件命名“中文别名”，方便业务人员查看数据。同时，与页面事件一样，点击“查看详情”可以查看某



个自定义事件的访问趋势。

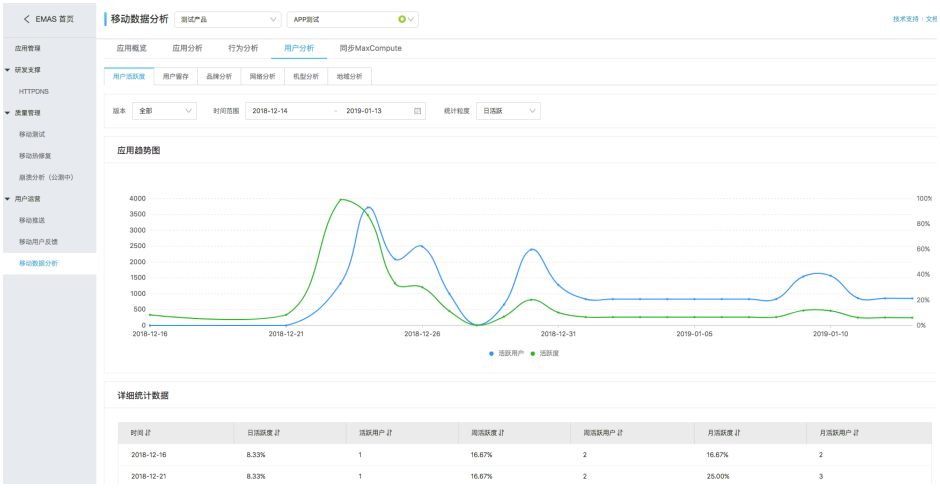
页面漏斗

页面漏斗提供了让开发者自定义页面路径转化分析的能力，比如设定A页面到B页面再到C页面为一个业务非常关心的用户转化路径，假设某天B页面做了新版改造，业务肯定非常关心是否有影响用户的转化率，再针对进一步细致的数据分析来优化产品体验。如下图创建漏斗，最多6个页面。

用户分析

用户活跃度

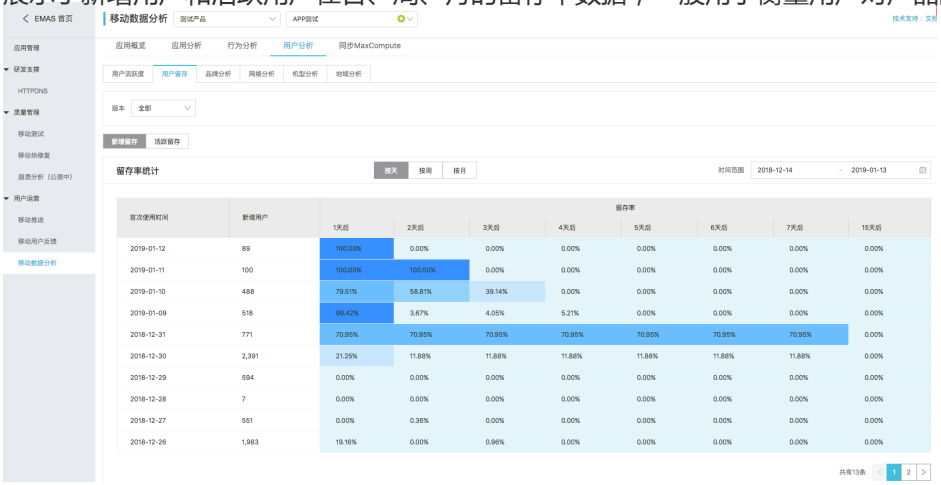
平台以2012-1-1开始以来的累计用户（即设备）作为基数，无论是日、周、月活跃度都可以用于衡量用户对APP的访问黏性。以活跃用户和活跃度双指标来显示APP的活跃趋势情况。支持版本、时间范围、统计粒度等



维度的聚合展示。

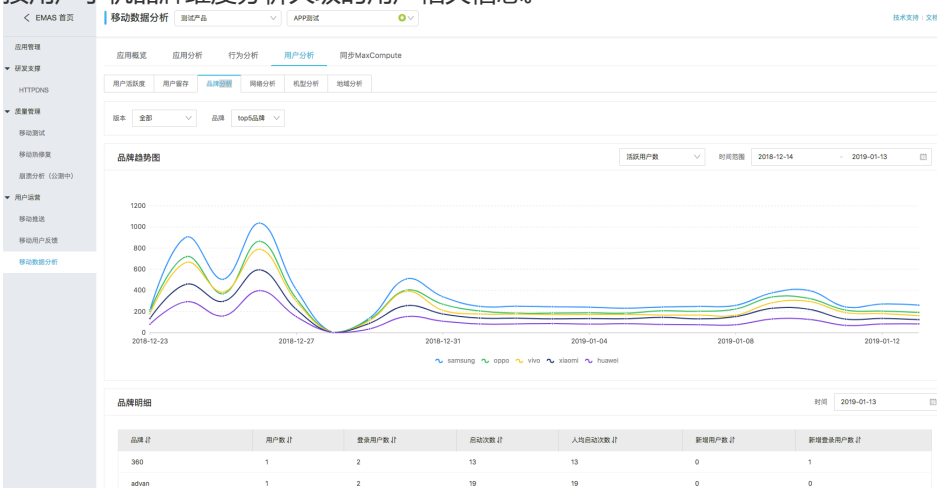
用户留存

展示了新增用户和活跃用户在日、周、月的留存率数据，一般用于衡量用户对产品的黏性好坏。



品牌分析

按用户手机品牌维度分析天级的用户相关信息。



网络分析

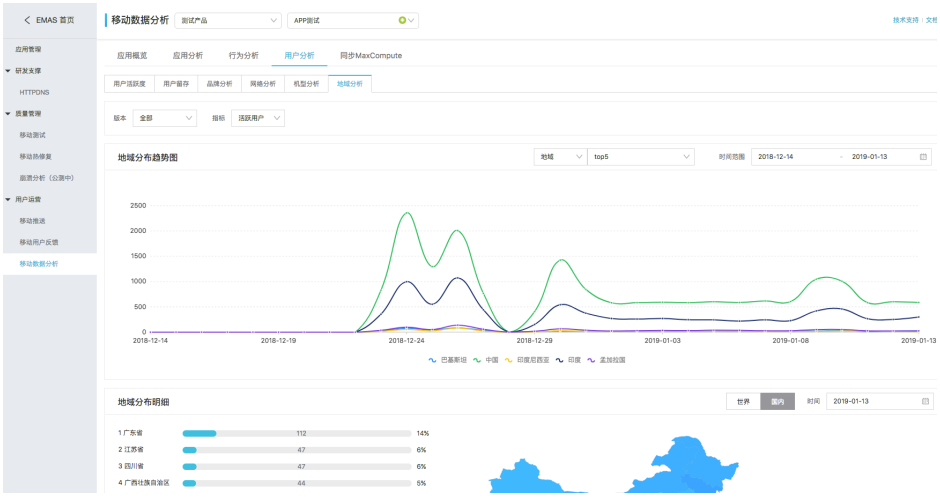
按Wi-Fi、4G、3G、2G、2G/3G、UNKNOWN等网络维度分析指定时间段的用户变化趋势。

机型分析

按用户手机机型维度分析天级的用户相关信息。

地域分析

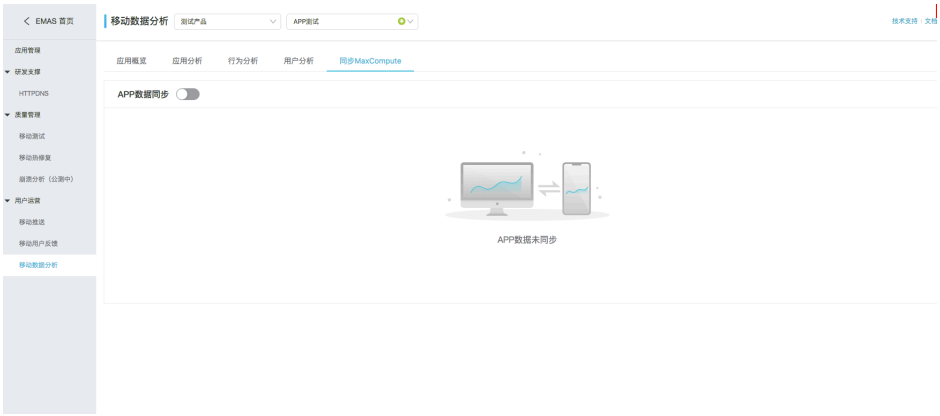
地域分析可以查看按地域分布维度、天级的活跃用户、新增用户、登录会员、新增登录会员、新注册会员、人均启动次数、活跃度等信息的图表展示。地域是通过采集的IP地址来解析的。



同步MaxCompute

同步MaxCompute

APP数据同步，支持将账号下所有 App 数据增量同步（15分钟间隔）到 MaxCompute 指定 region 的 project 中，延迟约为一小时。



操作路径

1. 使用主账号
2. 采访数加控制台： https://data.aliyun.com/console
3. 点击 个人信息
4. 点击 修改AccessKey信息
5. 输入 Access Key ID 和 Access Key Secret。如果没有 key 和 secret 或者不知道，可以点击访问控制获取
6. 采访移动数据分析 - 同步MaxCompute

7. 点击APP数据同步

上报的间隔与策略：

前台30s，后台5分钟，无网络情况下存储3天最多存储9000条日志。

日志数据内容字段

当前开放字段有：

字段名	*类型*	*注释*
app_id	string	以appkey@os的形式表示
app_name	string	app_id对应的app中文名称
app_version	string	app的应用版本号
channel	string	应用分发渠道
imei	string	移动设备国际身份码的缩写
imsi	string	国际移动用户识别码
brand	string	手机或终端的品牌
device_model	string	手机或终端的机型
resolution	string	手机或终端的屏幕分辨率
os	string	操作系统，如: Android、iPhone OS
os_version	string	操作系统的版本
carrier	string	移动运营商，如：中国移动、中国联通、中国电信
access	string	连接的网络，如：2G、3G、Wi-Fi
access_subtype	string	网络类型，如：HSPA、EVDO、EDGE、GPRS等
network_type	string	根据access,acess_subtype转化后的网络类型
school	string	根据client_ip如果为校园网解析出的学校
client_ip	string	客户端ip
longitude	string	经度，目前SDK没有采集，有需求请联系我们
latitude	string	纬度，目前SDK没有采集，有需求请联系我们
country	string	根据client_ip解析出的国家或地区

province	string	根据client_ip解析出的省、直辖市、自治区
city	string	根据client_ip解析出的地级市
district	string	根据client_ip解析出的所在区域，对应如 华南，西北等
session_id	string	用户的一次会话id
reach_time	string	到达日志服务器的时间，此时间可作为日志时间直接使用，格式为：yyyyMMddHHmmss
event_id	string	埋点的事件ID，事件ID为2001，page是表示当前页面，arg1表示上一个页面；事件ID为19999，page是默认page_extend,可埋点重写，arg1表示自定义事件名称
page	string	页面
arg1	string	事件参数，事件ID为1010时为上次启动的前台停留时间，单位毫秒
arg2	string	事件参数，事件ID为1010时为此次启动之前的后台停留时间，单位毫秒
arg3	string	事件参数
args	string	事件参数，调接口setProperty()等埋点的KV属性串
local_time	string	终端时间(格式为yyyy-mm-dd hh24:mi:ss)
local_timestamp	string	终端时间(格式为数字型的unix时间,精确到毫秒,可通过from_unixtime函数转换成日期)
utdid	string	服务端生成的设备唯一标识符
user_nick	string	长登录会员名称，长登录是指只要登录一次就会记住该设备最近一次登录会员，即使该设备下一次打开App且没有登录，其日志也会记录该设备最近一次登录会员
user_id	string	长登录会员id
short_user_nick	string	短登录会员名称，短登录是指当前处于登录状态的会员
short_user_id	string	短登录会员id
ds	string	分区字段，表示日期，一般格式

		为yyyymmdd
hh	string	分区字段，表示小时，一般格式为hh
mm	string	分区字段，表示分钟，一般格式为mm,15分钟一个间隔，如00 15 30 45

事件ID（即event_id）类型如下：

事件ID	***含义***	***是否自动采集***
2001	页面事件	自动采集
19999	自定义事件	调接口 setEventLabel() 埋点
1010	退出事件	自动采集，一般用于计算启动次数
1006	注册事件	调接口 userRegister() 埋点
1007	登录事件	调用接口 updateUserAccount() 埋点
3002	性能事件	网络请求
3003	性能事件	带宽利用率
3004	性能事件	错误详情
3005	性能事件	CAS加速效果
3015	性能事件	域名劫持事件
66602	性能事件	自定义性能事件
1	性能事件	android手机java crash
61006	性能事件	android手机native crash，iOS crash

事件规则：

事件ID	**page**	**arg1**	**arg2**	**arg3**	**args**	**上报时机**
1010	在运行过程中切换页面则取当前页面，程序退出取Page_UT	本次前台运行时间，单位毫秒	这次启动前的后台停留时间，单位毫秒	无	无	停留在后台时或下次启动
2001	现页面ID	上一跳页面ID	上一个控件名	页面停留时间，单位毫秒，不算后台停留时	页面参数KV对，k1=v1,k2=v2	上报时机为页面离开后

				间		
19999	现页面ID	event_id	无	使用时长	自定义事件参数KV对，k1=v1,k2=v2	点击后上报

日志中的三类时间：

1. local_time: 设备的时间，因为上报延迟，可能与另外两个时间相差较大
2. reach_time：服务器时间，日志到达服务器的时间
3. partition_time（ds,hh,mm）: 分区时间，采集服务器数据到达 Maxcompute 源表的时间。

实时日志

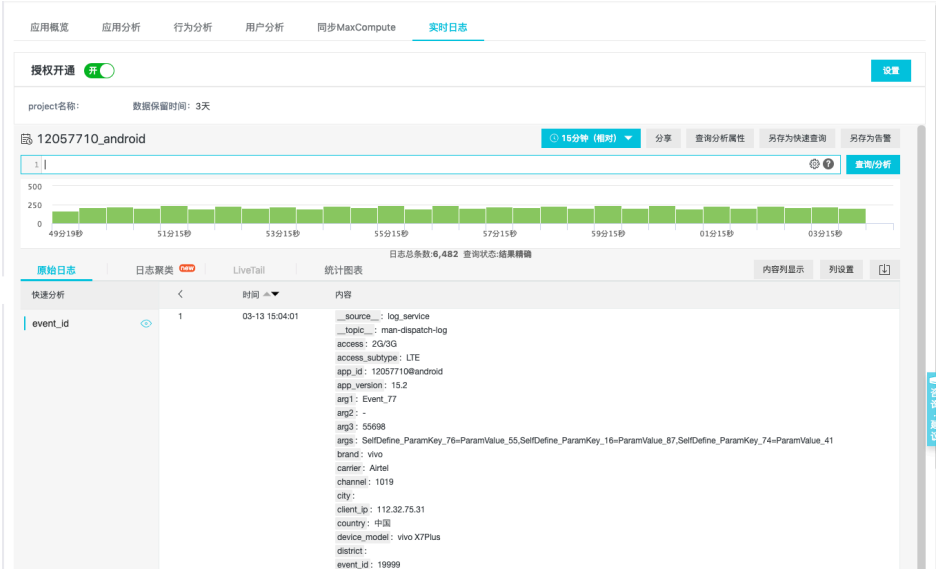
开启实时日志

这里需要三个步骤

1. 开通阿里云 SLS 日志服务
2. 授予数据分析写日志服务权限
3. 选择 Region

操作成功后，我们会帮您在您的日志服务中自动创建好 Project 和 Logstore, 其中 Project 名称为 man-{user_id}-project-{region_id}, Logstore 为 {appkey}_{platform}。注：该 Logstore 为打标 Logstore，商业化之前不对用户收费，也就是目前免费。商业化前我们会咨询客户是否继续同步，在征得客户同意后才会收费，商业化日期yue为19年5-6月。

您既可在控制台中看到实时日志搜索框。



使用实时日志

目前，您可以对实时日志进行如下方面的使用

1. 验证埋点：对于初次接入或者测试应用。
2. 投递 OSS，maxcompute：对于想有离线存储计算的需求。
3. flink 主动消费：有较强研发能力，且有自主实时分析，实时洞察的需求。
4. 搜索框搜索分析（目前只开启了event_id索引，商业化后会根据客户选择是否开启全部索引）：只是简单的统计，洞察。

其中2，3，4 请移步 SLS（日志服务） 文档

日志格式

字段名	*类型*	*注释*
app_id	string	以appkey@os的形式表示
app_name	string	app_id对应的app中文名称
app_version	string	app的应用版本号
channel	string	应用分发渠道
imei	string	移动设备国际身份码的缩写
imsi	string	国际移动用户识别码
brand	string	手机或终端的品牌
device_model	string	手机或终端的机型
resolution	string	手机或终端的屏幕分辨率
os	string	操作系统，如: Android、iPhone OS

os_version	string	操作系统的版本
carrier	string	移动运营商，如：中国移动、中国联通、中国电信
access	string	连接的网络，如：2G、3G、Wi-Fi
access_subtype	string	网络类型，如：HSPA、EVDO、EDGE、GPRS等
network_type	string	根据access,acess_subtype转化后的网络类型
client_ip	string	客户端ip
country	string	根据client_ip解析出的国家或地区
province	string	根据client_ip解析出的省、直辖市、自治区
city	string	根据client_ip解析出的地级市
district	string	根据client_ip解析出的所在区域，对应如 华南，西北等
session_id	string	用户的一次会话id
reach_time	string	到达日志服务器的时间，此时间可作为日志时间直接使用，格式和maxcompute 有差异，这里为 timestamp
event_id	string	埋点的事件ID，事件ID为2001，page是表示当前页面，arg1表示上一个页面；事件ID为19999，page是默认page_extend,可埋点重写，arg1表示自定义事件名称
page	string	页面
arg1	string	事件参数，事件ID为1010时为上次启动的前台停留时间，单位毫秒
arg2	string	事件参数，事件ID为1010时为此次启动之前的后台停留时间，单位毫秒
arg3	string	事件参数
args	string	事件参数，调接口setProperty()等埋点的KV属性串
local_time	string	终端时间(格式为yyyy-mm-dd hh24:mi:ss)
local_timestamp	string	终端时间(格式为数字型的unix时间,精确到毫秒,可通过from_unixtime函数转换成日期

		)
utdid	string	服务端生成的设备唯一标识符
user_nick	string	长登录会员名称，长登录是指只要登录一次就会记住该设备最近一次登录会员，即使该设备下一次打开App且没有登录，其日志也会记录该设备最近一次登录会员
user_id	string	长登录会员id
short_user_nick	string	短登录会员名称，短登录是指当前处于登录状态的会员

更多日志规则可移步 同步 Maxcompute。