SNG
社交网络事业群

基于社交网络的大规模网络攻击自动对抗技术

成杰峰
geoffcheng@tencent.com
腾讯社交网络BG

QCon 全球软件开发大会
INTERNATIONAL SOFTWARE DEVELOPMENT CONFERENCE
BEIJING 2017

大纲和自我介绍

SNG
社交网络事业群

1 互联网安全第一环：恶意帐号

3 腾讯恶意帐号智能识别系统

3 全量社交网络规模的图计算内幕

4 从恶意帐号识别到大规模网络攻击的自动对抗

5 恶意库在互联网安全的应用

QCon 全球软件开发大会

主办方 Geekbang. InfoQ

我是“网络信息安全建设最佳实践”专题讲师

腾讯社交网络运营部专家工程师

成杰峰

安全第一环：恶意帐号

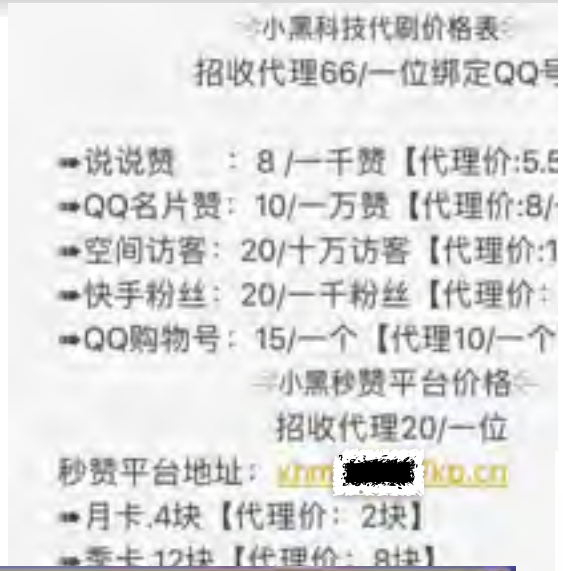
SNG 社交网络事业群

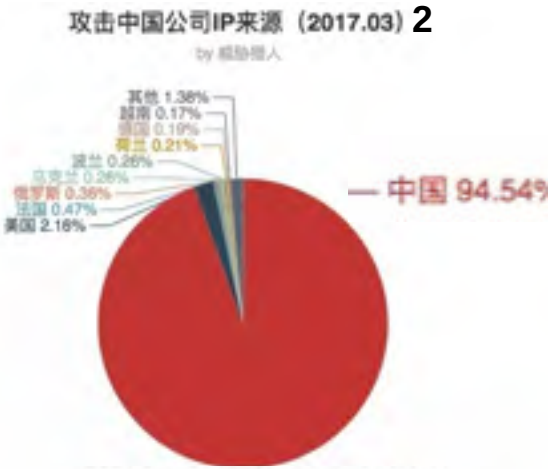
51.8%的全网流量来自于自动机，28.9%为大规模攻击产生的恶意网络流量[1]

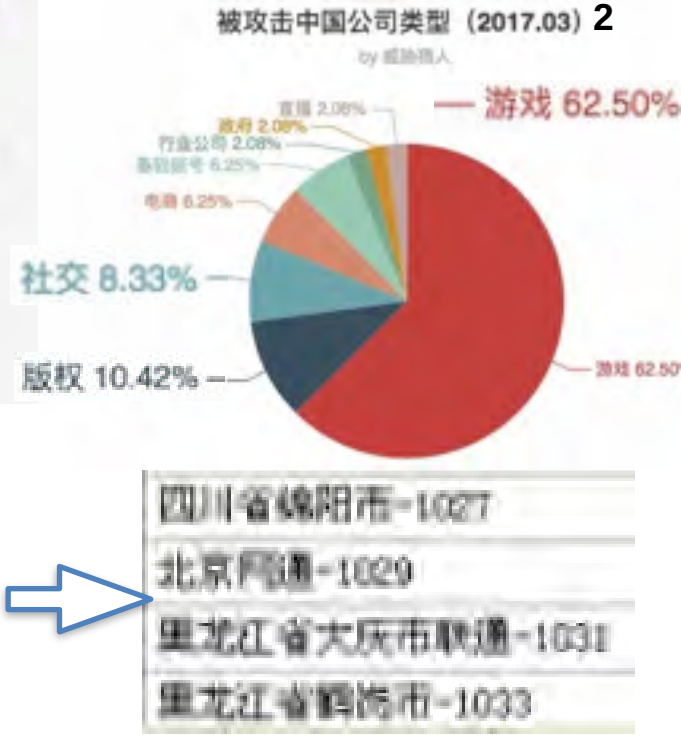
- 中国是最大的攻击目标/发源国
- 攻击主要受影响的行业：电游,娱乐和社交
- 通过帐号发起的恶意威胁互联网信息安全
- 养号和盗号[fake(sybil)/compromised account]

针对账号的安全制本之道

- 恶意帐号的操控人五花八门
- 恶意帐号的目的各异
- 养号行为无规则和统计特性
- 自动分类无法在大盘准确识别







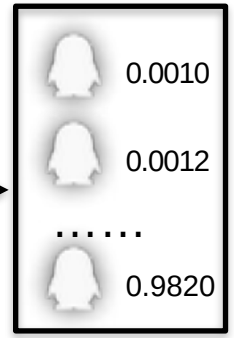
1 Bot Traffic Report 2016

2 http://www.cebnet.com.cn/20170407/102380792.html

腾讯恶意帐号智能识别和安全服务

全量社交关系链

高可扩展性
SybilRank算法



社交账号SybilRank打分

按登录异常条件
筛选高排名帐号

静态养号

请求流水

静态养号做种在海量请求流水数据中聚类扩展计算大盘中的活跃养号

大批活跃养号团伙

机器学习分类模型自动分类识别

恶意、可疑和温和的团伙类别

按业务逻辑定制
打击和监控策略

基于团伙标签的安全服务策略

腾讯云

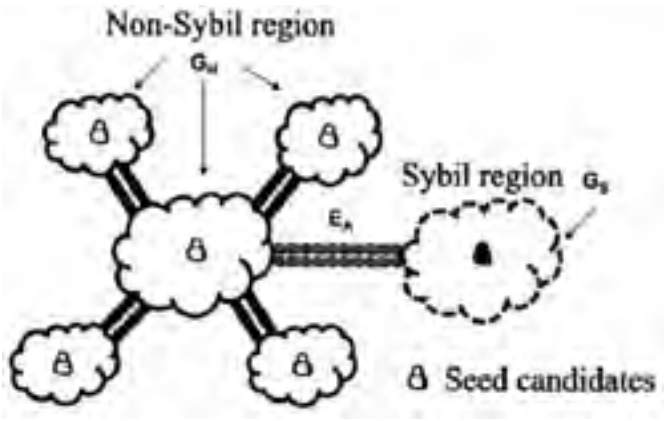
祝融安全系统

社交网络大数据挑战

挖掘静态养号账户：基于网络SybilRank打分[1]的图挖掘方案：I-图聚类；II-算打分；III-选养号

- 关系链图有18亿个顶点,495亿条边，目前最先进的可扩展图聚类pSCAN[2]也仅能处理4千万个顶点，12亿条边的大图
- SybilRank使用的是基于社区模度的图聚类方法[3]，其最好扩展版本[4]能处理1亿个顶点，37亿条边

高可扩展性图处理



1 Xiaowei Yang, and Tiago Pregueiro: Aiding the detection of fake accounts in large scale social online services, In NSDI, 2012.

2 Lijun Chang et al.: pSCAN: Fast and exact structural graph clustering, In ICDE, 2016.

3 Vincent D. Blondel et al.: Fast unfolding of communities in large networks, In J. Stat. Mech., 2008.

4 Hiroaki Shiokawa et al.: Fast algorithm for modularity-based graph clustering, In AAAI., 2013.

高可扩展性处理基础：图计算

图数据之间彼此互连，依赖性强；平台专业性强

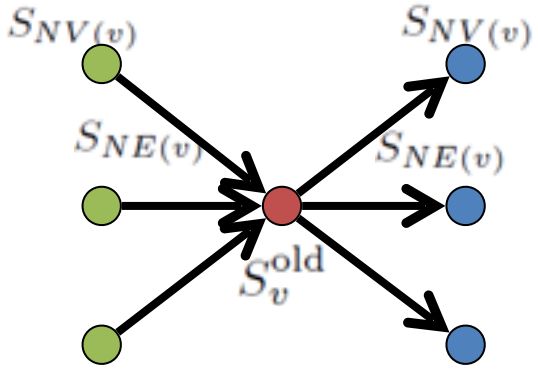
- MapReduce和Spark等主流大数据平台并不适合
- 图计算平台成为了图挖掘和相关机器学习算法的新方法
- 用顶点程序 (equiv. Map/Reduce)达到自动化分布并行平台独立
- 分布式内存图计算 vs 磁盘图计算
  - **Jiefeng Cheng**, Qin Liu, Zhenguo Li, et al. VENUS: Vertex-Centric Streamlined Graph Computation on a Single PC. In ICDE, 2015.

Pregel GraphLab

PROJECT PEGASUS

顶点程序

$$S_v^{new} = \text{VertexUpdate}(S_v^{old}, S_{NV(v)}, S_{NE(v)}).$$



分布式 vs 磁盘

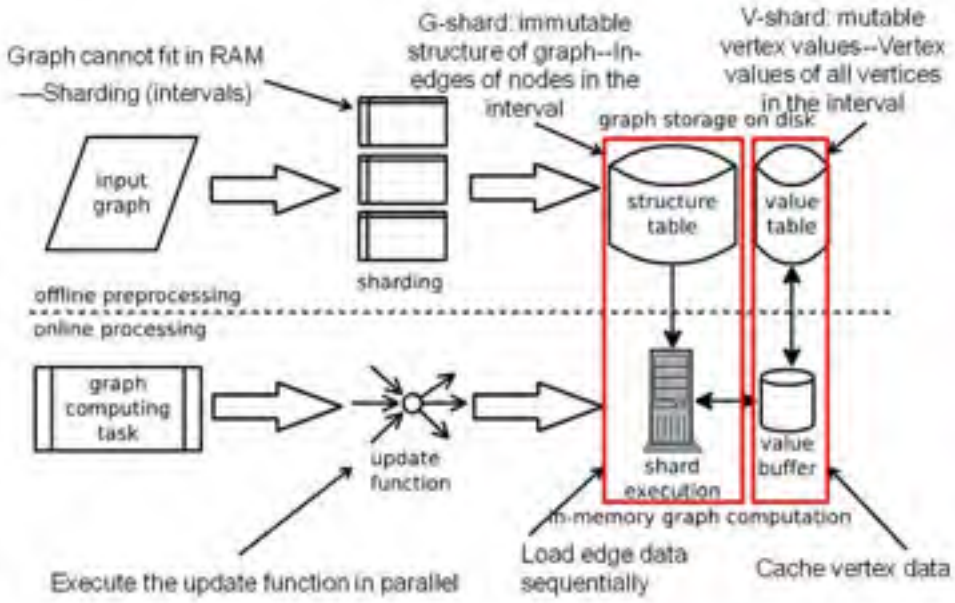
VENUS vs Spark: PageRank on Twitter (41M 顶点，1.4B 边)

system	resource	speed
Spark	50 servers, 100 CPU	8.1minutes
VENUS	1 PC, 8CPU, hard disk	5minutes

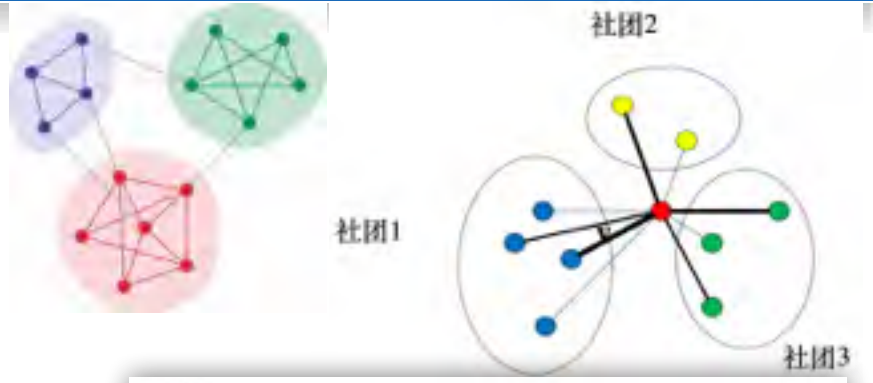
VENUS vs GraphChi: PageRank on clue-web (1B 顶点，43B 边)

GraphChi	X-Stream	VENUS
4.3hours	7.4hours	1.8hours

磁盘图计算系统VENUS



高可扩展性图聚类



Fast unfolding算法的图计算实现

- 以模块度最大化的目标把社交网络分成多个社区
- 模块度 = 各社区内总边数 – 各社区内随机应有的边数

$$Q = \frac{1}{2m}\sum_{i,j}[A_{ij} - \frac{k_i k_j}{2m}]\delta(c_i, c_j) \qquad \delta(c_i, c_j) = \begin{cases} 1 & when\ c_i == c_j \\ 0 & else \end{cases}$$

所有边的权重之和

顶点i和j之间边的权重

与顶点j相连的边的权重之和

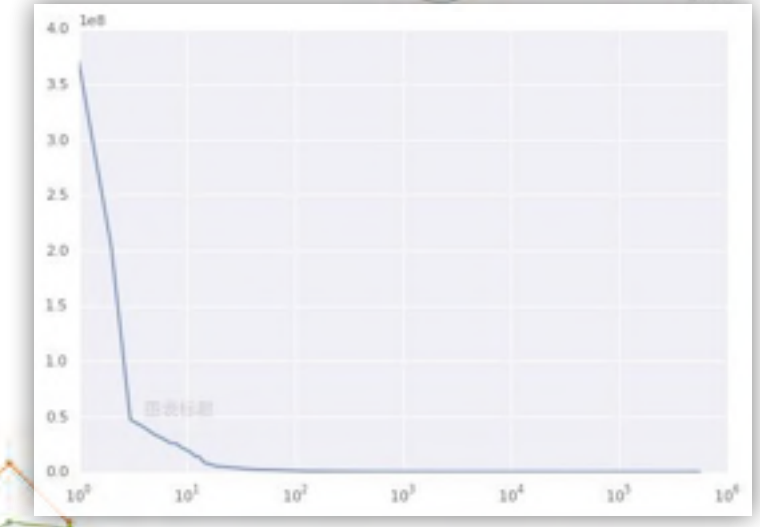
- (i)初始化(ii)按模块度增益调整每个顶点v至相邻社区，重复直到所有顶点的所属社区不再变化(iii)压缩同一个社区所有顶点成一个顶点并重复前一步直到模块度收敛

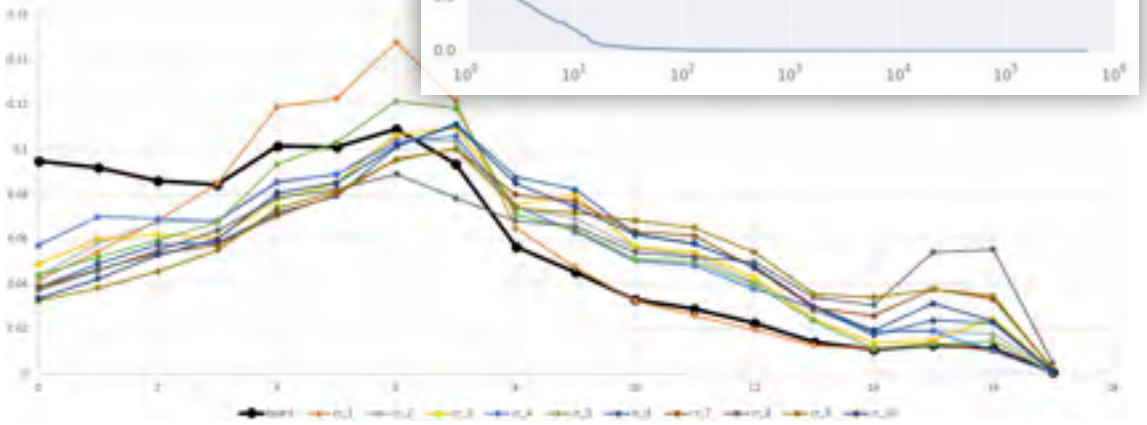
模块度增益

$$\Delta Q = \frac{k_v[C]}{2m} - \frac{k_v \sum_{tot}^{C}}{2m^2}$$

社区C与顶点v相连的边的权重之和

社区C的边的权重之和





- 图计算中的顶点状态是每个顶点所属社区，维护一个全局字典$\sum_{tot}^{C}$
- 顶点程序收集$k_v[C]$，并对每个相邻社区计算模块度增益
- 耗时119小时，3.1千万个社区，前3大社区大小分别为3.7亿、2亿、4.7千万，两大社区30%的顶点，2万个连通分量，最大18.8亿，3.9万微连通分量孤立号码
- 社区大小和对应的社区满足power-law分布

高可扩展性SybilRank计算

对图聚类发现的每个社区选出一个白种子（好人）
- 基于业务数据：投诉、异常登录、打击历史等

SybilRank 算法图计算实现
- 图计算中的顶点状态是每个顶点的SybilRank打分

$$T^{(0)}(v)=\begin{cases}\frac{T_G}{S} & \text{v是白种子}\\ 0 & \text{v不是白种子}\end{cases}$$

$$T^{(i)}(v)=\sum_{(u,v)\in E} w_{(u,v)}\frac{T^{(i-1)}(u)}{k_u}$$

给定图中分值的总和

图中白种子的个数

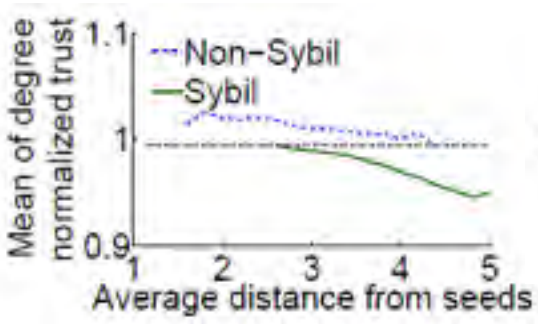
第i-1次迭代中u的打分

u的邻边的权重之和

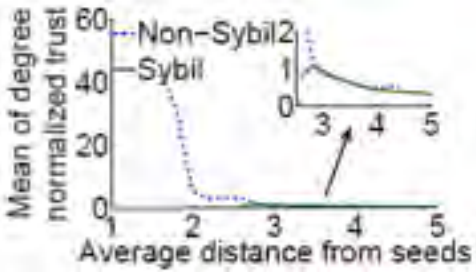
v和u间的边的权

- (i)初始化(ii)根据顶点状态和边的权值更新状态(iii)计算$O(log(|V|))$次迭代
- 需时62.65个小时

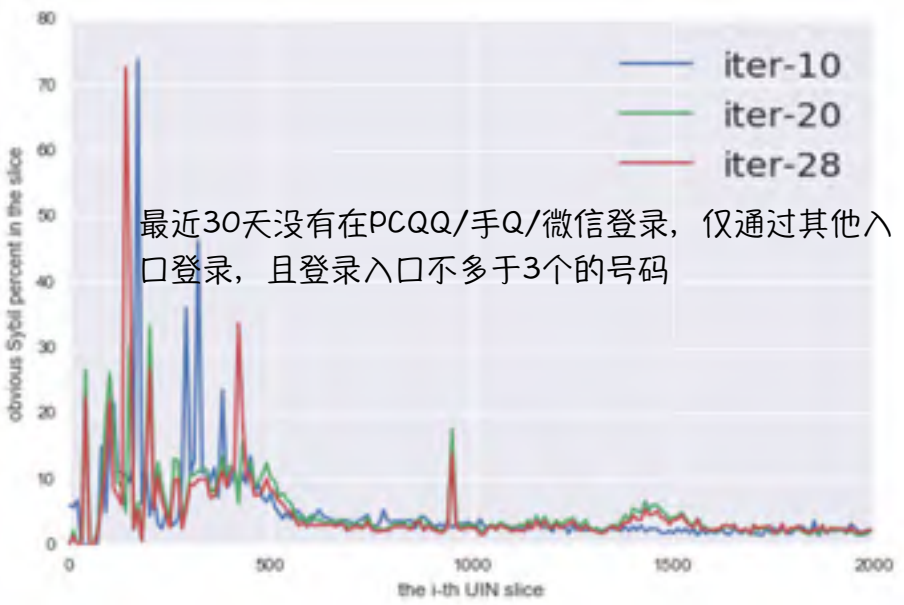
理想



实际



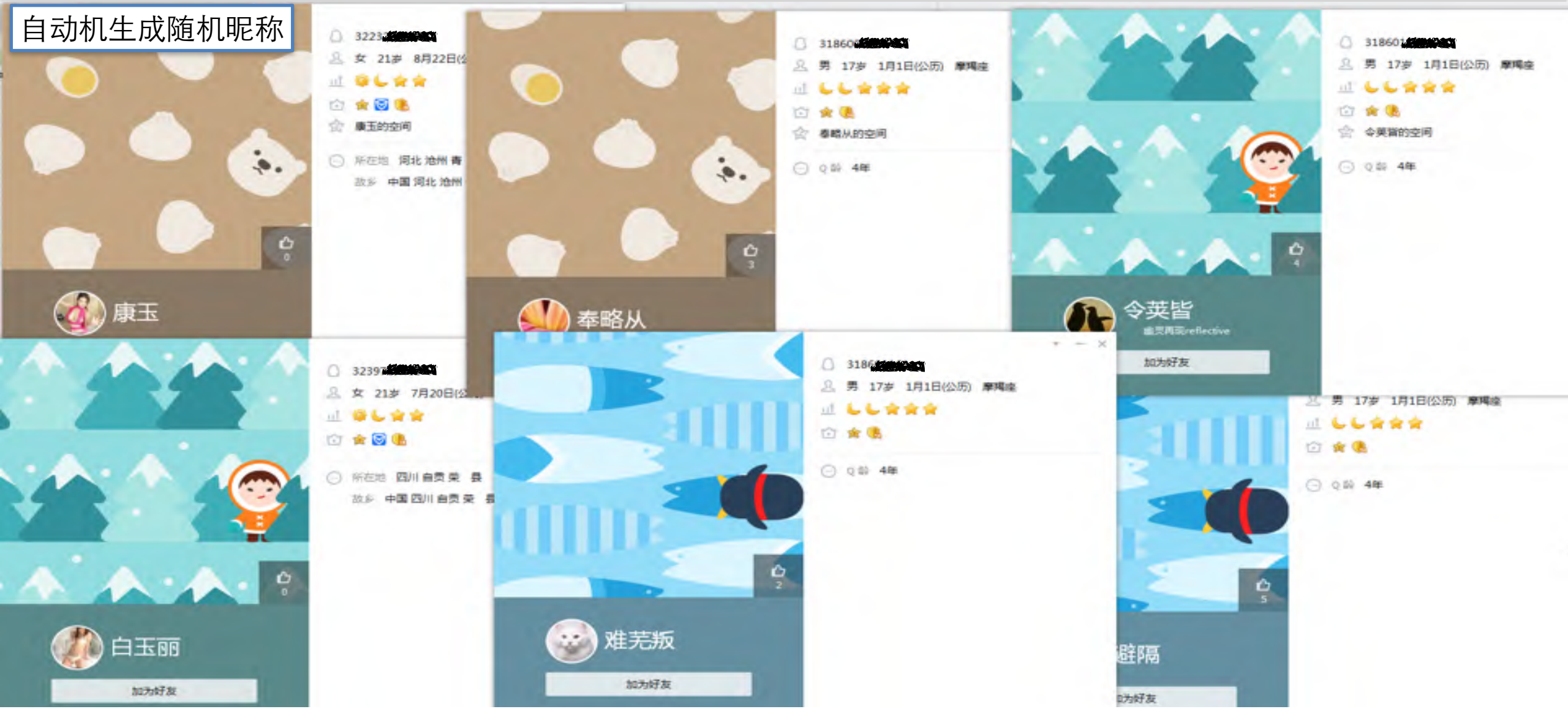
用打分排名划分出的高可信养号区
- 分析排名靠前的号码，使用登录入口行为异常，设备聚集等几个条件辅助判断定出高可信养号区，号码量在1千万左右
- 静态养号：针对在线活跃攻击的覆盖率很低



静态养号典型案例

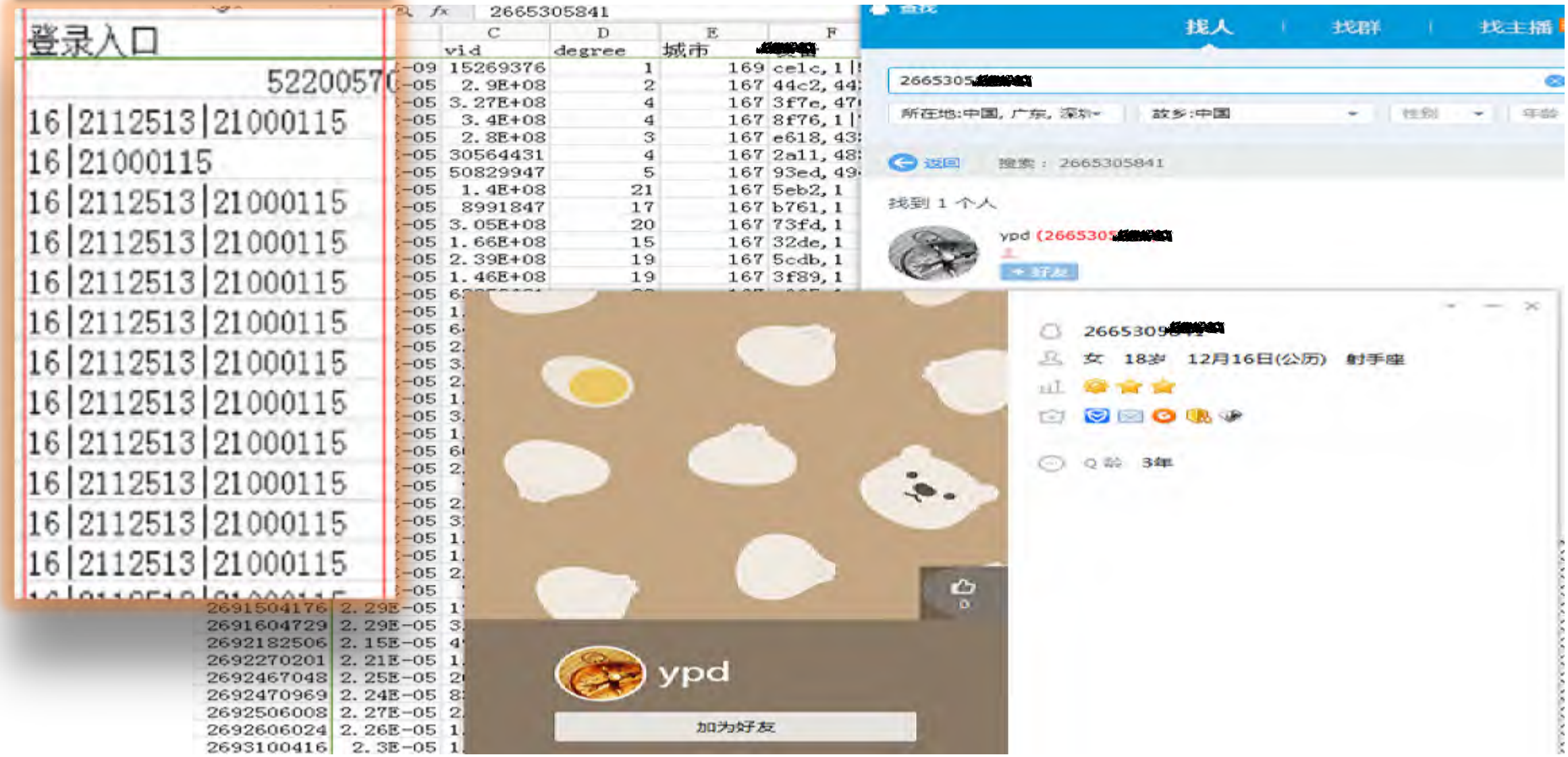
SNG
社交网络事业群

自动机生成随机昵称



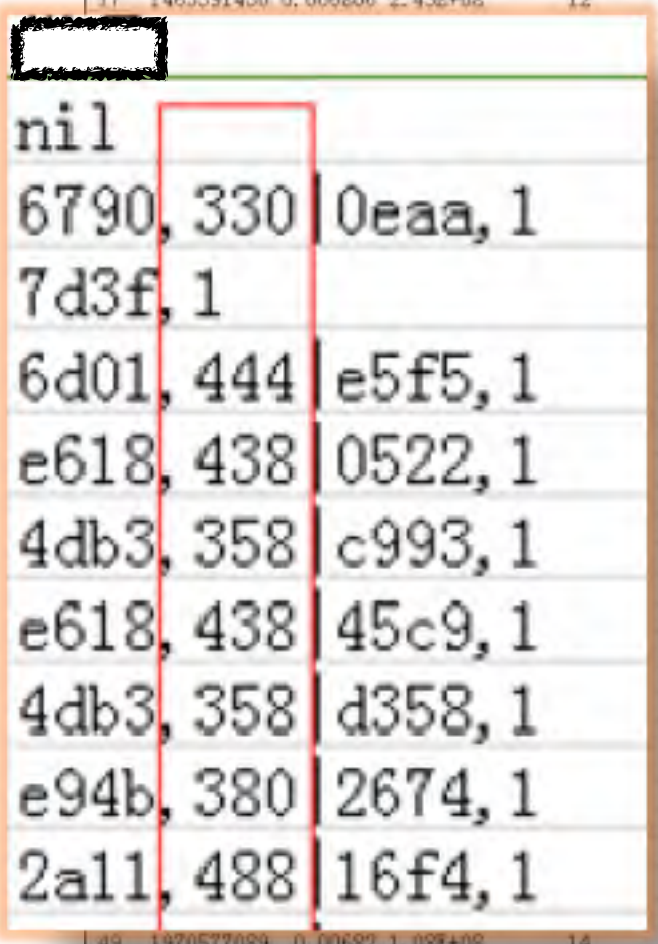
静态养号典型案例

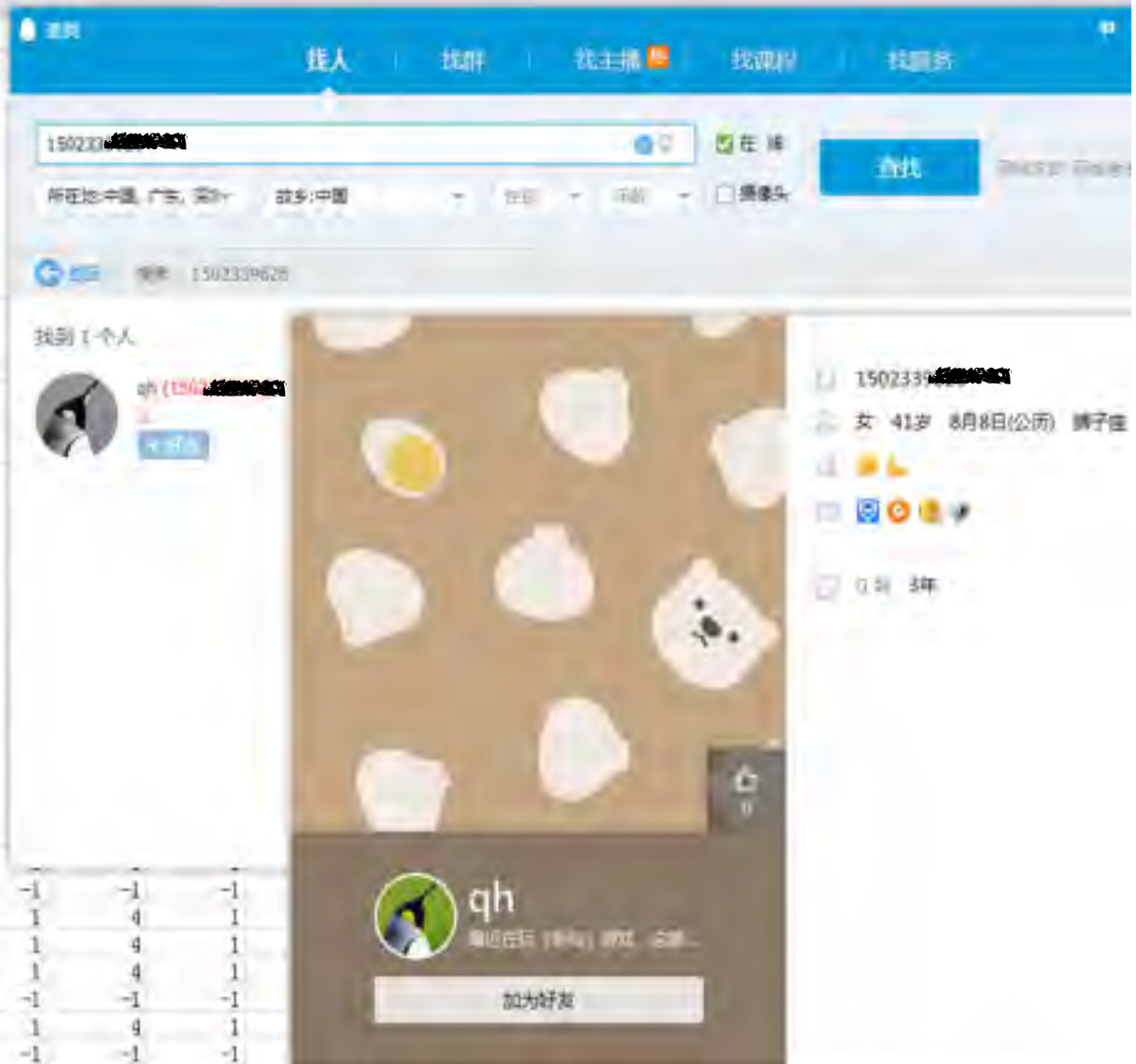
非常用登录入口上有聚集



静态养号典型案例

登录IP上有聚集





静态养号典型案例

大量不同虚假公司名

uin	score	vid	degree	城市	[redacted]	登录入口	hashes
227331[redacted]	0.41774	52361367	33	366|367	3f2d,5	16	11,12,13,14,
10574[redacted]	.417713	198522283	603	358	8a77,3|00b5,2|be9c,3	1|16	11,12,13,41,
27212[redacted]	.417739	297731159	144	354	1007,5	1	11,12,13,41,
12198[redacted]	.417712	161454275	607	353	96c4,2|5b88,3	1|16	11,12,13,41,
28485[redacted]	.417733	164152638	311	353	f4f2,4	1	12,13,41,111
3769[redacted]	.417747	97666292	446	353	76d4,10	1	11,12,13,41,
7050[redacted]	.417725	46638621	487	353	d8dc,6	16	11,12,13,41,

静态养号典型案例

利用已经失效的公司的号码

静态养号典型案例

不再活跃的商业号码

从静态养号到活跃养号团伙

SNG 社交网络事业群

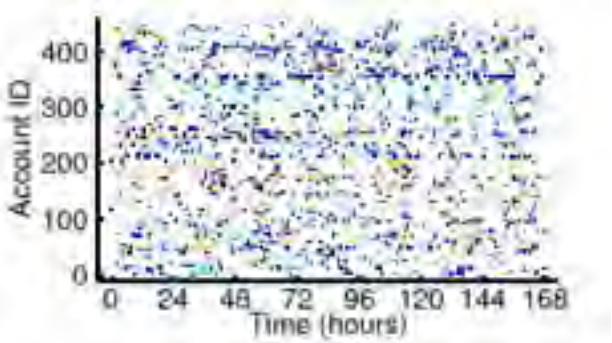
恶意批量行为和养号团伙识别
- 批量协同的大规模恶意通常具有爆发和分布特征
- 具有反复批量相似行为的养号群组形成团伙
- 帐号两两相似度计算：密集／反复性请求产生的噪音和有效时间窗产生的重复

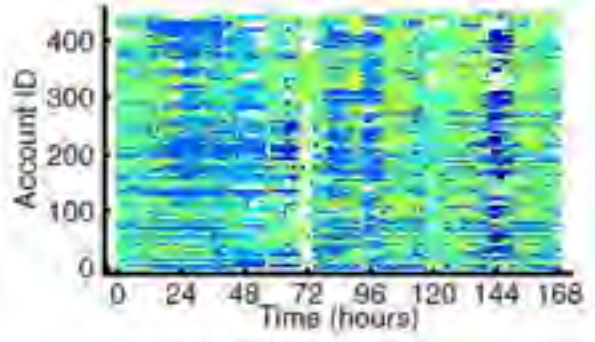
静态养号做种在海量请求流水数据中聚类
- 可计算性:

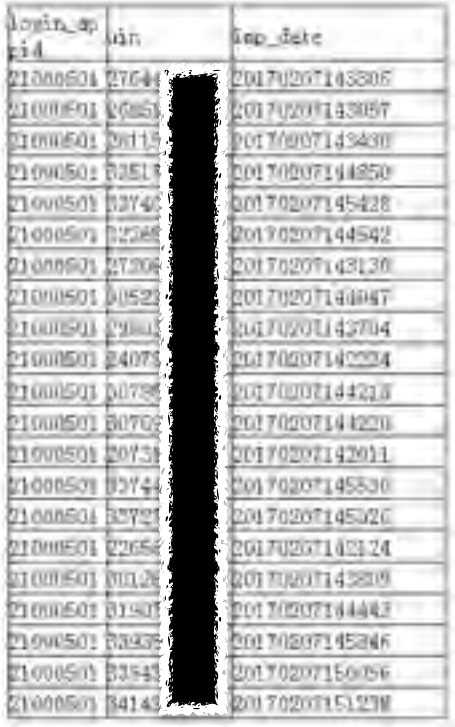
$$O(|V|^2) \rightarrow O(|V|)$$

- 健壮性：仅养号容易具有长时间反复批量相似行为(含盗养号)
- 可运营性：一天业务总请求量～500亿，用1个入口5天的登录请求(~2亿)可匹配出700亿对相似养号用户

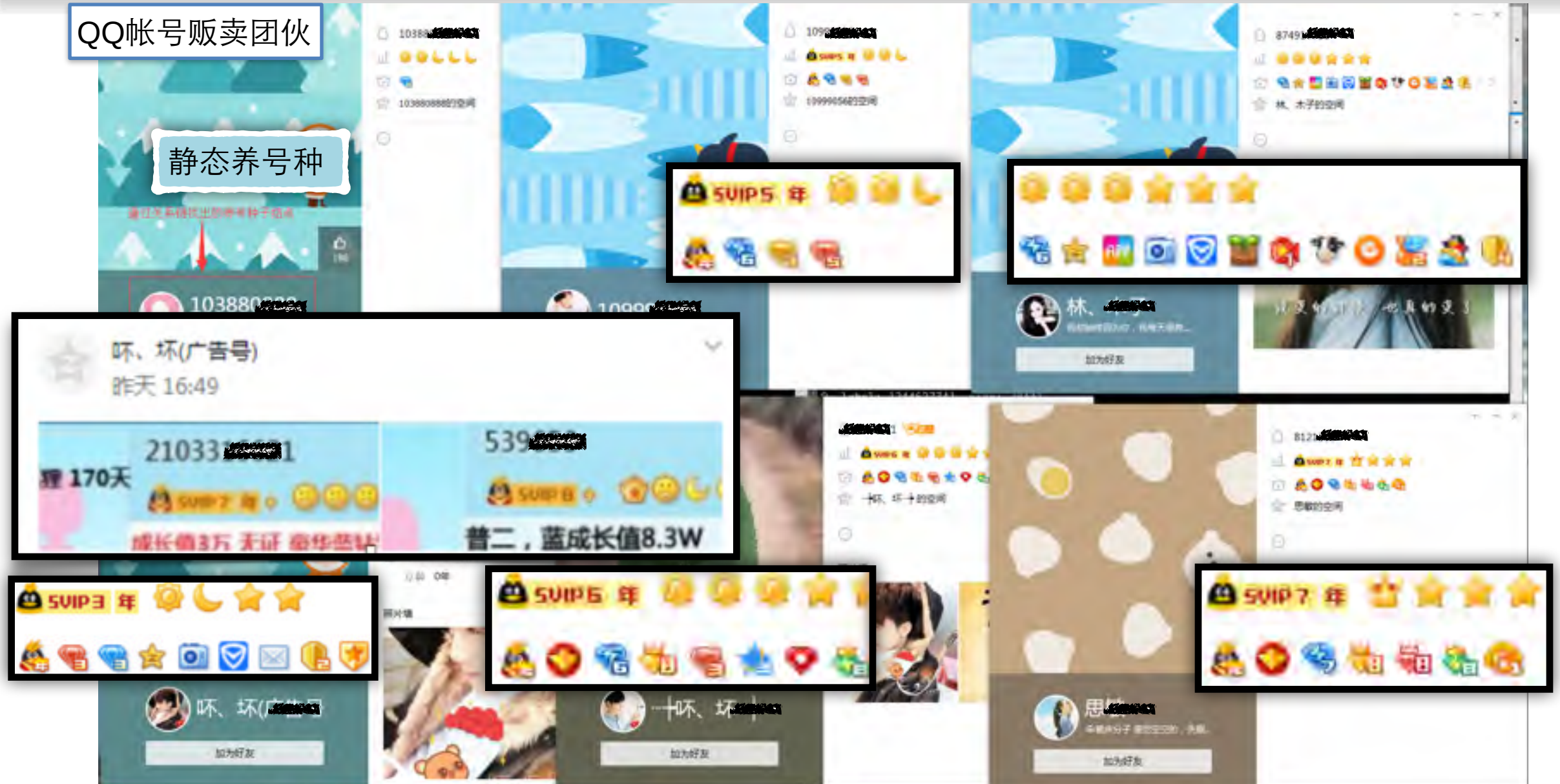
基于登录请求可自动识别～2000万的活跃养号
- 团伙大小大致分布：>1000 ~40%
  >500 ~30%
  其余 ~30%







活跃养号团伙典型案例


QQ帐号贩卖团伙
静态养号种
SVIP5 年
SVIP3 年
SVIP6 年
SVIP7 年
普二，蓝成长值8.3W
加为好友

活跃养号团伙典型案例

玲玲＋数字 只在同一IP登录，无其它请求

每个团伙通常有真人boss

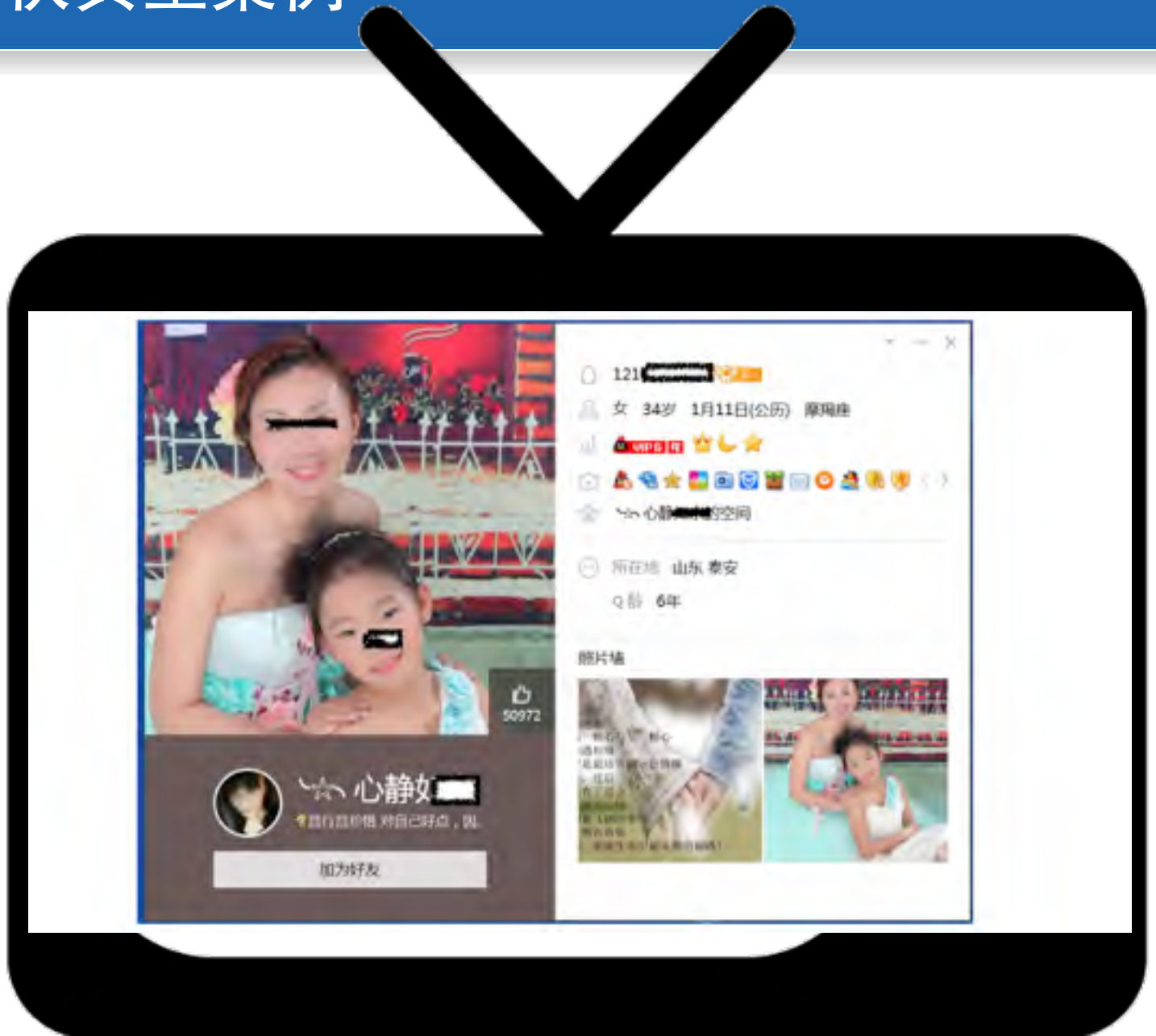
包含盗养号（标黄），有加群、群消息、好友消息产生，但是在该登录入口的行为高度匹配

活跃养号团伙典型案例

一群会说故事的nice养号团伙?

活跃养号团伙典型案例

盗养号团伙

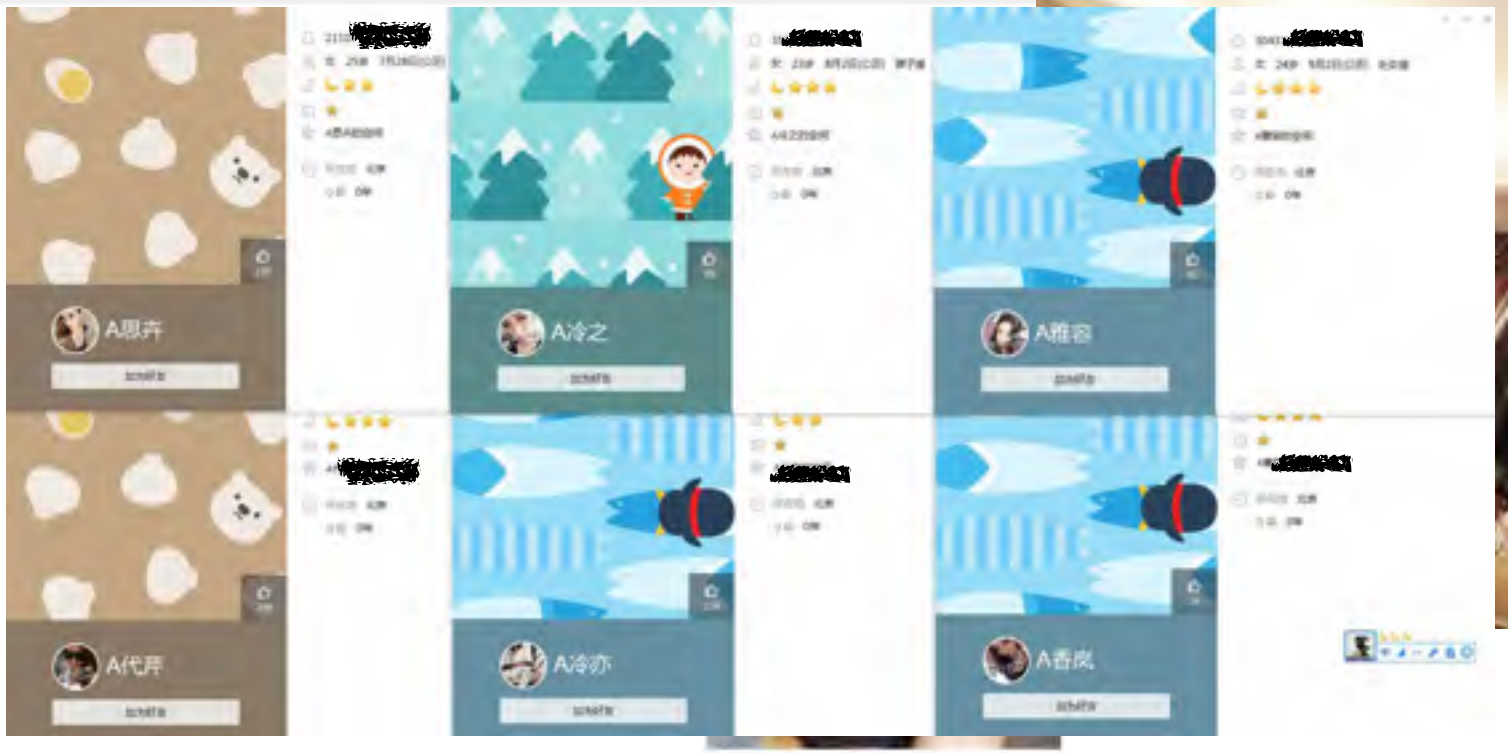


基于恶意帐号识别的恶意自动对抗

SNG 社交网络事业群

从养号作恶分类到团伙标签

- 柔性治理：好人偶然违例 vs 坏人严格处理
- 养号作恶种类
- 安全中机器学习难题的解决方案
  - ☻业务中的发现的养号：样本自动提取
  - ☻OCR图像识别难题的化解
  - ☻恶意URL识别难题的化解



基于团伙类别的各类安全策略

按业务逻辑定制打击和监控策略

恶意、可疑和温和的团伙类别

机器学习分类模型自动分类

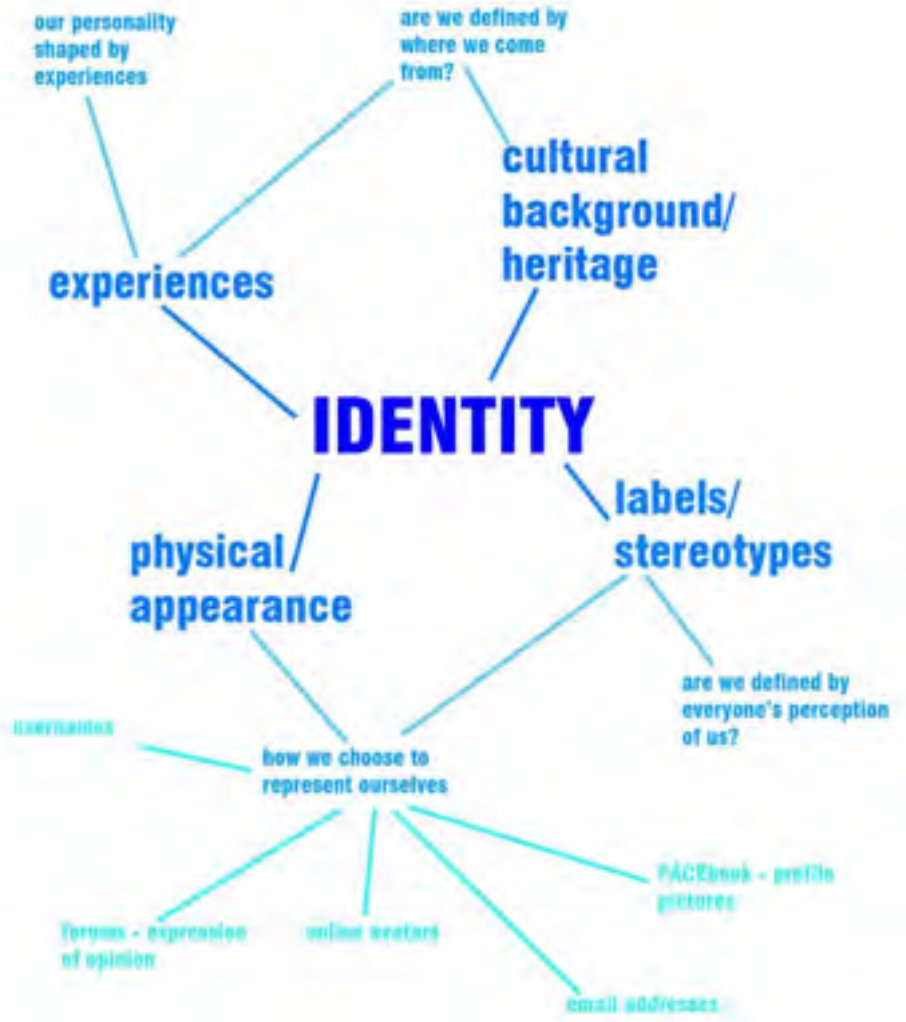
大批活跃养号团伙

基于恶意库的互联网安全应用

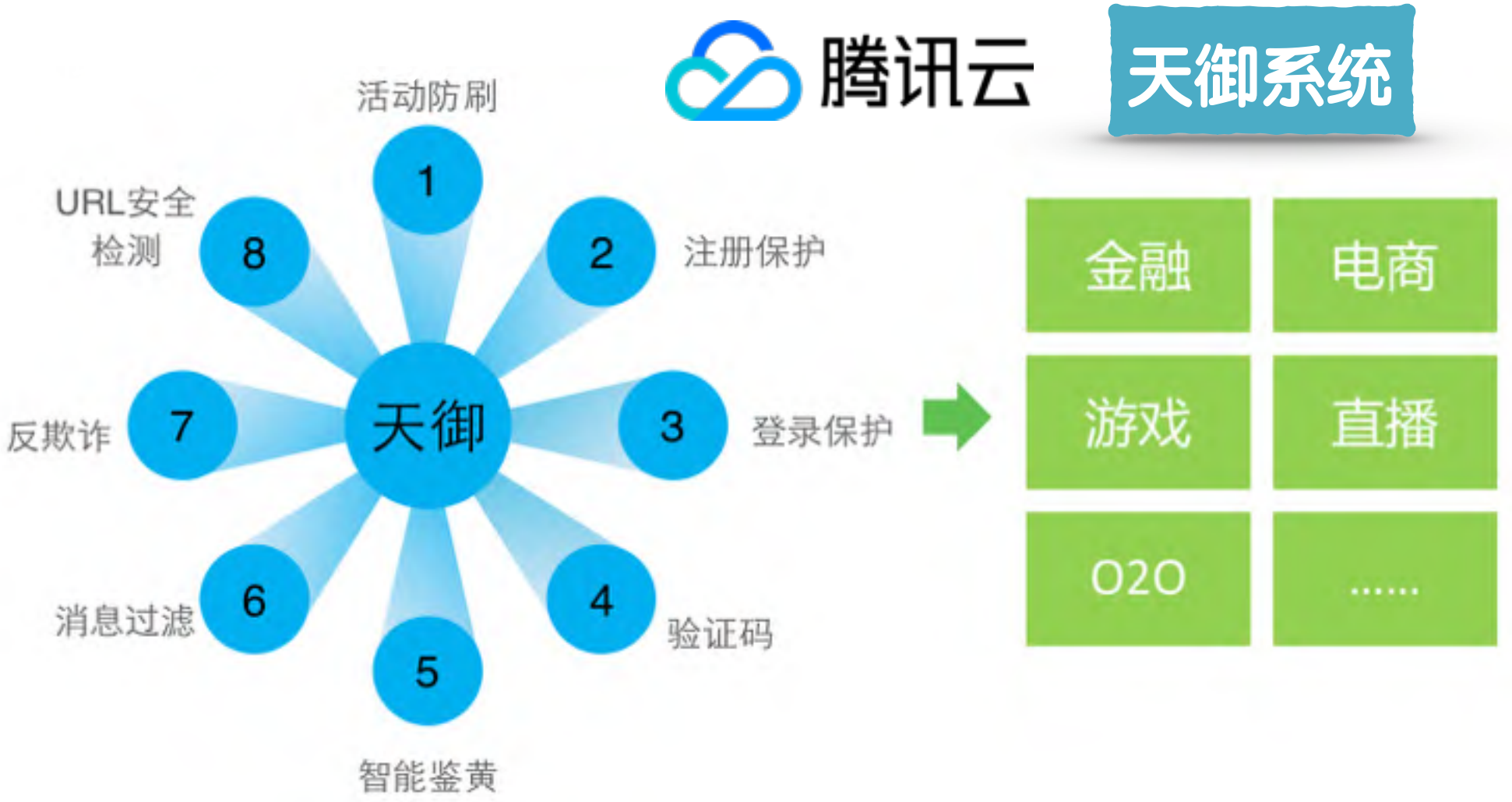
SNG
社交网络事业群

互联网金融安全：反欺诈

- 申请人帐号、手机号、设备、IP地址都是图中的结点
- 图中的边是发生的联系
- 组织性精密的金融欺诈团伙：多个实体、多组行为
- 知识图谱可用来更好地查询复杂的关联信息



腾讯云反欺诈


腾讯云
天御系统
活动防刷
1
URL安全
检测
8
2
注册保护
天御
反欺诈
7
3
登录保护
6
消息过滤
4
验证码
5
智能鉴黄
金融
电商
游戏
直播
O2O
.....

END