



图谱可视化与关联分析实践

赵丹

2016-09

海云数据

HYDATA



HYDATA
海云数据
— 大数据行业运营商 —

大数据的行业运营商 中国大数据可视分析的领导者

海云数据专注于利用大数据分析 with 计算机视觉技术
为用户运营数据、分析业务，帮助用户更加快速灵活的应用大数据。

"HYDATA is a team full of passion.
I like the creativity on treating data"



Microsoft—Bill Gates

User evaluation

"希望你们继续努力，在发展经济的数据战场上继续当好尖兵和突击队。"



国家总理
李克强



中航信党委书记
崔志雄

"数据可视化是大数据的核心部分，海云数据将在航空领域大有可为。"



北京市金融局局长
霍学文

"海云数据在大数据可视化的领域的技术让人惊喜，值得学习。"

HYDATA

在世界前500强的企业中
已经有56家企业购买了我们的产品与服务



我们的愿景

帮助用户更加快速、灵活地运营大数据



利用数据可视化思维



探索行业大数据价值

问题

从数据中寻找可能的关联（两两组合，回归分析）

从大量关联中发现模式（目视探索、交互筛选）

图可视化主要用在第二种分析，从大量关联中发现模式；偶尔也可用在第一种，从一种关联模式发现另一种关联模式（举例：从词共现模式中发现知识）

可视分析中图的扩展

节点和关联上加标签

节点颜色和图标，关联颜色、线型、线宽度

增加空间背景

以特殊形式自动排列



图本身是一种空间模式

图所表达的实体所处的空间形成空间背景

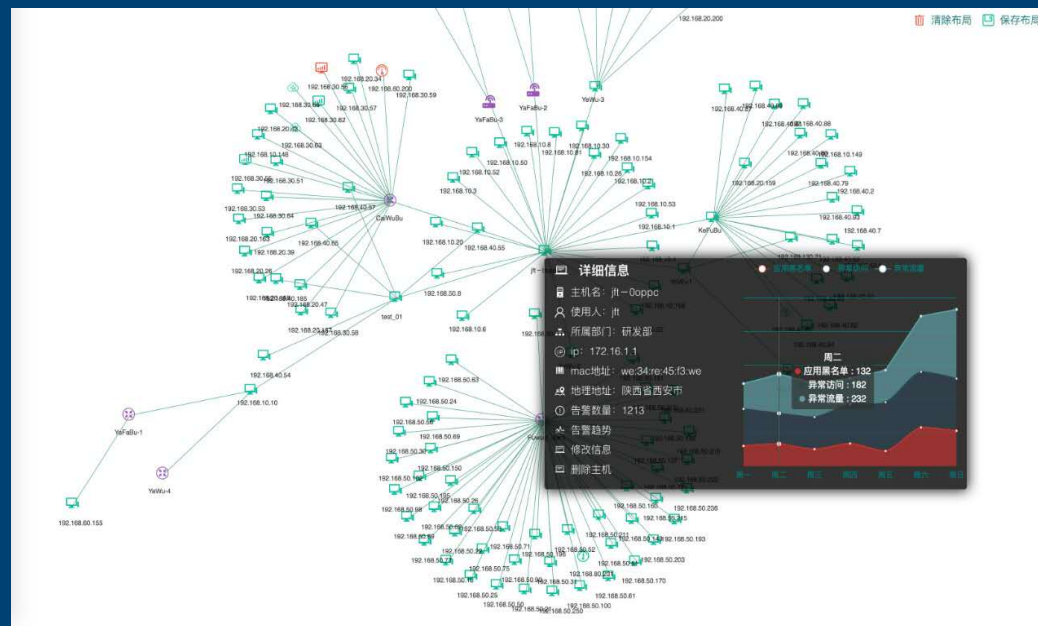
- a. 地理空间背景
- b. 人际关系空间背景
- c. 网络空间背景
- d. 其他专业空间背景，特定科研领域

图的空间背景

地理空间背景

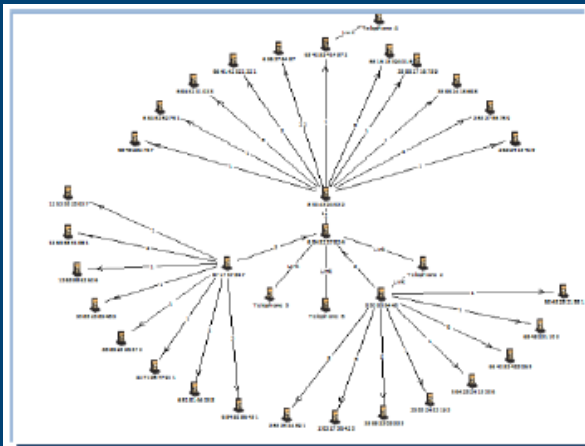


网络空间背景

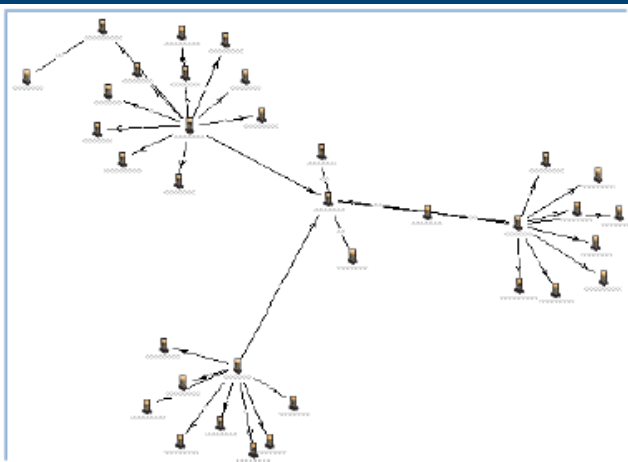


排列方式

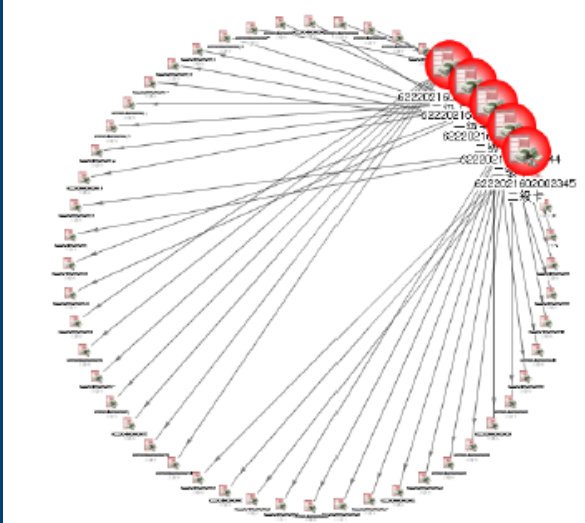
分组布局



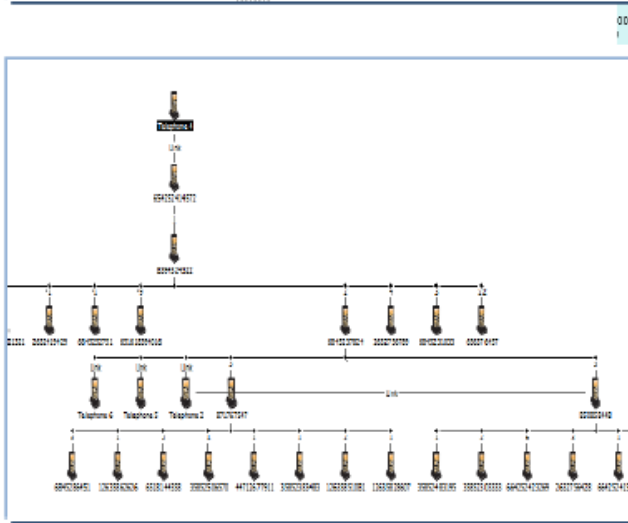
网络布局



圆形布局



层次布局



系统漏洞分析

公安技战法

- a. 群集：犯罪团伙结构
- b. 链路：关系人
- c. 空间：位置关系
- d. 时序：事件发展

定义关联（以公安行业为例）

 **同行入住分析**
不同房间号同一宾馆，入住时间前后5分钟。

最新任务: 2013-09-25 16:16:29 完成 9/9

[查看](#)

 **入住同天同房号**
入住同一天同一房号

最新任务: 2013-09-07 23:51:25 完成 5/5

[查看](#)

 **国内老乡同住**
同一天入住同一酒店的老乡。

最新任务: 2013-09-07 20:22:19 完成 2/2

[查看](#)

 **进港同订票号分析**
进港同订票号分析

未开启任务 完成 0/0

[查看](#)

 **进港同订票且老乡**
进港同订票且老乡

最新任务: 2013-09-27 09:48:51 故障 0/2

[查看](#)

 **离港同订票号分析**
离港同订票号分析

最新任务: 2013-09-07 20:22:50 完成 1/1

[查看](#)

 **离港同订票且同乡**
离港同订票且同乡

最新任务: 2013-09-07 20:23:03 完成 1/1

[查看](#)

 **进港后1天就离港**
进港后1天就离港

最新任务: 2013-09-07 20:23:15 完成 1/1

[查看](#)

 **旅客同行航班**
旅客同行航班

最新任务: 2013-09-07 20:23:27 故障 2/2

[查看](#)

 **国内同行分析**
国内旅客同时入同一房间或者同一宾馆开房时间前后10分钟。

最新任务: 2013-09-07 23:51:40 完成 4/4

[查看](#)

 **民航同飞分析(同航班)**
找出目标人物出现3上次以上乘坐同一天同一航班的同乘旅客。...

最新任务: 2013-09-16 12:40:14 完成 11/11

[查看](#)

 **民航同飞分析(同目的地)**
根据目标人物，找出48小时内1次以上飞往同一目标地的旅客；...

最新任务: 2013-09-16 12:37:52 完成 10/10

[查看](#)

可视分析解决方案

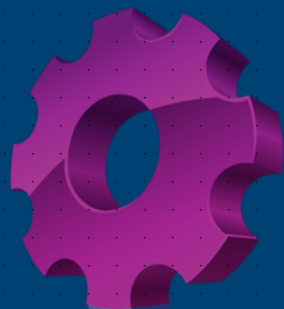
数据收集

数据处理

数据分析

数据可视化

数据产品



可视分析流程



登录

取消

为了实现这一目标，海云独立开发了一款大数据可视化工具帮助用户快速处理数据、全面分析数据和科学运营数据

可视分析案例（公安行业）

使用部门

- 情报、刑侦、技侦等部门

业务痛点

- 数据信息过于复杂，学习成本高，可视化能力弱
- 数据量大，警员人工作业压力大，容易遗漏关键数据线索
- 分析与可视化脱节，无法直接判断关联关系
- 主流产品是海外技术（IBM I2），安全风险大



Analytics

图易分析

Log In



嫌疑人跟踪分析



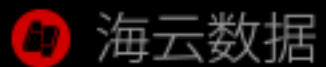
Analytics

图易分析

Log In



串并案分析



感谢关注！
Thank you!

Email: zhaodan@hiynn.com