



聚·变

第二届顺丰信息安全峰会分论坛

—— AI安全与及隐私保护(2) ——

图算法实践之设备评分

肖景芬

信息安全工程师

目 录

- 01 你为什么需要设备维度评分
- 02 唯品会的设备维度评分
- 03 图算法设备维度评分的实现
- 04 设备维度评分展望

01 你为什么需要设备评分

会员注册

已注册可直接登录

请输入手机号码

请输入验证码

获取验证码

密码由6-20位字母，数字和符号

请再次输入上面的密码

☐

我已阅读并接受以下条款：[《唯品会服务条款》](#) [《隐私条款》](#) [《唯品支付用户服务协议》](#) [《唯品信用服务协议》](#)

立即注册

账户登录

手机号/用户名/邮箱

密码

☒ 记住用户名

忘记密码

登录

 更多 

免费注册

betu

折扣

满1件打7.5折,满2件打7折,满3件打6.5折,封顶

¥10

领取

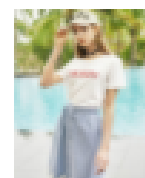
¥20

领取





🕒 商品将保留 19分50秒 ×



betu百图新款
撞色字母印花

1 ¥119

S

1件商品配送至上海市

¥119

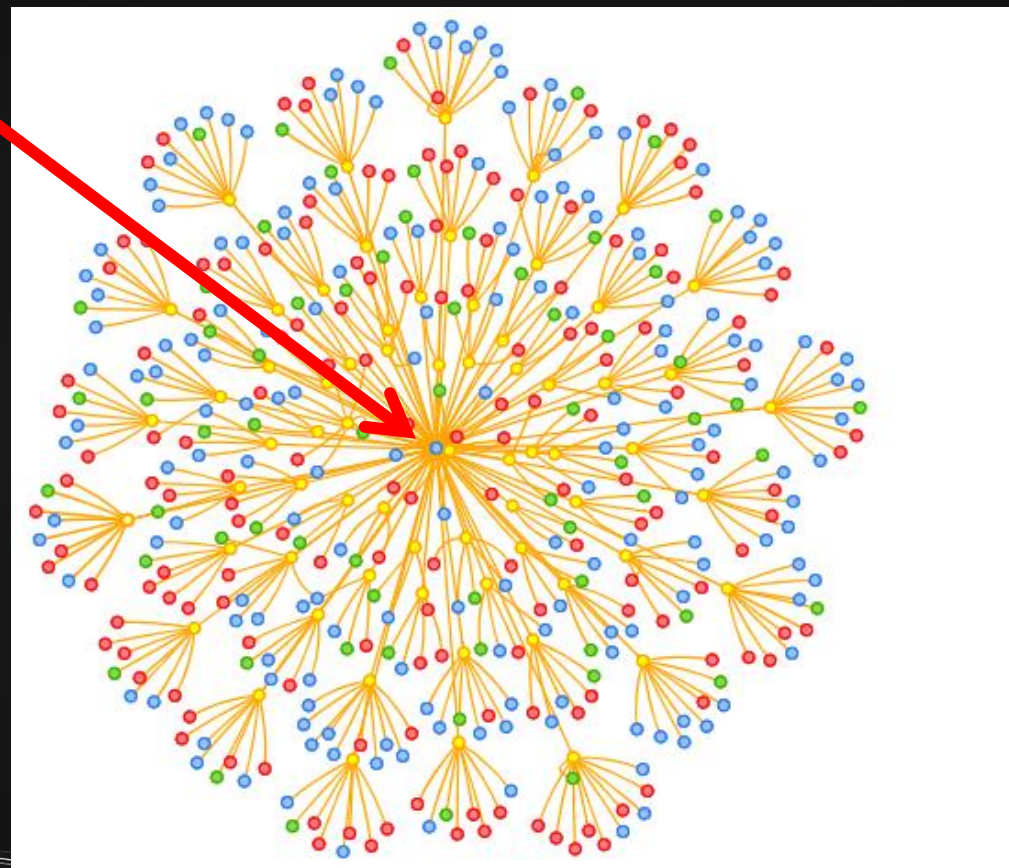
去购物袋结算



你怀疑过这个埋点吗？

log_time	mid	use	activity_name	activity_property	start_time	end_time
2018-06-15 23:45:52	nvs	(NU	active_te_push_connection_status	{"ispushcreate":"0","isconnected":	2018-06-15 21:02:40.119	(NULL)
2018-06-15 23:46:02	nvs	(NU	active_te_push_connection_status	{"ispushcreate":"0","isconnected":	2018-06-15 21:25:43.120	(NULL)
2018-06-15 23:30:51	nvs	189	(NULL)		2018-06-15 23:30:49.092	(NULL)
2018-06-15 23:45:08	nvs	151	active_te_interface_finished	{"name":"user_msg_type","msgtyp	2018-06-15 23:45:06.450	(NULL)
2018-06-15 23:39:24	nvs	(NU	active_te_push_connection_status	{"ispushcreate":"0","isconnected":	2018-06-15 23:39:22.713	(NULL)
2018-06-15 23:41:46	nvs	(NU	active_te_push_connection_status	{"ispushcreate":"0","isconnected":	2018-06-15 23:41:44.287	(NULL)
2018-06-15 23:36:38	nvs	189	active_pro_detail_sku	{"goods_id":"485108906","size_id"	2018-06-15 23:36:36.068	(NULL)
2018-06-15 23:39:40	nvs	135	active_te_video_download_result	{"res_type":"adv","res_id":"583993"	2018-06-15 23:39:17.385	(NULL)
2018-06-15 23:40:09	nvs	wei	active_te_snapped	{"type":"3"}	2018-06-15 23:40:05.288	(NULL)
2018-06-15 23:56:34	nvs	134	(NULL)		2018-06-15 23:56:32.199	(NULL)

▶ ▶
您查询的数据已经超过5000条，全部数据请下载后查看
当前显示 1 - 30 条记录 共 5000 条记录





你需要设备评分

02 唯品会的设备维度评分

- 特点一：分数制

评分分数表示设备的异常程度, 分数越大表示设备越异常

1. 分数制让用户可以很直观的看到设备有多异常
2. 业务方根据自己的业务需要，决定自己的业务场景去使用评分

- 特点二：无满分制

1. 百分制评分，值域为 $[0,100]$ ，区分度不够。

2. 无满分制评分，值域为 $[0,+\infty)$ 。

让评分的区分度更大，业务方可以对及其恶劣的设备进行某些特别的处理，
比如使用该设备的用户进行冻结

- 特点三：深入本质

1. 一设备多用户:

2. 挖掘“用户群”

提高黑产成本，评分更具解释性

03 图算法设备维度评分的实现



采集数据

imei

imsi

ccid

mac地址

cpu频率

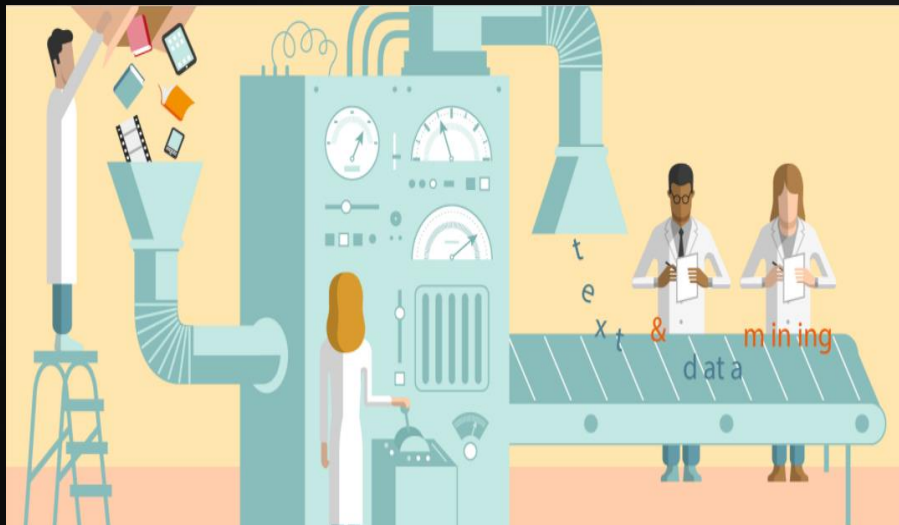
cpu型号

蓝牙地址

网络制式

屏幕分辨率


...



结合数据唯一性粗筛选数据

imei	----- > unique
imsi	----- > unique
ccid	----- > unique
mac地址	----- > unique
cpu频率	----- > not unique
cpu型号	----- > not unique
蓝牙地址	----- > unique
网络制式	----- > not unique
平衡陀螺仪	----- > not unique

...

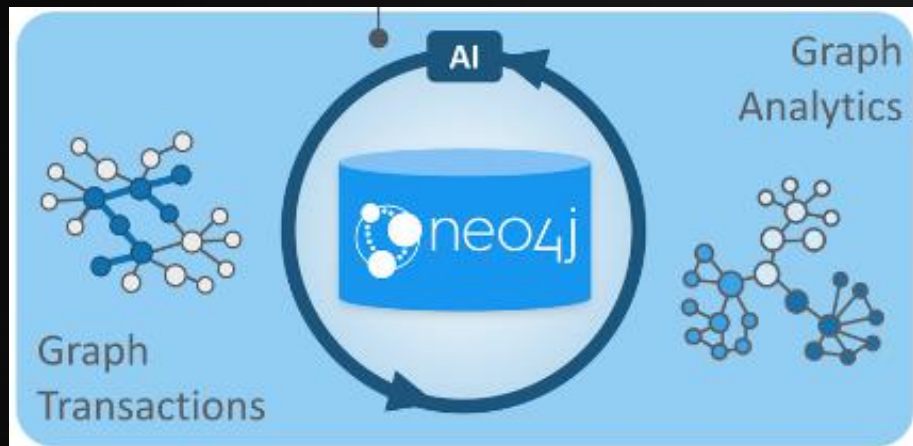


	I	J	K	L	M	N	O	P
1	kitchen	meal	pc	dog	nose	newspaper		
2	5	5	1	1	5	1	5	2
3	5	5	3	5	4	1	5	1
4	5	5	4	5	5	4	5	2
5	5	5	5	5	5	1	5	1
6	5	5	2	5	5	3		2
7	2	2	1	3	2	1		1
8	5	5	5		5	3		1
9	2	2	2	3	2	2		2
10	3	5	1	3	5	1		1
11	5	5	2	5	5	1		2
12	1	3	1	5	5	1		1
13	4	5	4	5	5	2		2
14	5	5	1	4	5	1		1
15	5	5	3		5	1	5	1
16	5	5	1		5	1	5	2
17	5	5	3	3	5	2	5	1
18	5	4	1	5	5	1	5	1
19	5	3	2	4	5	2	5	2
20	4	4	4		3	1	4	1

通常：维度中topN数据清洗

我们：借助已有的用户标签数据进行清洗

数据清洗



数据入图数据库neo4j

文档丰富

社区版免费

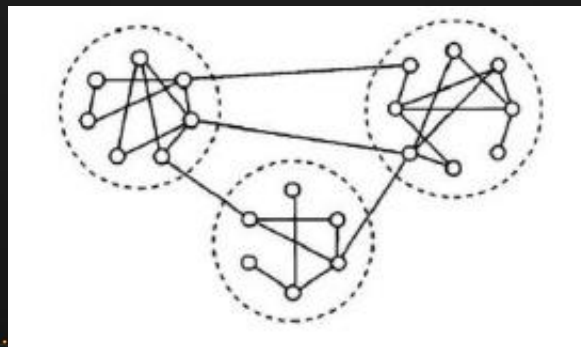
性能：

- 节点 300+亿
- 边 300+亿

设备评分

1. 分群：社区发现算法

A community is a subgraph containing nodes which are more densely linked to each other than to the rest of the graph or equivalently, a graph has a community structure if the number of links into any subgraph is higher than the number of links between those subgraphs. [Newman, et. al. 2004]



设备评分

2. 计算群内数据统计特征

① 度数统计特征: 单节点入度,出度/群内节点入度,出度的均值,方差,偏度,峰度

① 节点统计特征: 总节点数,属性节点占比

3. 根据统计特征计算评分

04 设备维度评分展望

新增其他节点

1. 引入uid相关属性节点: phone, bank_card, ID_card and so on
2. 引入业务场景属性节点: register, login, coupon, order and so on
3. 引入环境属性节点: ip , network_type and so on

提供其他功能

1. 挖掘黑产团伙
2. 补充用户维度评分未检测出的异常
3. 检测出孤立ip群
4. 提供图特征优化现有机器学习模型,尝试图卷积算法(GCN)



THANK YOU