山东省计算中心
SHANDONG COMPUTER SCIENCE CENTER

面向多源异构数据的电子取证关键技术

王连海

山东省计算中心（国家超级计算济南中心）研究员

2014-09-25

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

关于我

王连海
山东省计算中心（国家超级计算济南中心）
二级研究员
山东省计算机网络重点实验室 总工

32号）

主要内容

- 背景
- 国内外研究现状
- 研究工作
- 下一步的工作

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

背景

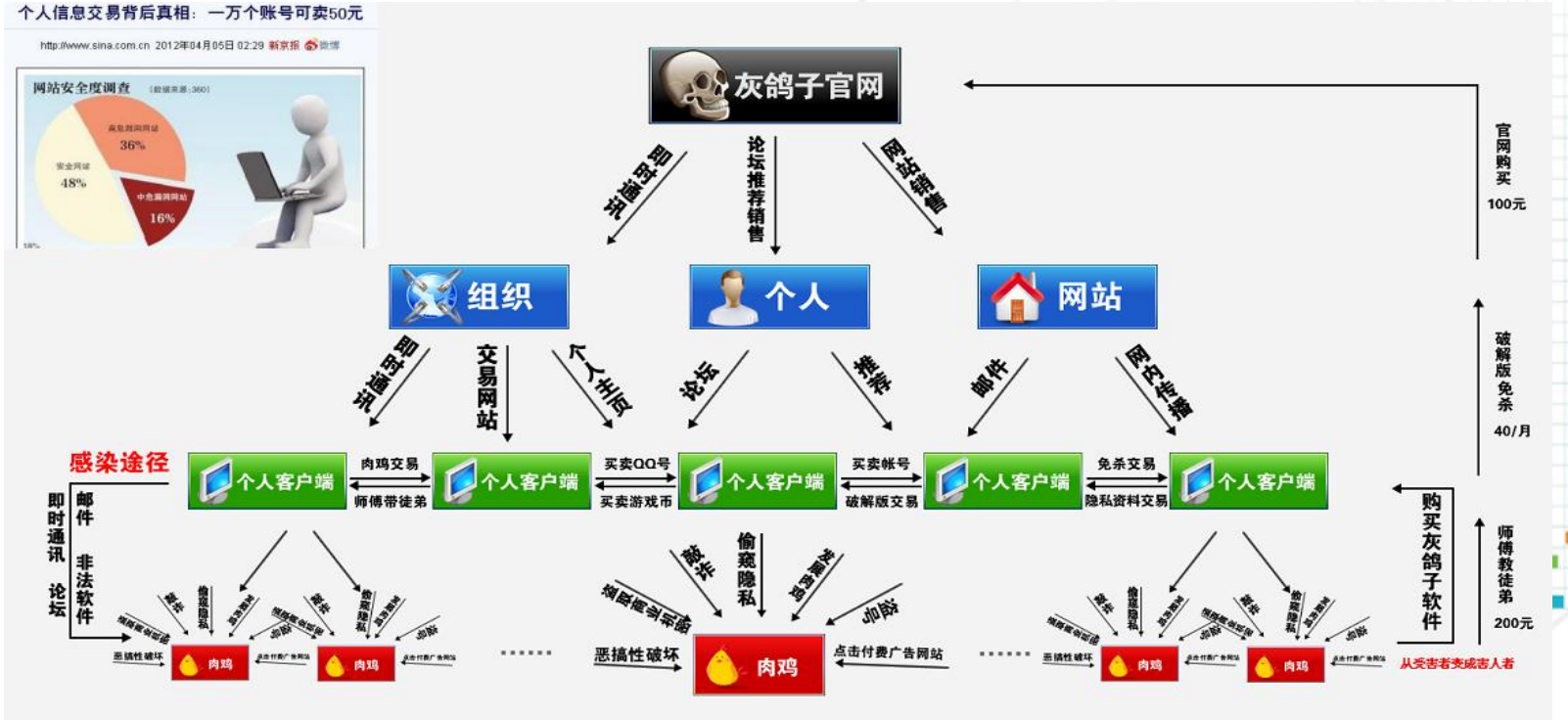
移动互联网日新月异

云计算飞速发展

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

背景

新信息时代在给人们生活带来便利的同时，也使得网络犯罪活动更加猖獗，逐渐形成一系列黑色产业链。



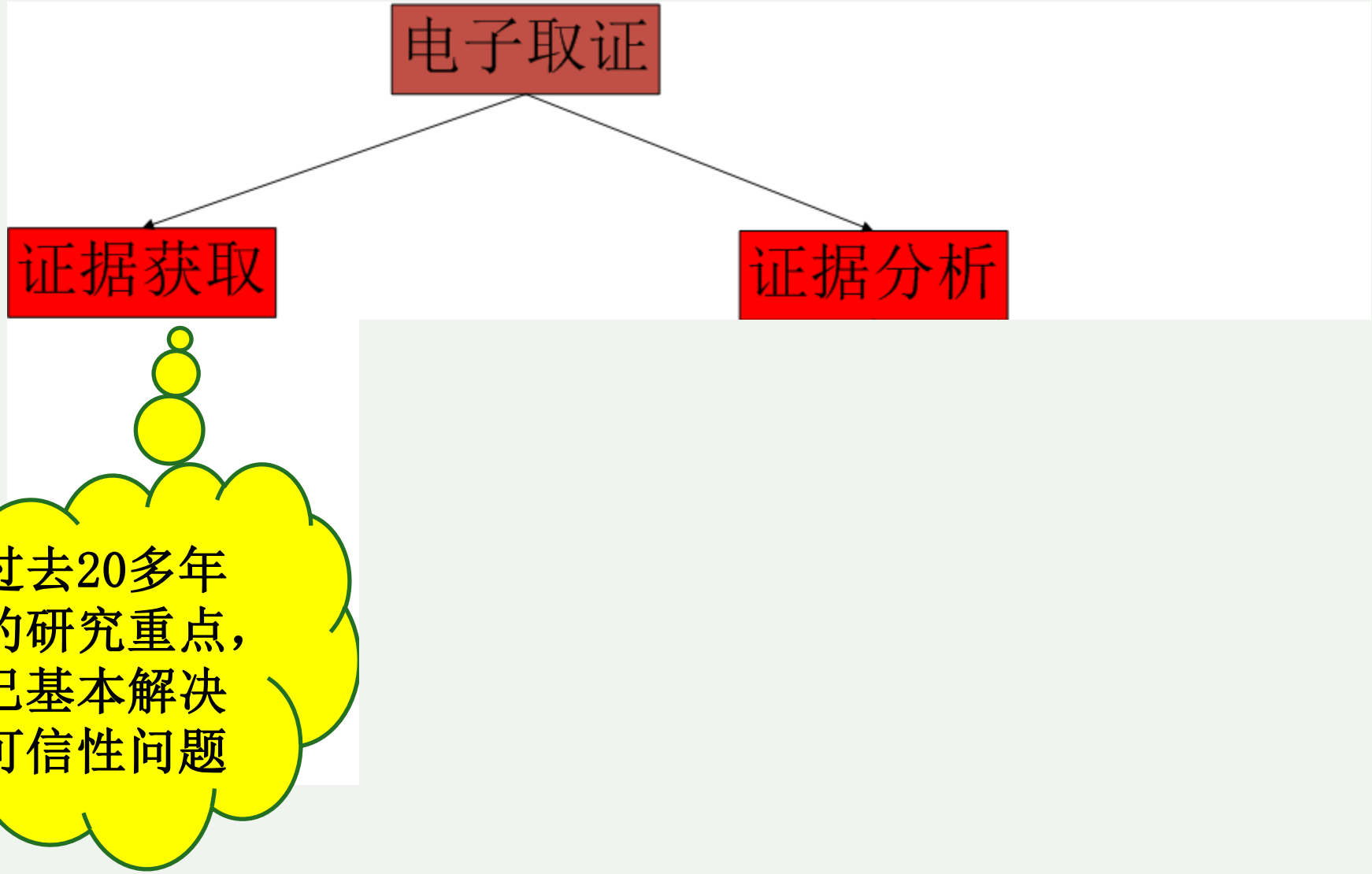
背景

- 要打击新信息时代的犯罪，需要有真实、可靠、完整的电子证据。
- 但是，传统的取证分析方法自动化程度低、可信性差，往往需要人工参与，难以应对海量多源异构证据分析。

因此需要研究新的电子取证分析方法应对多源异构的证据。

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状

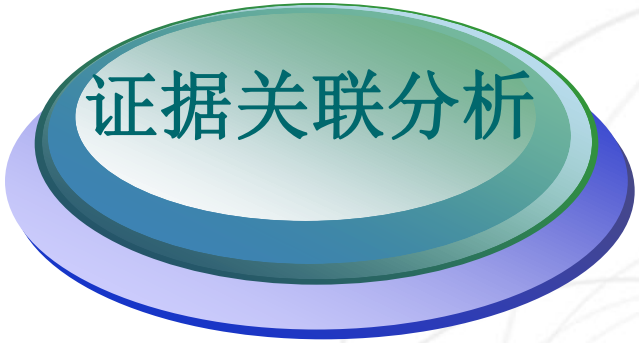


SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状



李辉等提出基于交互式知识发现的入侵事件关联分析方法

Wang W.等提出了基于攻击组及攻击场景进行推理的取证分析方法

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状

自动化分析

- 冯登国等给出具有自学习能力的专家系统分析证据
- Gladyshev运用有限状态机进行事件自动化分析
- Abbott J .等提出了事件自动重构方法
- 伏晓等提出基于自动证据分析的层次化入侵场景重构方法

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状

证据融合

许榕生等使用基于粗糙集理论的关联规则来研究数字证据的提取

徐晓滨等提出基于随机集理论的多源信息统一表示与建模方法

Case A.等开发了一种工具FACE，实现了从多种数据源中自动发现证据。

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

国内外研究现状

国际标准ISO/IEC27037建立了可信的取证流程和可信的证据获取保存方法

孙国梓等研究了基于可信概率的证据推理方法

王连海等研究了在线证据的可信获取分析等问题

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

现有方法的局限性

1

自动化程度低，缺乏实战意义的可信高效自动化分析技术，难以适应当前电子证据海量、多源异构化的趋势。

2

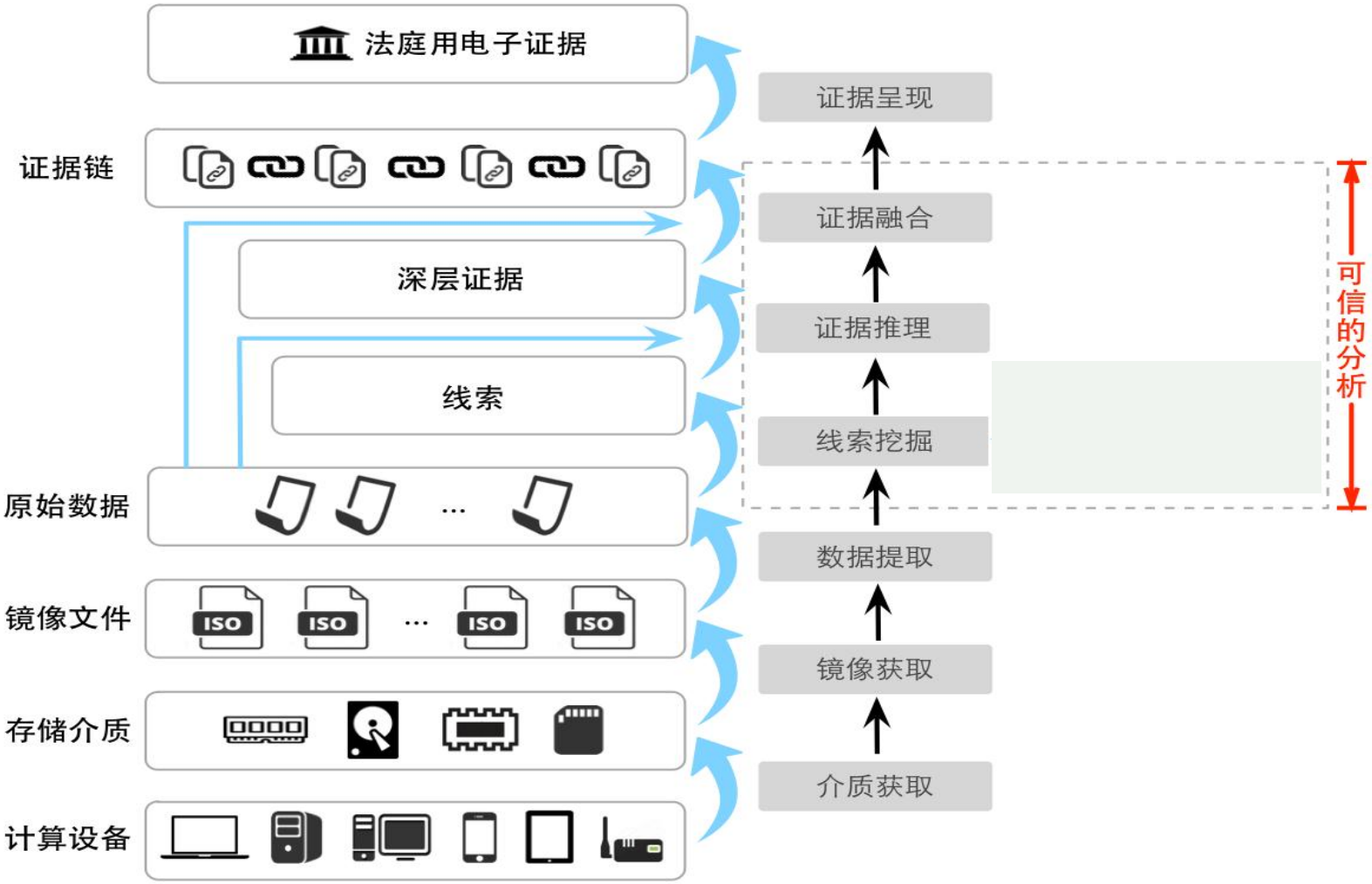
对多源异构证据的支持度不高，多源异构证据的融合程度低，得到的分析结果往往以偏概全，严重影响了电子证据的可信性。

3

对关联、推理、融合阶段的可信性没有充分考虑，所得证据的正确性难以评估。

研究工作

取证框架


法庭用电子证据
证据链
深层证据
线索
原始数据
镜像文件
存储介质
计算设备
ISO
证据呈现
证据融合
证据推理
线索挖掘
数据提取
镜像获取
介质获取
可信的分析

SHANDONG COMPUTER SCIENCE CENTER 山东省计算中心

云计算平台中证据的获取

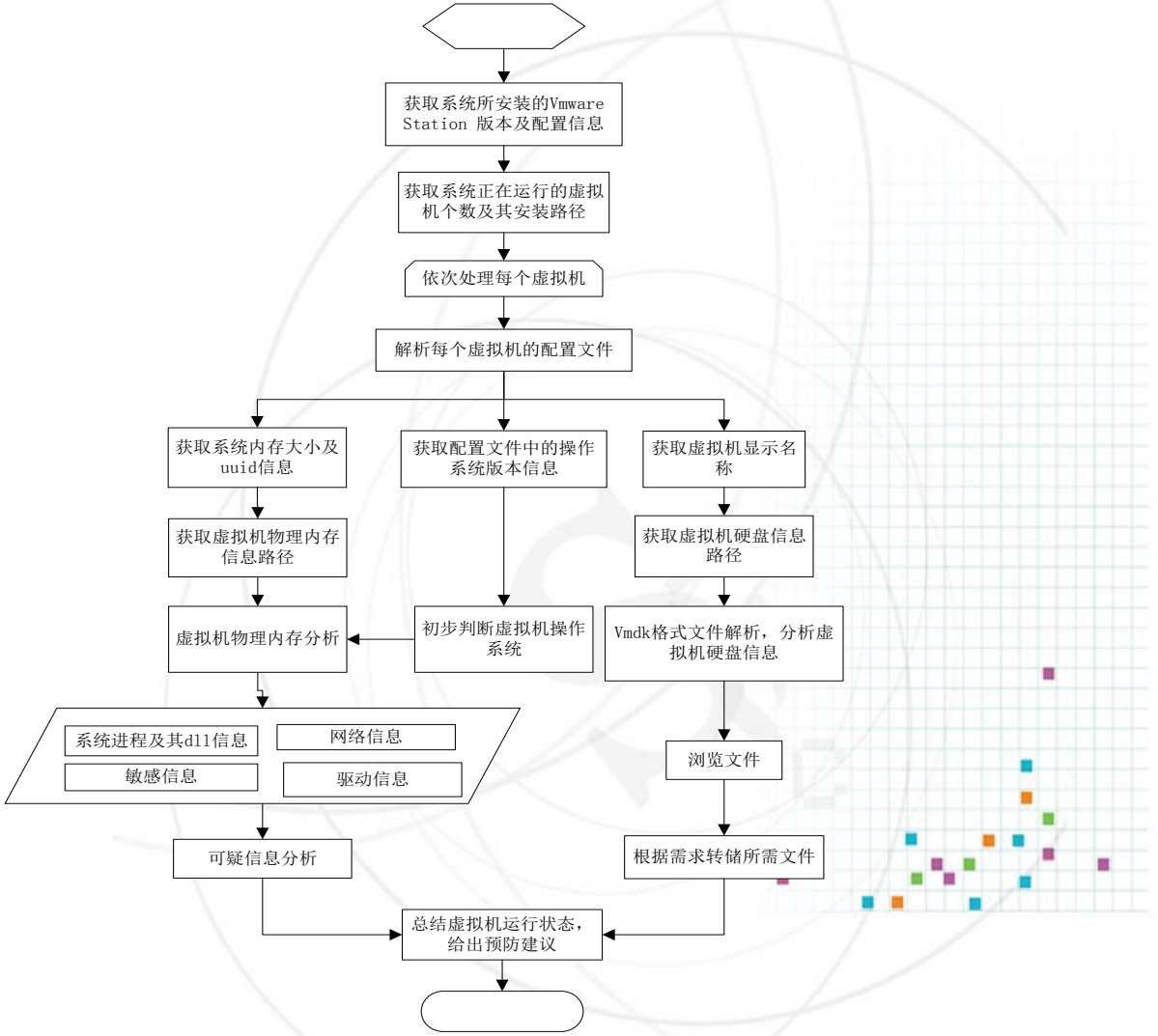
- 云计算安全性不容乐观，各种“云中犯罪”呈上升趋势。
- 传统网络环境下的证据获取已有基本解决方案，而针对云计算平台的方法较少。
- 云计算平台中违法犯罪活动的取证关键在于如何在不影响其正常运转的同时，获取云平台中各虚拟机中的证据。

我们研发了基于物理内存分析的云计算监控、取证系统用于云平台相关证据的获取分析。

SHANDONG COMPUTER SCIENCE CENTER 山东省计算中心

云计算平台中证据的获取

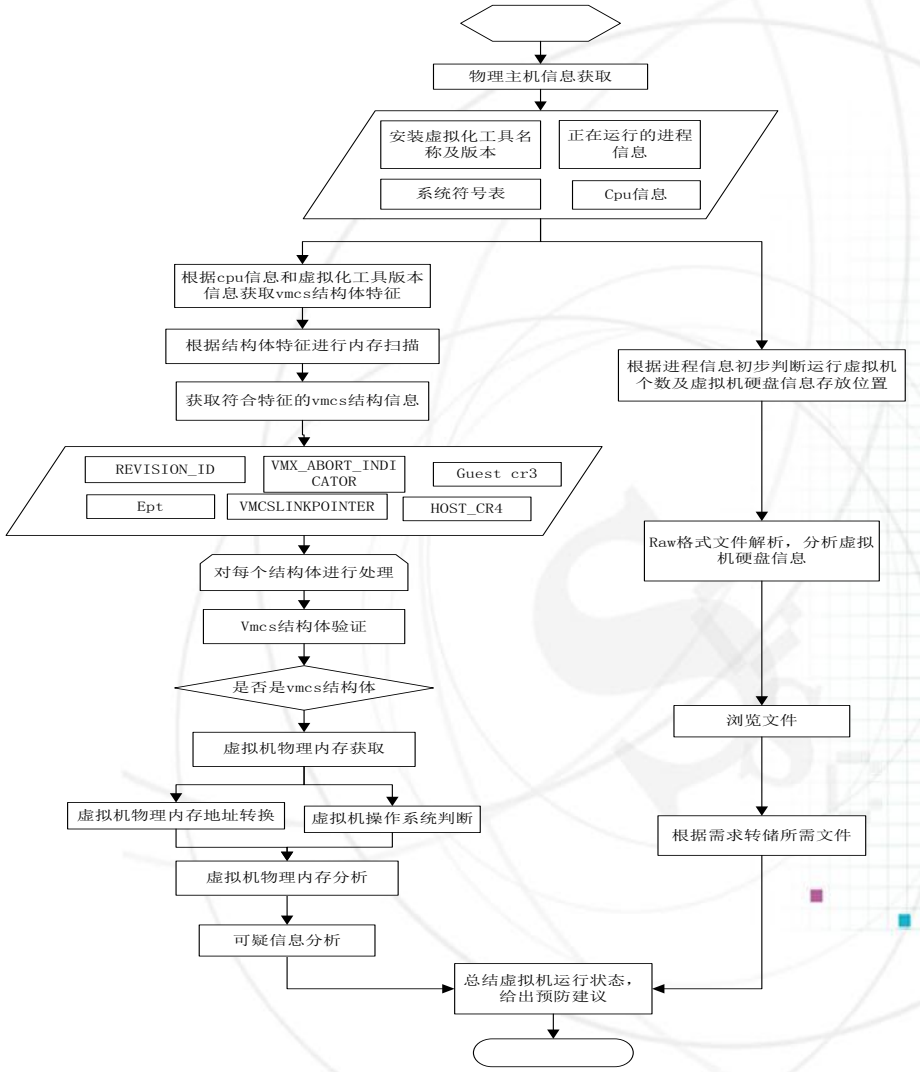
- Vmware Station下虚拟机物理内存和硬盘信息的获取和分析



SHANDONG COMPUTER SCIENCE CENTER 山东省计算中心

云计算平台中证据的获取

- KVM下虚拟机物理内存和硬盘信息的获取和分析



SHANDONG COMPUTER SCIENCE CENTER 山东省计算中心

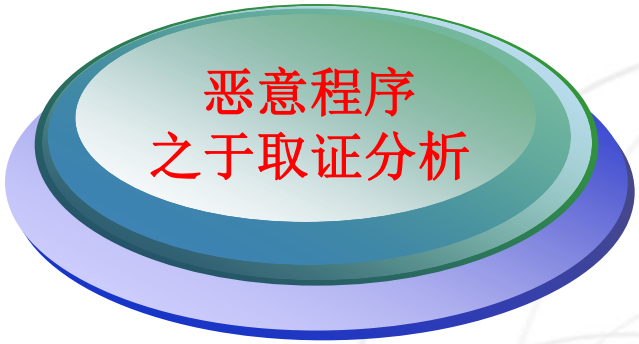
云计算平台中证据的获取

关键技术

- 基于物理内存分析的云计算监控管理模型
- 基VMCS结构体特征值匹配
- 虚拟机物理地址到物理机物理地址的转换
- 虚拟机虚拟地址到物理机物理地址的转换
- 基于物理内存分析的可疑信息分析
  - ApiHook分析
  - 签名分析
  - 恶意进程检测
  - DLL注入检测

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

恶意程序的判定



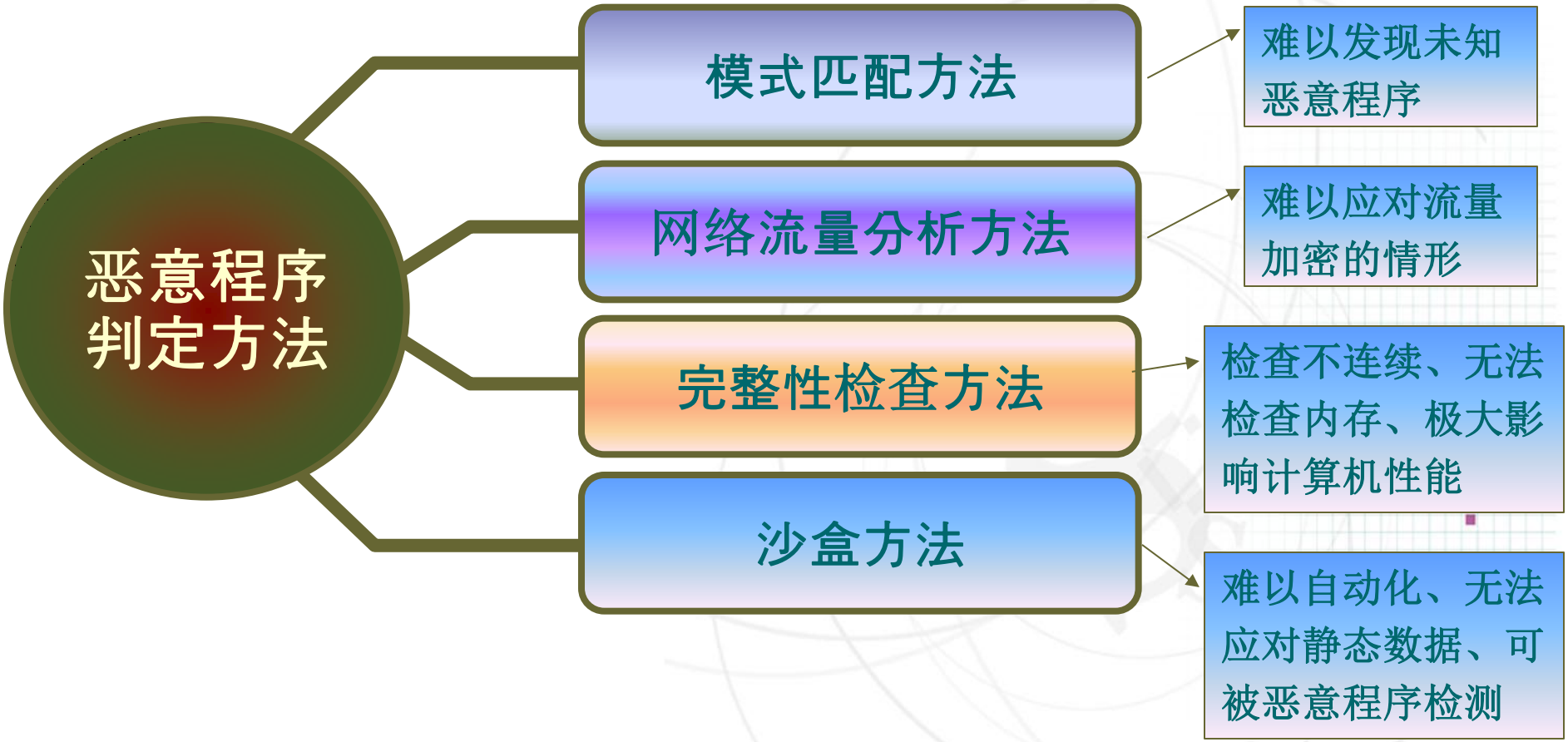
恶意程序在很大一部分违法犯罪活动中起至关重要的作用，其往往成为左右证据分析结果的重要因素。

恶意程序往往阻碍取证分析，其通过隐藏或擦除犯罪活动的痕迹使得取证分析非常困难。

因此在取证分析中需要首先判定是否存在恶意程序。

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

恶意程序的判定

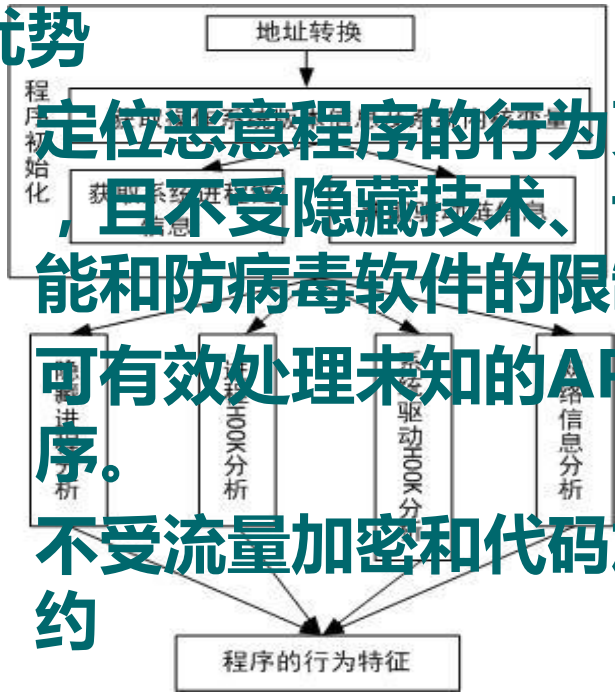


SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

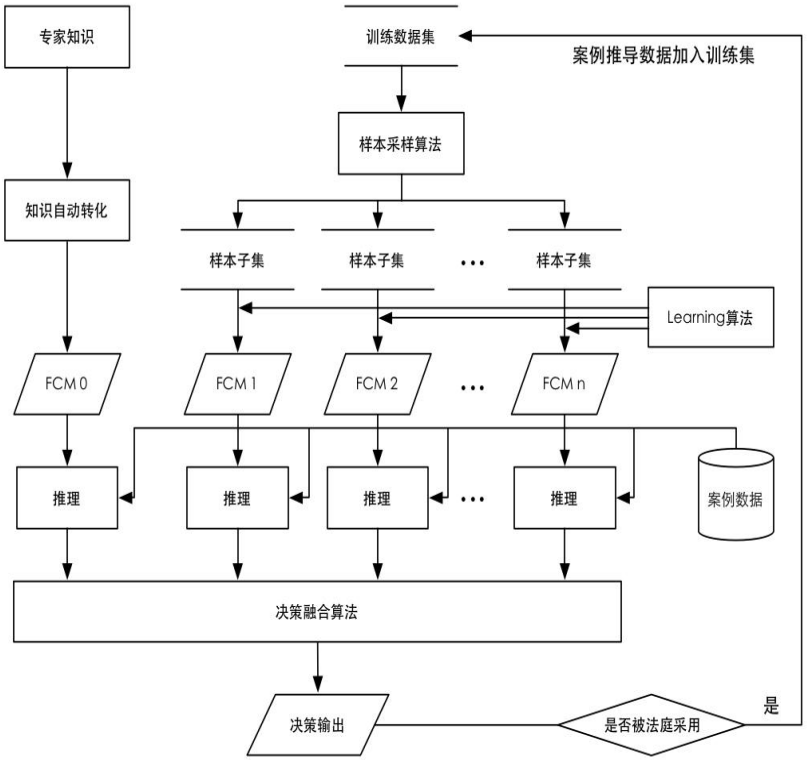
恶意程序的判定

基于内存分析和模糊认知的恶意性判定判定方法

基于内存分析的程序行为特征提取



基于模糊认知的程序恶意性判定



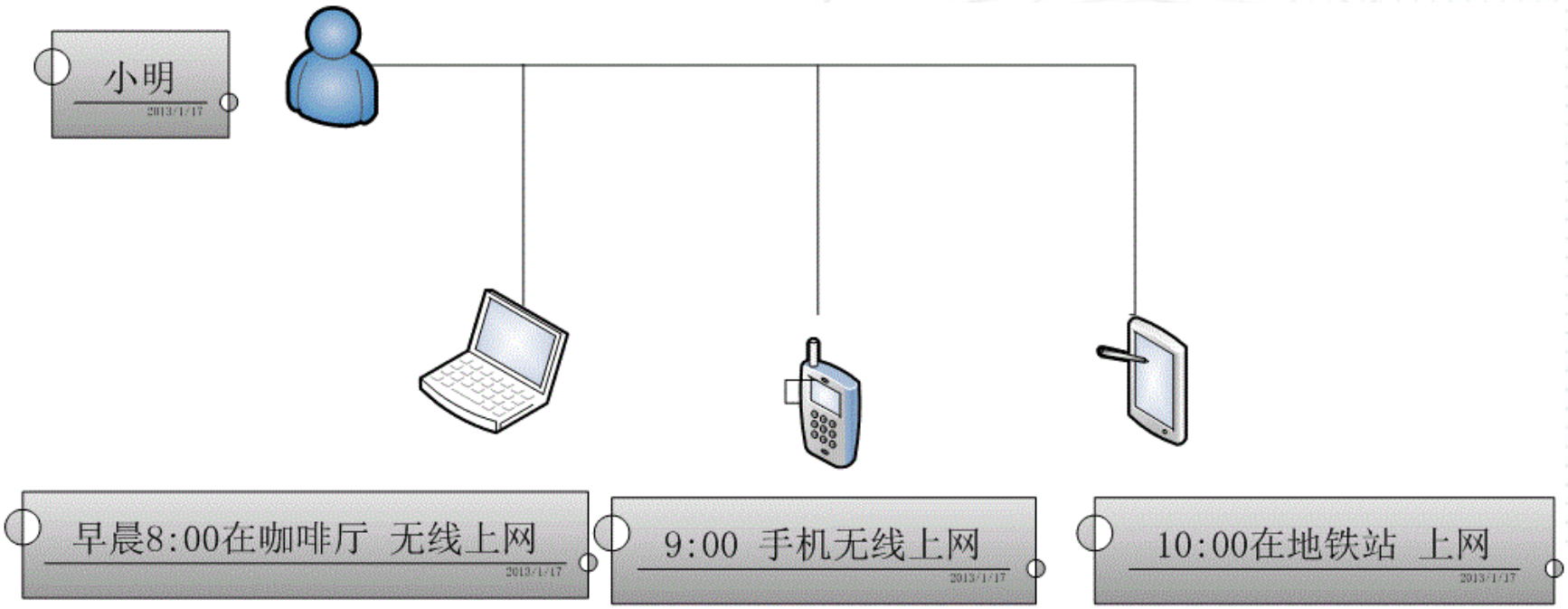
优势

- **定位恶意程序的行为更加准确，且不受隐藏技术、计算机性能和防病毒软件的限制。**
- **可有效处理未知的APT恶意程序。**
- **不受流量加密和代码加密的制约**

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

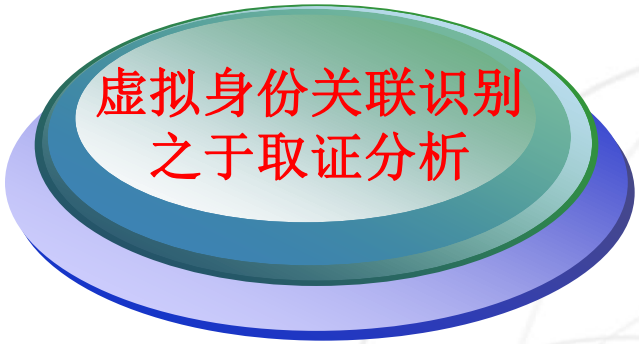
虚拟身份的关联识别

- 移动互联网的便捷性、匿名性特点使虚拟身份关联识别存在很多困难



SHANDONG COMPUTER SCIENCE CENTER 山东省计算中心

虚拟身份的关联识别



虚拟身份众多造成案件侦破过程中对犯罪嫌疑人的身份难以追踪。

各种身份隐藏技术的应用使得犯罪嫌疑人的真实身份难以被发现。

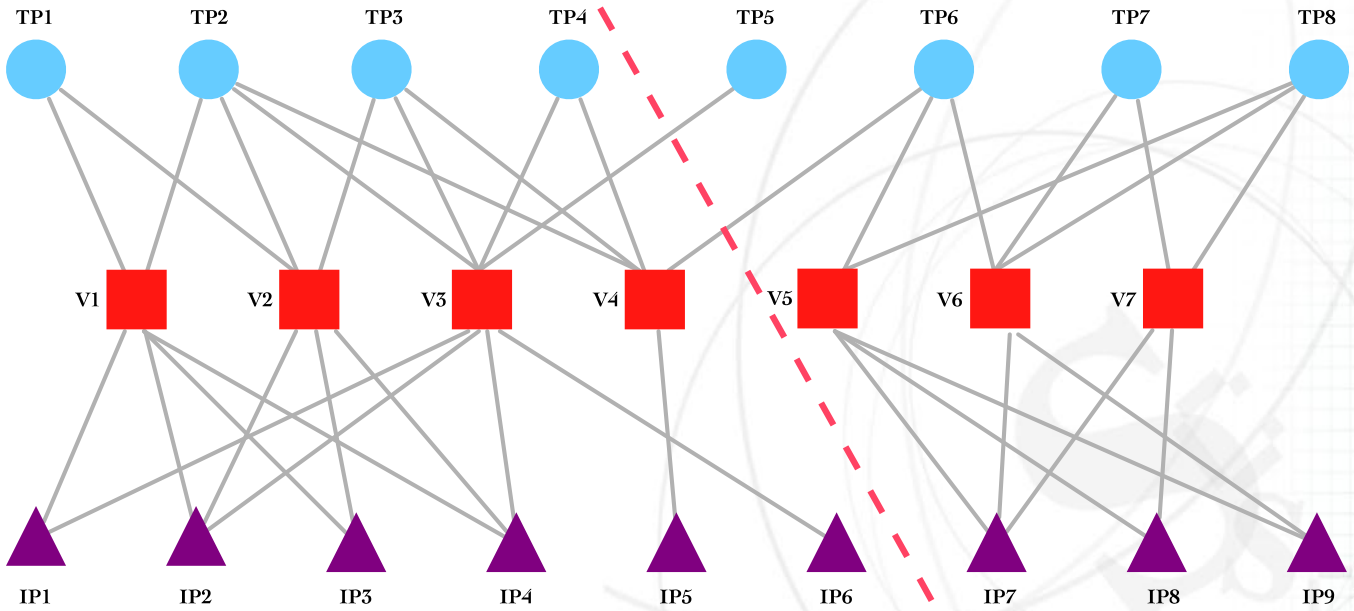
因此虚拟身份的关联识别是取证分析中亟需解决的关键难题。

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

虚拟身份的关联识别

- 针对于虚拟身份的关联问题，我们依据虚拟身份在时间、空间、行为、社交关系、资金流向等进行关联挖掘，并开发了相应的身份关联系统。

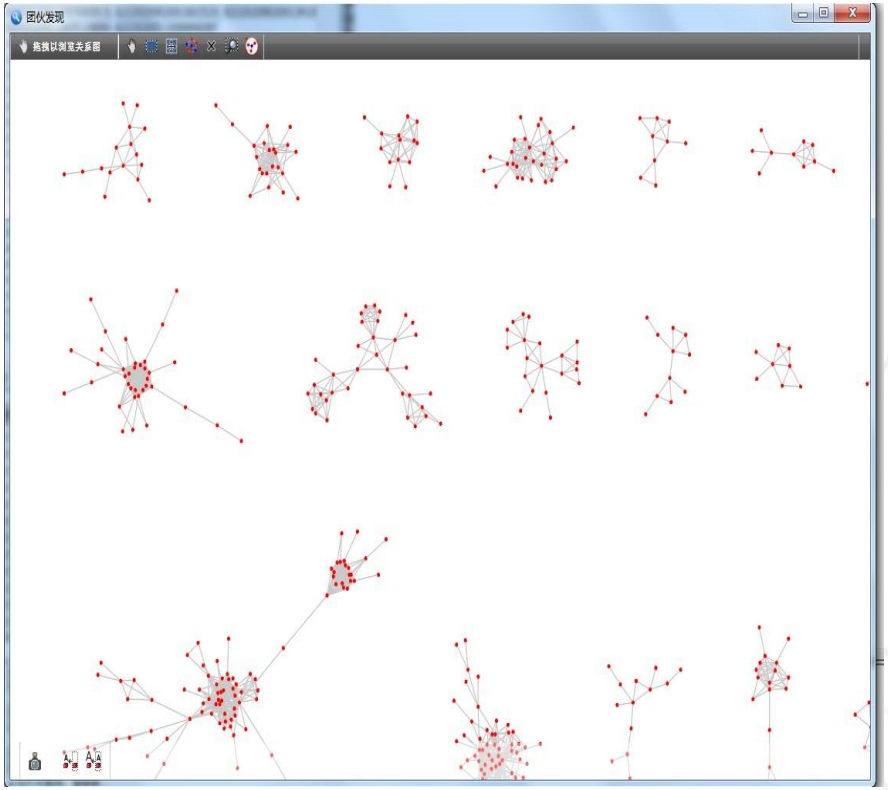
❑ 虚拟身份、时间、Ip关联


TP1
TP2
TP3
TP4
TP5
TP6
TP7
TP8
V1
V2
V3
V4
V5
V6
V7
IP1
IP2
IP3
IP4
IP5
IP6
IP7
IP8
IP9

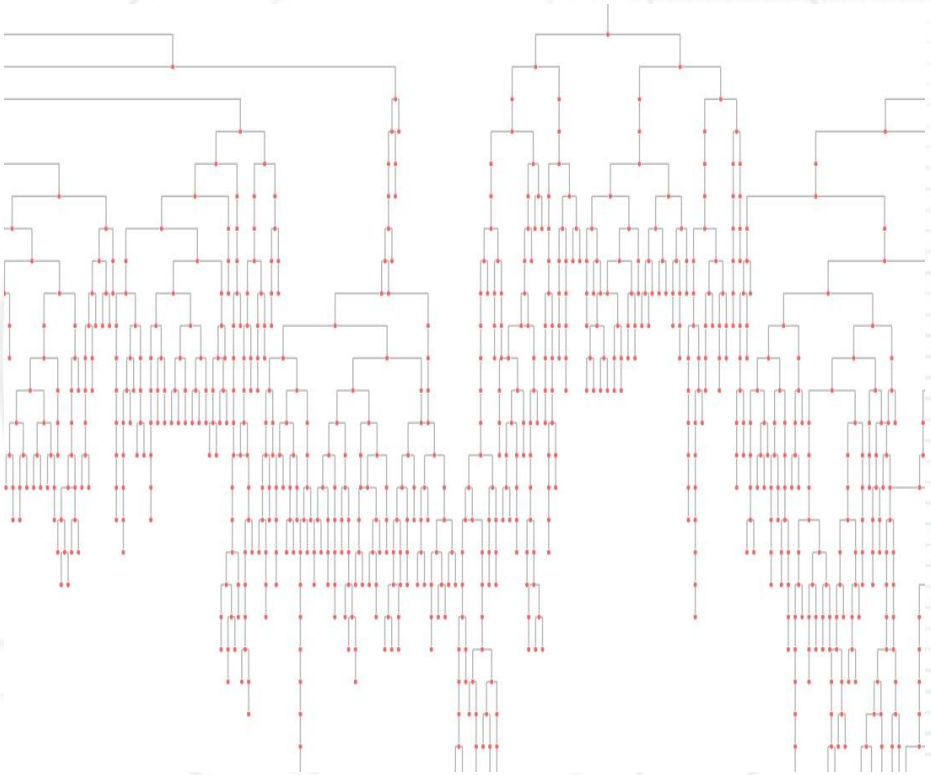
SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

虚拟身份的关联识别

- 犯罪团伙关联分析



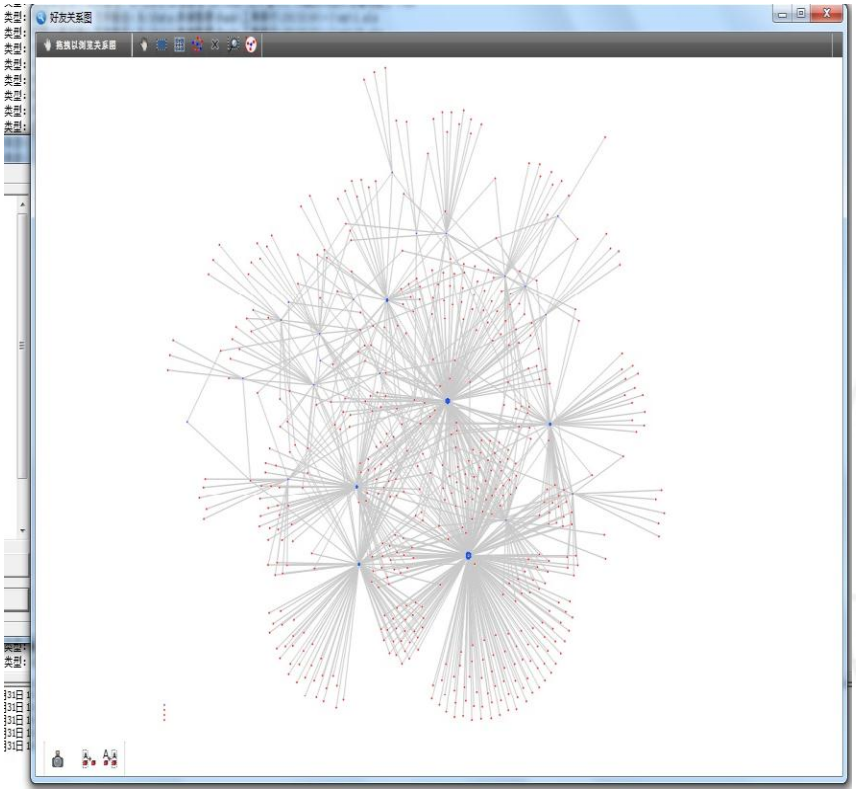
- 内部组织结构关联分析



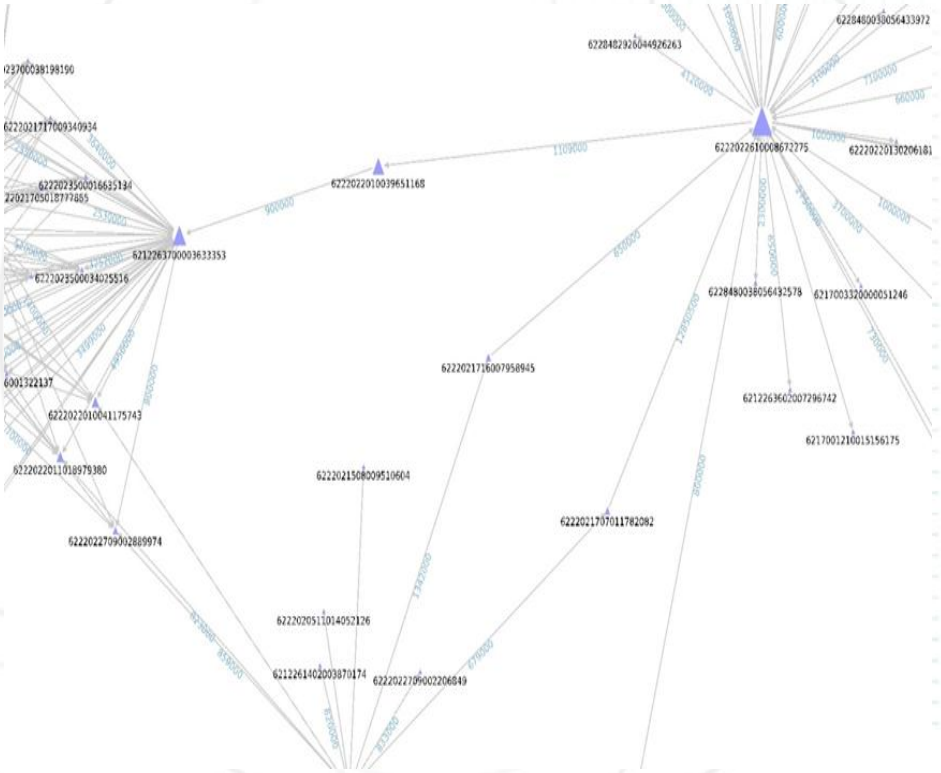
SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

虚拟身份的关联识别

- 好友关系关联分析



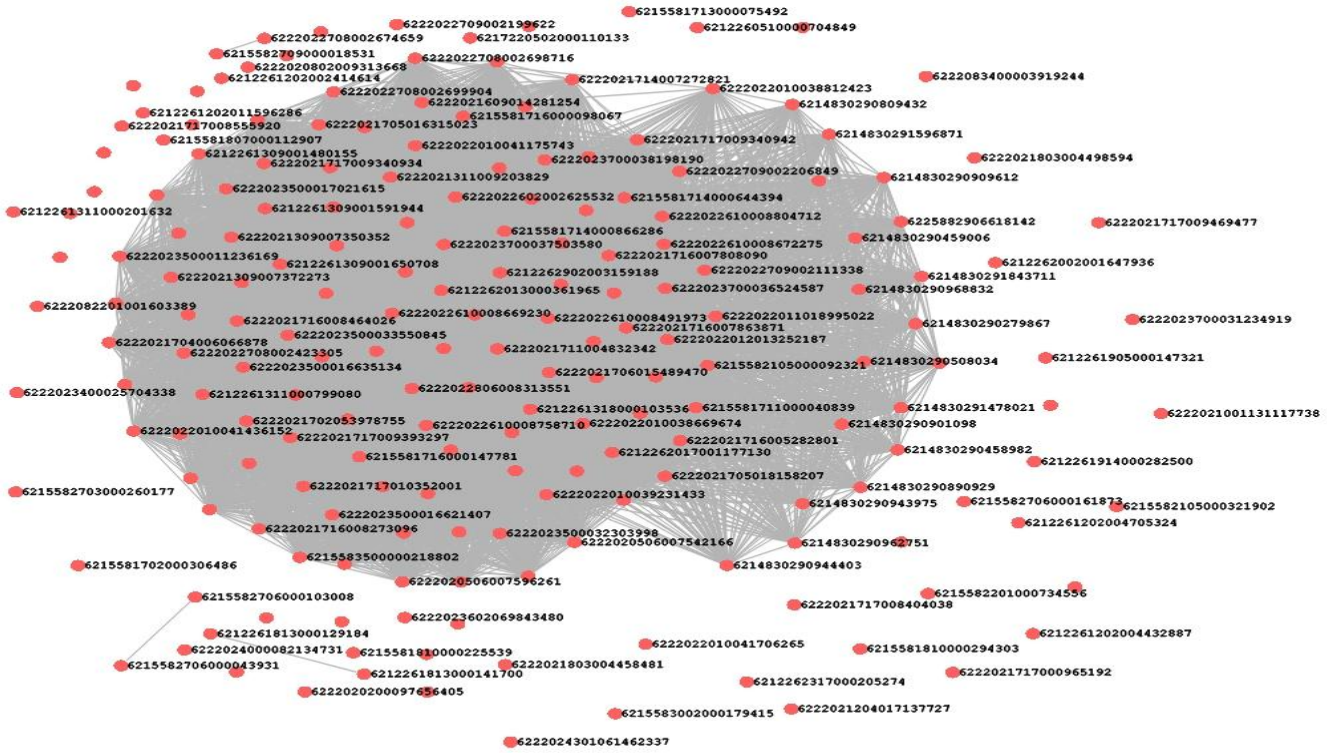
- 资金流向关联分析



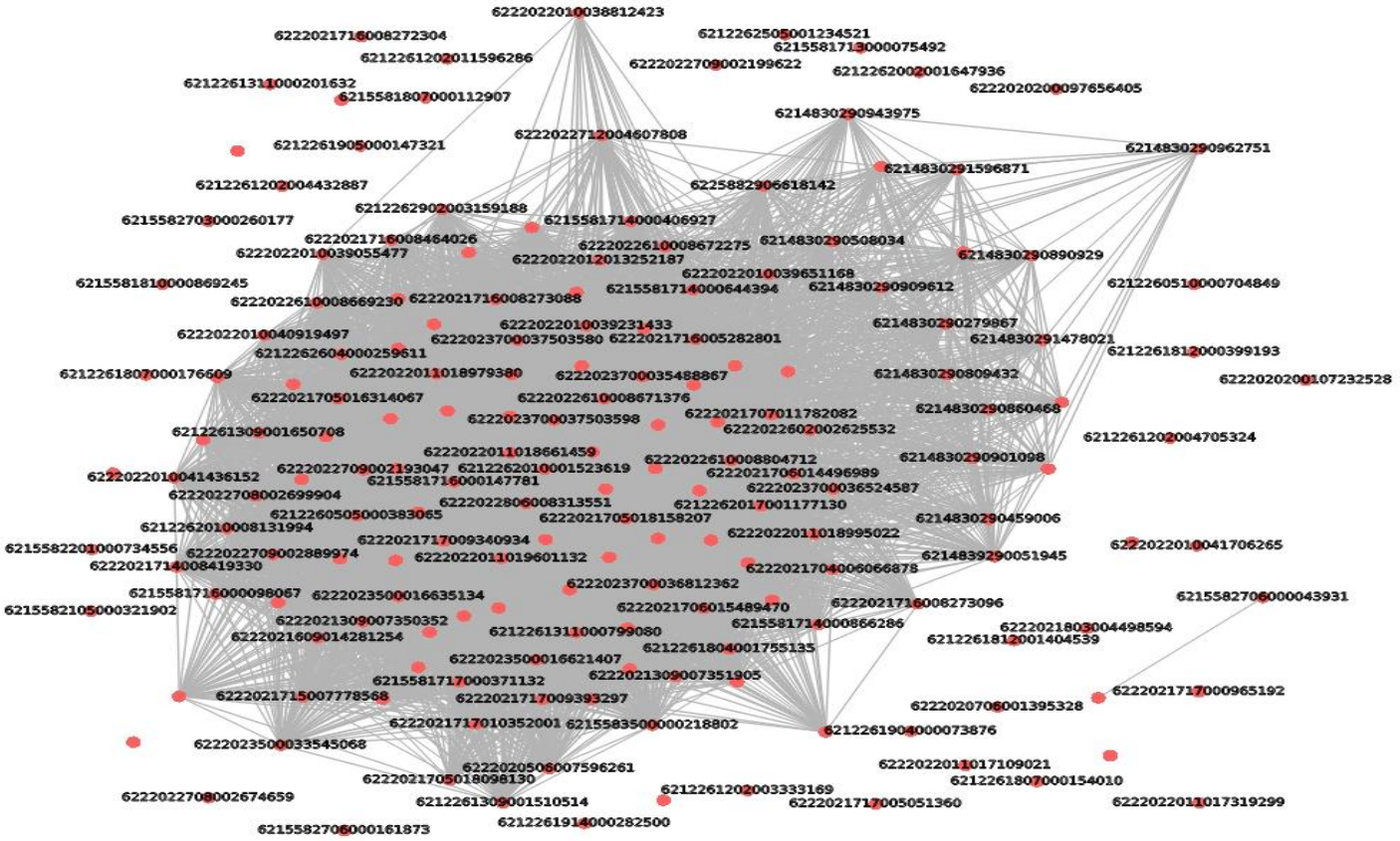
SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

虚拟身份的关联识别

- IP关联分析



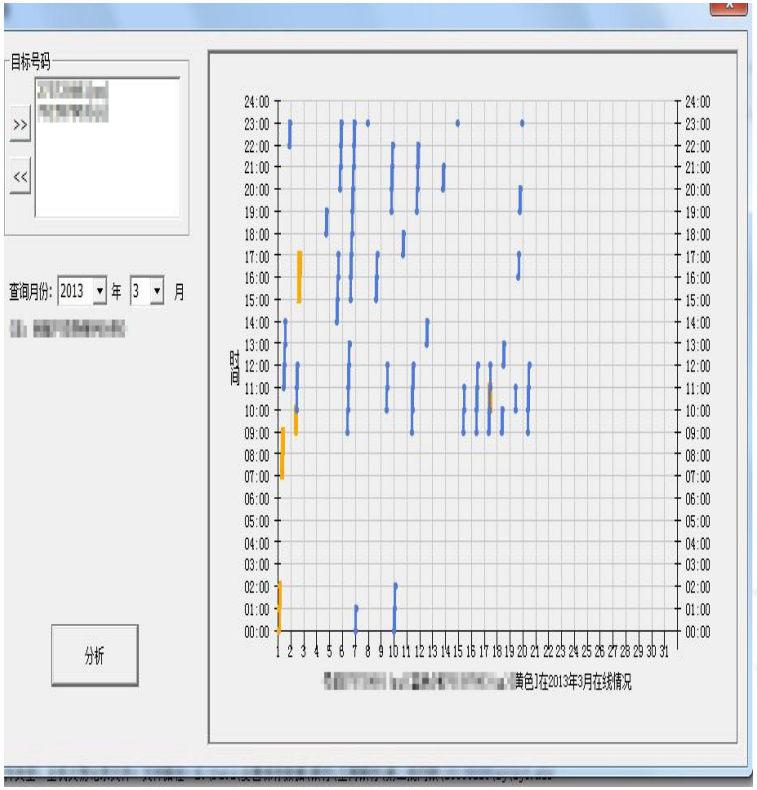
❑ Ip、MAC 地址的关联


6222022010038812423
6222021716008272304
6212262505001234521
6215581713000075492
6212262002011596286
6222022709002199622
6212262002001647936
6212261311000201632
6222020200097656405
6215581807000112907
6214830290943975
6222022712004607808
6212261905000147321
6214830290962751
6214830291596871
6212261202004432887
6225882906618142
6212262902003159188
6215582703000260177
6215581714000406927
6222021716008464026
6222022610008672275
6214830290508034
6222022010039055477
6222022012013252187
6214830290890929
6222022010039651168
6215581810000869245
6215581714000644394
6214830290909612
6212260510000704849
6222022610008669230
6222021716008273088
6222022010039231433
6214830290279867
6222022010040919497
6222023700037503580
6222021716005282801
6214830291478021
6212262604000259611
6212261812000399193
6212261807000176609
6222022011018979380
6222023700035488867
6214830290809432
6222020200107232528
6222021705016314067
6222022610008671376
6222021707011782082
6214830290860468
6222023700037503598
6222022602002625532
6212261309001650708
6212261202004705324
6222022011018661459
6222022610008804712
6214830290901098
6222022010041436152
6222022709002193047
6212262010001523619
6222021706014496989
6215581716000147781
6222023700036524587
6222022708002699904
6222022806008313551
6212262017001177130
6212260505000383065
6222021705018158207
6214830290459006
6212262010008131994
6222022011018995022
6215582201000734556
6222022709002889974
6222022011019601132
6214839290051945
6222022010041706265
6222021714008419330
6222021704006066878
6222023700036812362
6215581716000098067
6222023500016635134
6215582105000321902
6222021706015489470
6222021716008273096
6215582706000043931
6222021309007350352
6215581714000866286
6222021803004498594
6212261811001404539
6222021609014281254
6212261311000799080
6212261804001755135
6222023500016621407
6222021309007351905
6215581717000371132
6222021715007778568
6222021717009393297
6222020706001395328
6222021717000965192
6222021717010352001
6215583500000218802
6212261904000073876
6222023500033545068
6222020506007596261
6222022011017109021
6222021705018098130
6212261807000154010
6212261202003333169
6222022708002674659
6212261309001510514
6222021717005051360
6222022011017319299
6212261914000282500
6215582705000161873

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

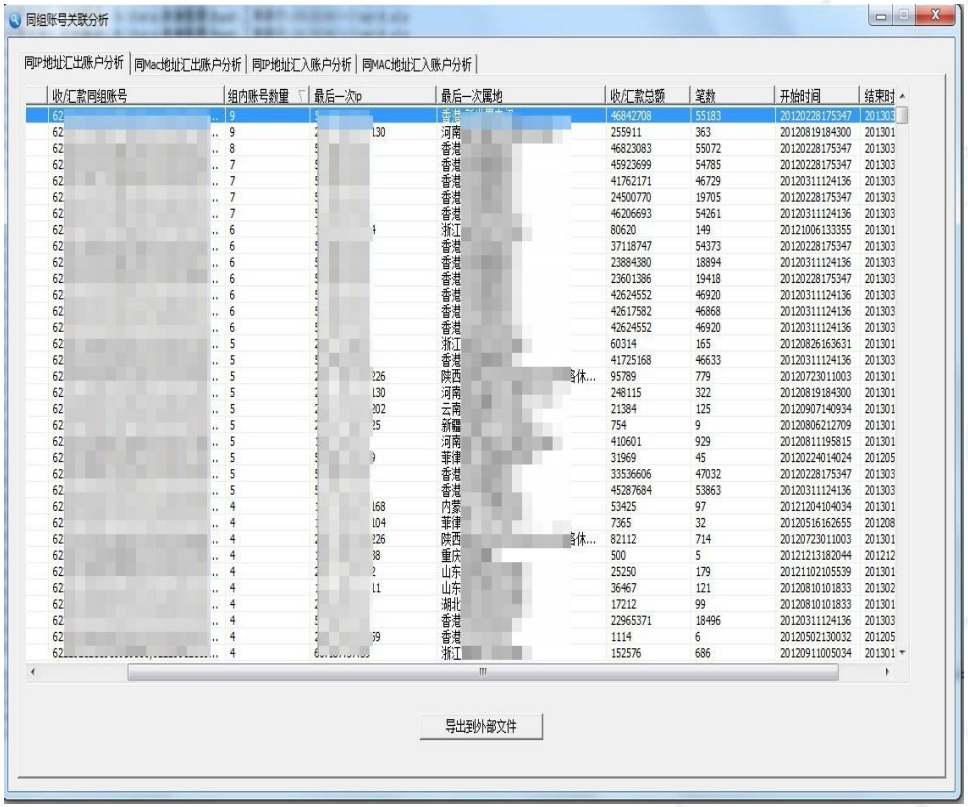
虚拟身份的关联识别

- 时间线分析



- 地理位置分析

• 网银账号分析

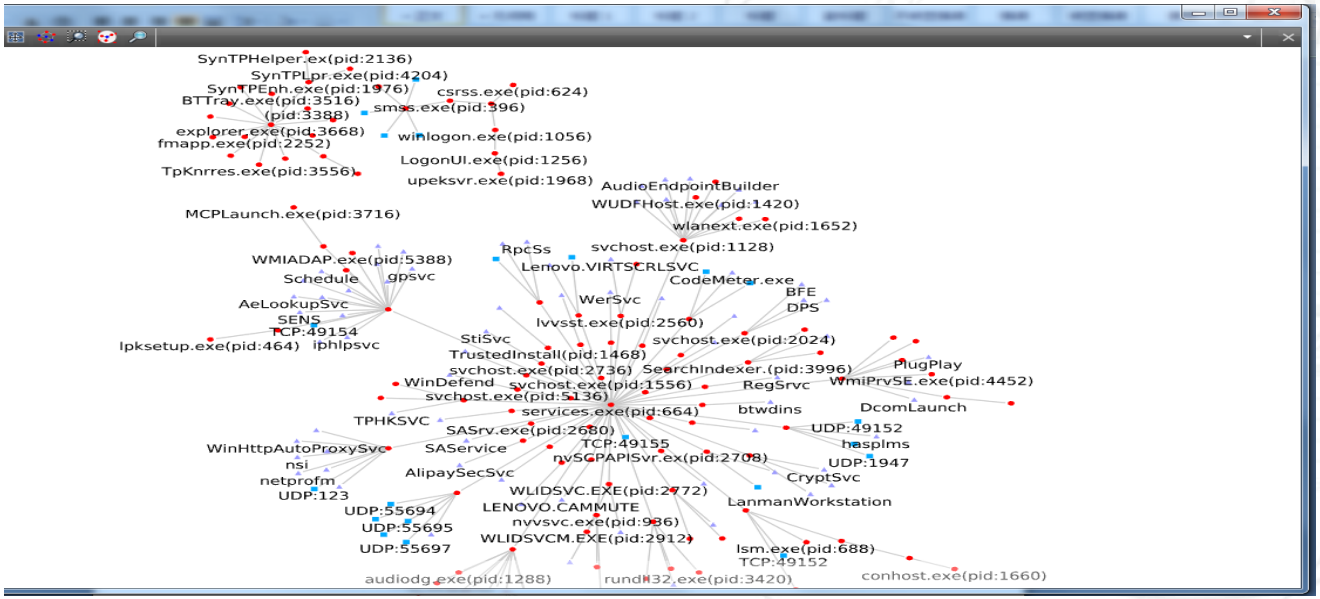

同组账号关联分析
同IP地址汇出账户分析
同Mac地址汇出账户分析
同IP地址汇入账户分析
同MAC地址汇入账户分析
收/汇款同组账号
组内账号数量
最后一次ip
最后一次属地
收/汇款总额
笔数
开始时间
结束时
导出到外部文件

SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

证据的关联分析

- 证据的关联是证据分析的核心与关键，我们依据获取数据特征建立证据间关联关系，并初步开发实现。

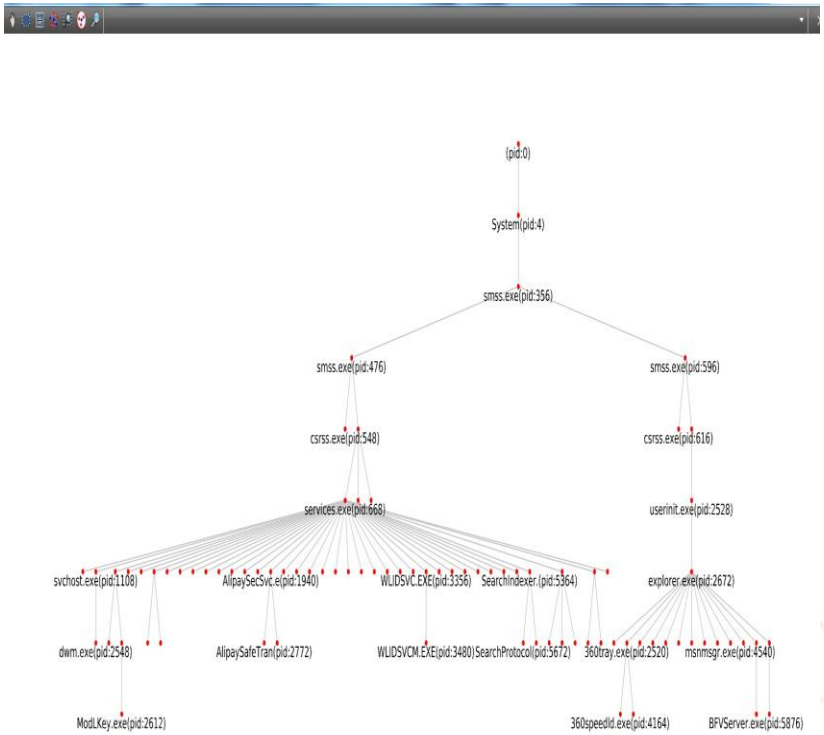
- **进程、网络端口、系统服务关联分析**



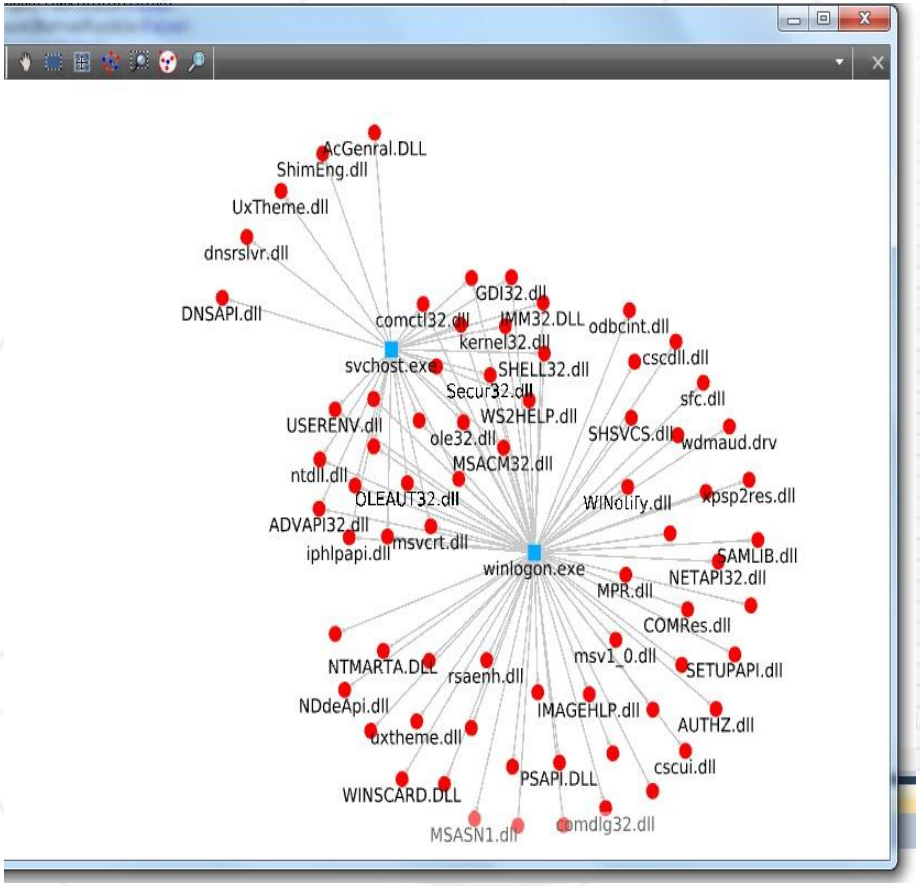
SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

证据的关联分析

- 进程链关联分析



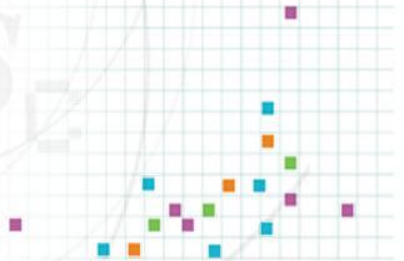
- DLL关联分析



SHANDONG COMPUTER SCIENCE CENTER
山东省计算中心

下一步的工作

- 电子取证高可信方法的构建
- 多源异构证据的推理方案
- 多源异构证据的融合方法



SHANDONG
COMPUTER SCIENCE
CENTER
山东省计算中心

Thanks!