bigreco大数据统计、分析、智能推荐系统

开发文档

根据某些系统的统计、分析、推荐需求，结合功能的独立化\模块化\标准化的要求，设计并开发了本系统。

# 架构设计

bigreco 大数据统计、分析、智能推荐系统采用分层架构方式进行设计，具有运行独立化、核心功能模块化、数据对象标准化（抽象化），使用简单（大部分使用直接嵌入JS然后插入相应事件监听即可），对象扩展方便，算法叠加容易等特点。

## 第一层:客户层

对于bigreco而言,所有用户均为一个客户存在，不于任何其他项目深度耦合,以保证本系统的独立性与开放性。

## 第二层:对象层

所有需要统计的数据,根据OOP设计思路,被划分为多个独立的对象,例如目前支持的 用户对象,商品对象(商品可以是任何商品或者书籍)等,根据以上设计，可保证系统的最大兼容性与后期二开便捷性。

## 第三层:事件层

所有对象发生的数据变化，均被视为一次事件而被系统记录，该层仅仅进行必要的数据记录以及数据预处理.以达到数据标准化的目的。

## 第四层:计算层

本层由多个不同的数据统计方法组成，与第二层第三层无直接隶属关系，每一个统计方法为一个模块，根据不同模块的启动时机与要求分为触发式模块与定时式模块.后续新增任何统计方法，即可对所有对象生效。

## 第五层:接口层

该层负责统计数据结果的输出。

# 管理接口（客户层）

## 应用创建接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=app\_create | | | |
| app\_name | varchar(50) | 不为空 | 应用名称 |
| app\_exp | varchar(200) | 不为空 | 应用简介 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，用户管理应用需要该密钥，必须是5-50位大小写字母数字，区分大小写。 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥，用户提交行为需要附带该密钥，必须是5-50位大小写字母数字，区分大小写 |
| app\_stat\_type | public|private | 可空，默认private | 该应用是否公开统计结果 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| app\_id | int | >0 | 创建的应用序号 |

## 应用修改接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=app\_modify | | | |
| app\_id | int | 不为空 | 需要修改的应用序号 |
| app\_pass | varchar(50) | 不为空 | 老的应用管理密码 |
| app\_name | varchar(50) | 空 | 新的应用名称 |
| app\_exp | varchar(200) | 空 | 新的应用简介 |
| app\_password | varchar(50) | 空 | 应用管理密码，用户管理应用需要该密钥，必须是5-50位大小写字母数字，区分大小写。 |
| app\_key | varchar(50) | 空 | 应用提交密钥，用户提交行为需要附带该密钥，必须是5-50位大小写字母数字，区分大小写 |
| app\_reco\_data | int | >1<90 | 本应用推荐1个周期自然天数，默认30，最大90天 |
| app\_stat\_type | public|private | 可空，默认private | 该应用是否公开统计结果 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |

# 对象定义（对象层）

## 用户对象

所有具有编号、帐号、密码等特征数据以及具有注册、登录等行为的资料，被统一取模为用户对象。

## 商品对象

所有具有名称、编号、价格、简介等特征数据以及具有购买、交易等行为的数据，被统一取模为商品对象。

## 企业对象

所有具有机构唯一识别号、名称、简介、经营等特征数据以及具有年报、月报等统计行为的数据，统一被取模为企业对象。

# 事件申报接口（事件层）

## 用户事件

### 用户添加|修改上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=user\_modify | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_password | varchar(50) | 不为空 | 应用管理密码 |
| user\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号，如果重复出现则表示修改，不区分大小写的字母数字、最长32位 |
| user\_sex | enum(男|女|保密) | 空 | 用户的性别 |
| user\_age | int | 空 | 用户年龄 |
| user\_phone | varchar(20) | 空 | 用户手机号码，建议填写，该属性可有效关联用户行为 |
| user\_reco | varchar(50) | 空 | 用户推荐来源名称 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_usermodify(user\_no,user\_sex,user\_age,user\_phone,user\_reco)  本函数为异步函数，可不用等待返回 | | | |

### 用户移除上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=user\_remove | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_password | varchar(50) | 不为空 | 应用管理密码 |
| user\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_userremove(user\_no)  本函数为异步函数，可不用等待返回 | | | |

### 用户登录上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=user\_login | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_password | varchar(50) | 不为空 | 应用管理密码 |
| user\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_userlogin(user\_no)  本函数为异步函数，可不用等待返回 | | | |

## 商品事件

### 商品添加|修改上报

该接口如果暂停某个商品的显示，需要修改所有与pro\_no相关的表

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=pro\_modify | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_password | varchar(50) | 不为空 | 应用管理密码 |
| pro\_no | varchar(32) | 不为空 | 商品在本应用中的唯一序号，如果重复出现则表示修改 |
| pro\_name | varchar(50) | 不为空 | 商品名称 |
| pro\_cover | varchar(250) | 空 | 商品封面图片地址，接受全路径 |
| pro\_tags | tinytext | 不为空 | 商品标签名称，以英文逗号分割多个 |
| pro\_show | bool | true | 该商品是否记录进入统计与显示 |
| pro\_grade | int | 空 | 商品显示权重，0-9，当显示为9时强制推荐，不建议超过10个 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_promodify(pro\_no，pro\_name，pro\_cover，pro\_tags，pro\_grade)  本函数为异步函数，可不用等待返回 | | | |

### 商品移除上报

如果该接口移除某个商品，需要同步移除该商品的所有数据

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=pro\_remove | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_password | varchar(50) | 不为空 | 应用管理密码 |
| pro\_no | varchar(32) | 不为空 | 商品在本应用中的唯一序号 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_proremove(pro\_no)  本函数为异步函数，可不用等待返回 | | | |

### 商品分享上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=pro\_share | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| pro\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号 |
| share\_name | varchar(50) | 不为空 | 分享渠道名称 |
| user\_no | varchar(32) | 空 | 用户在本应用中的唯一序号 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_proshare(pro\_no,share\_name,user\_no)  本函数为异步函数，可不用等待返回 | | | |

### 商品查看上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=pro\_play | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| pro\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号 |
| pro\_play | enum(info|play) | 选择，默认info | 查看方式，  info表示详情查看，  play表示内容查看 |
| pro\_seq | varchar(50) | 空 | 如果该商品仍有下级关联，可采用 no\_no\_no方式传入 |
| user\_no | varchar(32) | 空 | 用户在本应用中的唯一序号 |
| pro\_play\_status | enum(new|sleep) | 空 | 表示当前动作是新动作还是之前动作的心跳监听信号，如果是sleep则表示心跳监听信号 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_proplay(pro\_no,pro\_play,pro\_seq,user\_no,pro\_play\_status)  本函数为异步函数，可不用等待返回 | | | |

### 商品购买上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| POST bigrec/?act=pro\_sale | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| pro\_no | varchar(32) | 不为空 | 用户在本应用中的唯一序号 |
| pro\_sale\_no | varchar(32) | 不为空 | 本次购买的订单序号 |
| pro\_pay\_name | varchar(32) | 不为空 | 支付渠道名称 |
| pro\_money | decimal | 不为空 | 当前购买的实际金额数 |
| user\_no | varchar(32) | 不为空 | 购买用户在本应用中的唯一序号 |
| pro\_seq | varchar(50) | 空 | 如果该商品仍有下级关联，可采用 no\_no\_no方式传入 |
| pro\_sale\_referer | varchar(50) | 空 | 本次购买的来源名称,如果是被推荐购买则可以在此处记录推荐信息 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_prosale(pro\_no,pro\_sale\_no,pro\_pay\_name,pro\_money,  user\_no,pro\_seq,pro\_sale\_referer)  本函数为异步函数，可不用等待返回 | | | |

# 统计数据获取接口（接口层）

## 商品接口－自动推荐接口

算法：  
冷启动：当前应用的申报条数少于1万条，则成为冷启动。

推荐规则为：根据当前应用的所有标签，平均随机推荐 size\*100条，然后根据权重倒叙，获取每个标签的前 （size / 标签数） 条。以保证每个标签均有机会被推荐。

热启动：当前应用的申报条数大于1万条，则成为热启动。

推荐规则-用户：根据当前用户的浏览历史记录，获取关心的标签类型，然后根据获取到的标签得出每个标签的获取比例，然后根据比例获取每个标签的数据，加入待选池。根据待选池的数据以及每一条数据的热度，得出命中几率，然后根据随机数取得命中的序号并获取商品。

推荐规则-游客：根据游客的IP获取距离最近并且在统计时间内的访问记录，然后根据访问记录得出热门标签，然后根据热门标签获取数据序号并加入待选池。根据待选池的数据以及每一条数据的热度，得出命中几率，然后根据随机数取得命中的序号并获取商品。

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
|  | GET bigrec/?act=reco | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| user\_no | varchar(32) | 空 | 用户在本应用中的唯一序号，如果为空，则进入冷启动程序 |
| size | int | >1<100，默认10 | 推荐总数 |
|  | 返回数据 | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| val | [{}] | | |
|  | pro\_no | string | 商品序号 |
|  | pro\_name | string | 商品名称 |
|  | pro\_cover | string | 商品封面图片地址 |
| info | choice | array | 采用了什么推荐方法，  user=用户热推，  play=游客热推 |
|  | choices | int | 推荐方法挑选进入待选池的数量 |
|  | static | int | 权重为9的固定值进入待选池的数量 |
|  | supplement | int | 热推方法不足，按权重进入待选池的总数 |
|  | elected | int | 最终进入待选池的待选记录总数 |

### 推荐点击事件上报

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=reco\_up  当用户通过本程序的推荐商品进行访问的时候，需要进行本接口的提交动作 | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| user\_no | varchar(32) | 空 | 用户在本应用中的唯一序号 |
| pro\_no | varchar(50) | 不为空 | 商品在本应用的序号 |
| pro\_tags | varchar(100) | 不为空 | 申报的时候必须将该商品的分类标签传入 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| JS插件用法 | | | |
| 执行函数bigrec\_recoup (user\_no，pro\_no，pro\_tags)  本函数为异步函数，可不用等待返回 | | | |

## 用户对象－统计信息获取接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_user | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| date\_start | date | 统计开始日期 | 最远时间不超过一年 |
| date\_end | date | 统计结束日期 | 与开始时间跨度不超过６０天 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| user | {total} | int | 总共用户数量 |
|  | {total\_sale} | int | 总共付费购买的用户数量 |
|  | total\_man | int | 所有用户中的男性个数 |
|  | total\_woman | int | 所有用户中的女性个数 |
|  | {total\_reco} | int | 通过分销商引入用户数量 |
|  | {total\_active} | int | ３０天内活跃用户总数  对用户活跃度的定义为有任何操作与上报记录（排除管理端操作记录） |
|  | {total\_active\_view}  [  {user\_no:1，  login\_time:1234}  ] | array | ３０天内活跃用户前１００条记录 |
|  | total\_active\_man | int | 活跃用户中的男性个数 |
|  | total\_active\_woman | int | 活跃用户中的女性个数 |
|  | {total\_reco\_view}  [  {分销商名称:int}，  ] | int | 各个分销商的引入统计表 |

## 用户对象－购买排行榜查询接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_user\_sale | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| p | int | 可空，默认１ | 查询第几页 |
| s | int | 可空，默认100 | 每页显示数据量 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| user | [{  user\_no:{  user\_sex:男，  user\_age:，  user\_phone:，  user\_sale:0.00，  user\_login:2017-06-13  }}，…] |  | 根据购买记录排序 |

## 商品对象－统计信息获取接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_pro | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| pro\_no | varchar(50) | 可空 | 商品在本应用的序号，逗号分割多个，最多允许１００个 |
| show\_type | info(详情统计)  play(内容统计)  playtime(内容时长)  share(分享统计)  sale(购买统计) | 不为空 |  |
| date\_start | date | 统计开始日期 | 最远时间不超过一年 |
| date\_end | date | 统计结束日期 | 与开始时间跨度不超过６０天 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| info | {total} | int | 所有符合条件的点击数 |
|  | [{pro\_no}{total}] | int | pro\_no商品的所有点击数 |
|  | [{pro\_no}{hour}  {[17-6-15-15]..}] | int | pro\_no的小时点击数统计 |
| play | {total} | int | 所有符合条件的内容查看数 |
|  | [{pro\_no}{total}] | int | pro\_no商品的所有内容查看数 |
|  | [{pro\_no}{hour}  {[17-6-15-15]..}] | int | pro\_no的小时内容查看数统计 |
| playtime | {total} | int | 所有符合条件的查看时长 |
|  | [{pro\_no}{total}] | int | pro\_no商品的查看时长统计 |
|  | [{pro\_no}{hour}  {[17-6-15-15]..}] | int | pro\_no的小时内容查看时长统计 |
| share | {total} | int | 所有符合条件的分享数 |
|  | [{pro\_no}{total}] | int | pro\_no商品的所有分享数量 |
|  | [{pro\_no}{hour}  {[17-6-15-15]..}] | int | pro\_no商品的小时分享数统计 |
| sale | {total} | int | 所有符合条件的购买数 |
|  | [{pro\_no}{total}] | int | pro\_no商品的所有购买数量 |
|  | [{pro\_no}{hour}  {[17-6-15-15]..}] | int | pro\_no商品的小时购买数统计 |

## 商品对象－销量排行榜查询接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_pro\_sale | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| p | int | 可空，默认１ | 查询第几页 |
| s | int | 可空，默认100 | 每页显示数据量 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| pro | [{  pro\_no:{  pro\_no，pro\_name，  pro\_cover，pro\_sale  }}，…] |  | 根据购买记录排序 |

## 商品对象－浏览排行榜查询接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_pro\_info | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| p | int | 可空，默认１ | 查询第几页 |
| s | int | 可空，默认100 | 每页显示数据量 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| pro | [{  pro\_no:{  pro\_no，pro\_name，  pro\_cover，pro\_info  }}，…] |  | 根据详情查看次数排序 |

## 商品对象－内容查看时长排行榜接口

|  |  |  |  |
| --- | --- | --- | --- |
| 参数 | 类型 | 默认值 | 说明 |
| GET bigrec/?act=stat\_pro\_playtime | | | |
| app\_id | int | 不为空 | 应用序号 |
| app\_key | varchar(50) | 不为空 | 应用提交密钥 |
| app\_password | varchar(50) | 不为空 | 应用管理密码，如果应用配置为不公开统计结果，则需要该密码方可查看数据 |
| p | int | 可空，默认１ | 查询第几页 |
| s | int | 可空，默认100 | 每页显示数据量 |
| 返回数据 | | | |
| status | bool | true | 接口状态 |
| error | string | 空 | 如果出错，则返回出错说明 |
| pro | [{  pro\_no:{  pro\_no，pro\_name，  pro\_cover，pro\_playtime  }}，…] |  | 根据内容查看总时长排序 |

# JS调用方法

首先引入JS文件：bigrec/js/?app\_id=xxx&app\_key=xxx&app\_password=xxx

默认情况下app\_password可为空，除非使用的接口必须该数据