全息安全

- 互联网空间中的导弹防御系统

百度
云安全部
郝轶

思考 困扰 实践 态度

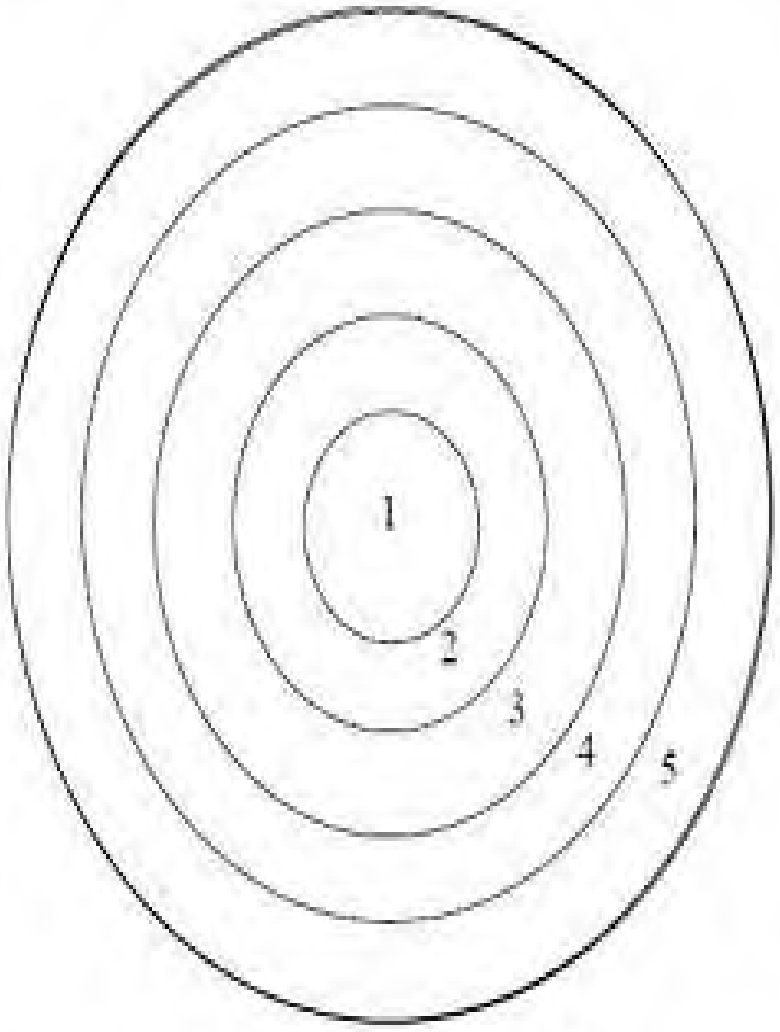
防护者的视角 - 塔防式的纵深防御

伊拉克战争 - 防护者视角

伊拉克战争 - 旁观者视角

伊拉克战争 - 攻击者视角

战争五环



1 领导 → 军政领导、防空预警系统、C^4ISR系统

2 系统关键要素 → 国家骨干企业、军工、石油、电力等目标

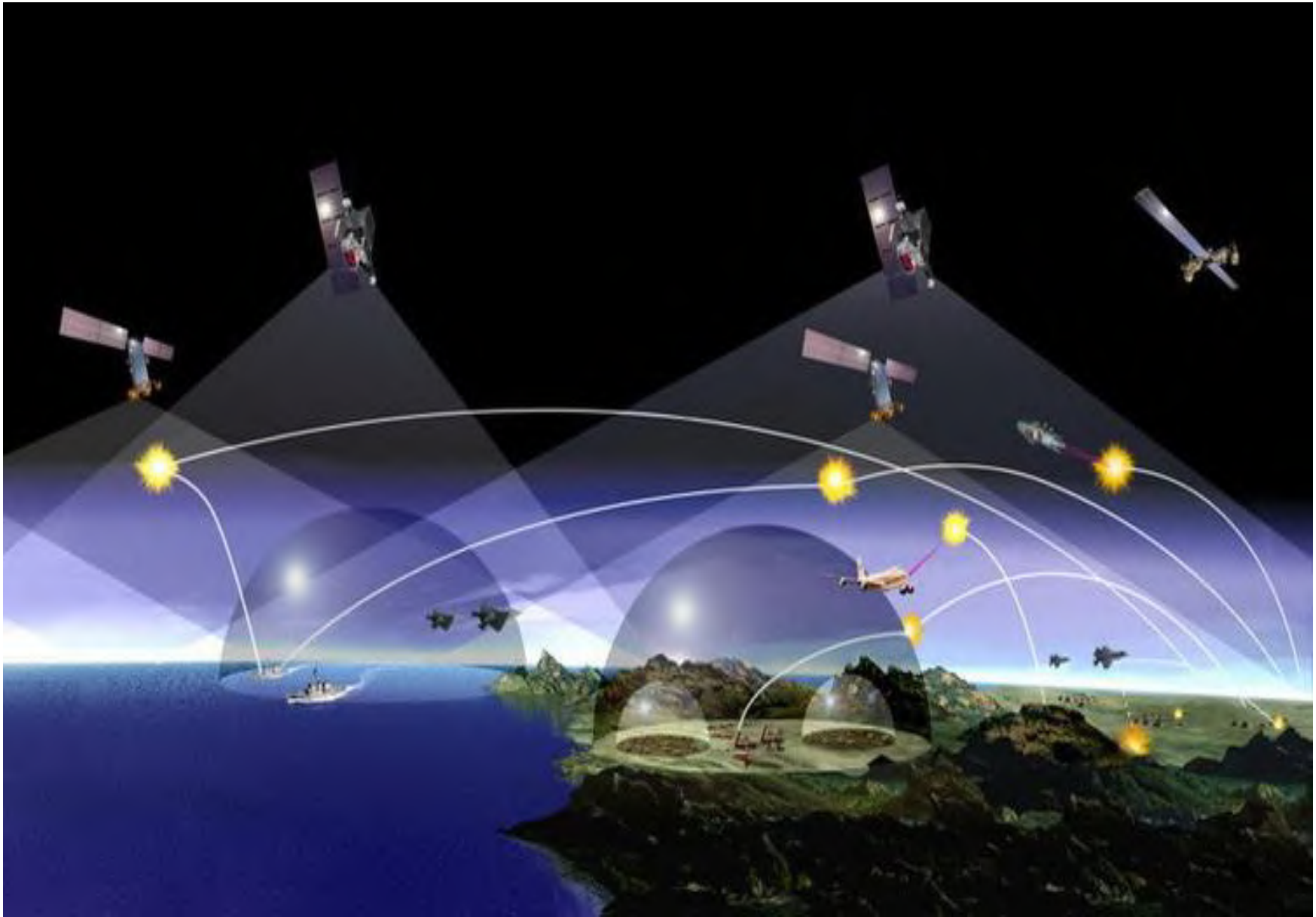
3 基础设施 → 道路桥梁机场等具有机动战争物资输送价值的目标

4 民众 → 通过心理战等来影响民众的精神心理状态

5 军队 → 军事人员、作战部队

斩首行动

反导弹系统



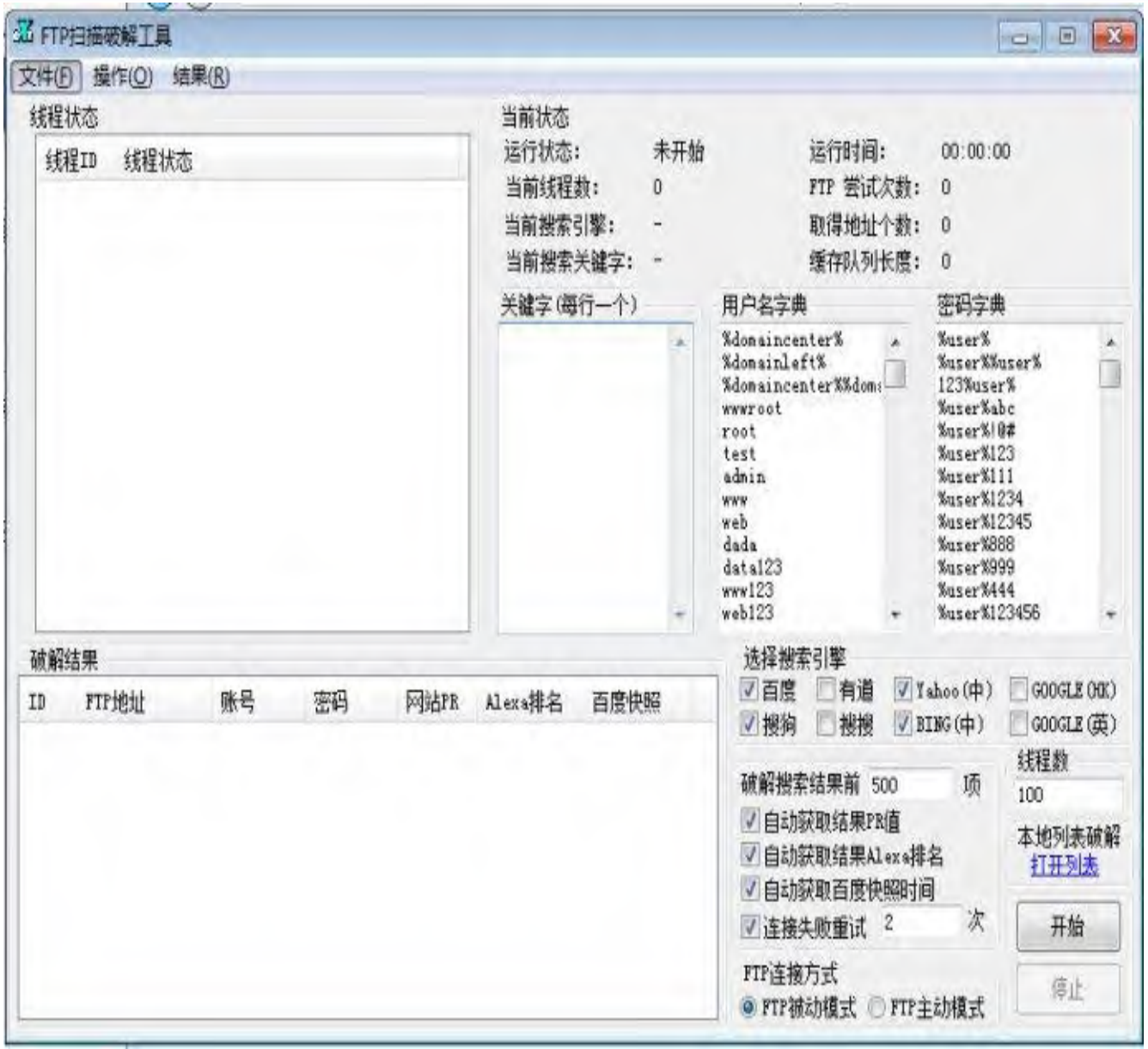
思考 困扰 实践 态度


互联网空间中的全息安全

内网：APT攻击场景

互联网：非持续性攻击场景

IDCC



最新提交 (62)

提交日期	漏洞名称
2015-11-27	广州联通某安全漏洞导致上万用户信息泄漏（宽带安装地址+身份证+姓名+电话+产品号码）
2015-11-27	厦门大学内网漫游
2015-11-27	每日视频播放量超过5亿的小咖秀任意登陆管理员漏洞
2015-11-27	土豆某系统SQL注入到Getshell
2015-11-27	宜搜科技多数重要系统文件读取+某系统命令执行(唤起厂商确认漏洞的良知)二
2015-11-26	天虹商场红领巾最新版APP多处越权(泄露用户信息/收货地址/订单记录)

最新确认 (1138)

提交日期	漏洞名称
2015-11-27	天融信TopScanner存在任意命令执行&文件遍历(无需登录)
2015-11-27	华数某系统存在多处SQL注入漏洞
2015-11-26	华数运维支撑系统SQL注入导致20多万用户信息泄露众多设备资源泄露
2015-11-26	华数某重要系统存在SQL注入
2015-11-25	虎扑分站某处SQL注入
2015-11-27	优酷某项目管理系统一枚弱口令涉及大量内部敏感信息

最新公开 (34040)

提交日期	漏洞名称
2015-10-11	万达飞凡APP某处SQL注入泄露电影业务数据
2015-10-13	齐乐聚所有分站存在SQL注入漏洞涉及65万用户信息
2015-10-13	秒针配置不当可导致大量内部信息泄露
2015-10-13	赶集网某重要系统账户控制不严导致大量内部信息可泄漏（影响多个内部站点）
2015-11-22	巨人网络某处DBA权限报错SQL注入
2015-08-26	中海达某设备产品存在设计缺陷(可导致敏感信息泄漏包括账号密码)

ATTACK ORIGINS

#	COUNTRY
44	China
20	Saudi Arabia
14	Turkey
8	United States
4	Taiwan
3	United Kingdom
2	Moldova
2	Spain
1	Venezuela
1	Sweden

ATTACK TYPES

#	PORT	SERVICE TYPE
37	22	ssh
21	137	unknown
15	23	telnet
9	50856	unknown
9	25	smtp
4	50864	unknown
3	445	microsoft-ds
2	80	http
1	5900	vnc
1	49570	unknown

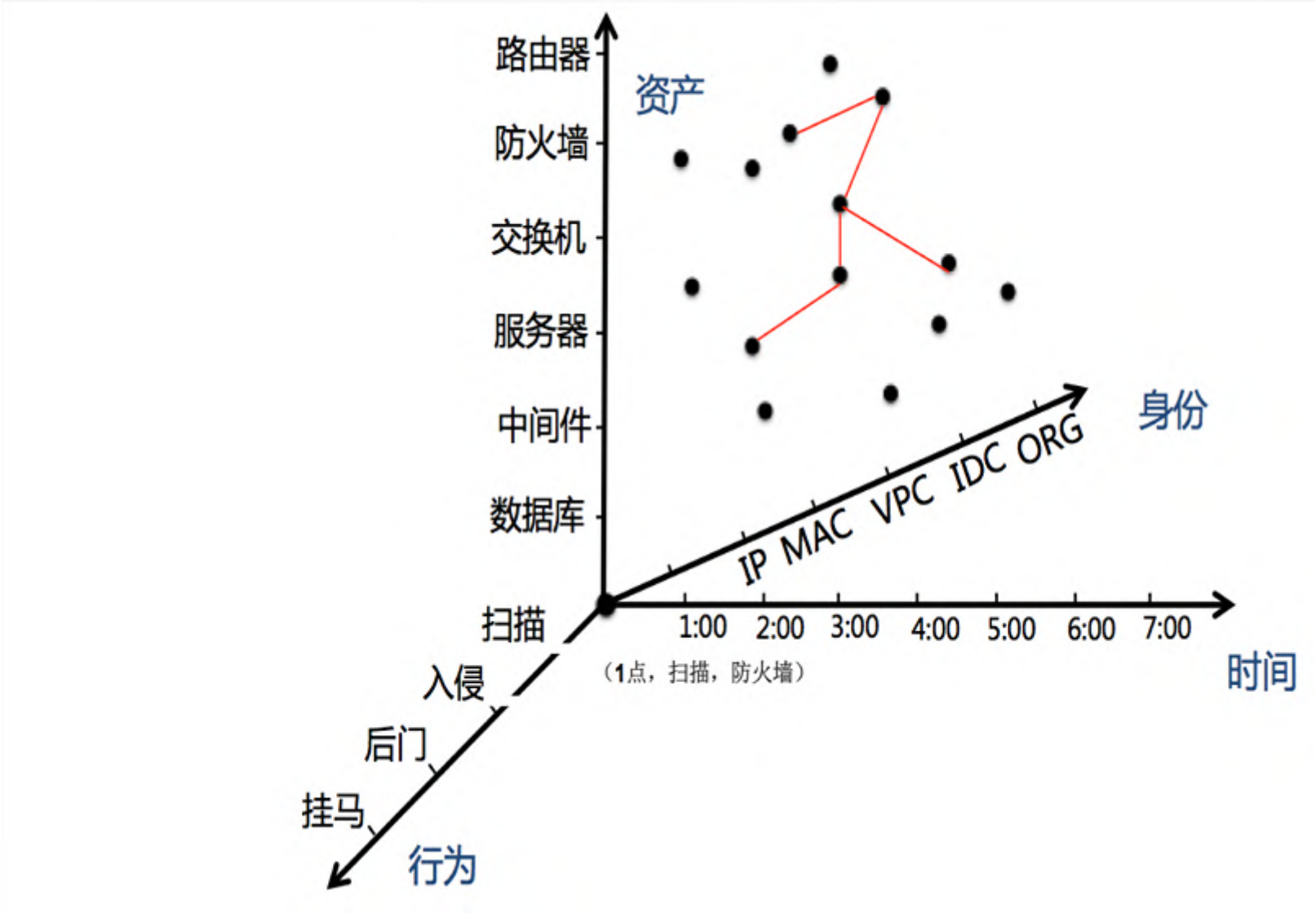
ATTACK TARGETS

#	COUNTRY
75	United States
20	Saudi Arabia
9	France
2	United Arab Emir...

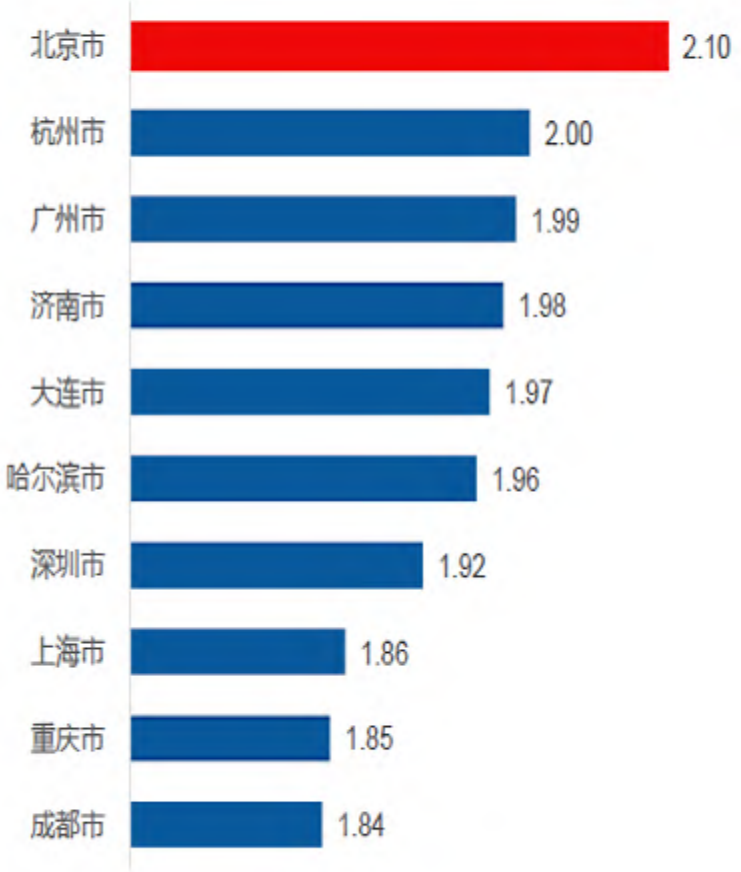
LIVE ATTACKS

TIMESTAMP	ATTACKER	ATTACKER IP	ATT
14:59:19.256	Zhengzhou Giant Computer Network Tech...	122.114.17.100	Zhe
14:59:19.266	Zhengzhou Giant Computer Network Tech...	122.114.17.100	Zhe
14:59:19.272	Zhengzhou Giant Computer Network Tech...	122.114.17.100	Zhe
14:59:19.279	Zhengzhou Giant Computer Network Tech...	122.114.17.100	Zhe
14:59:19.286	Streamyx-Biz-Eastern	218.111.91.121	O1
14:59:19.292	National Computer Systems Co.	46.151.209.24	Riya
14:59:19.530	National Computer Systems Co.	46.151.210.66	Riya
14:59:19.541	National Computer Systems Co.	46.151.215.165	Riya
14:59:19.623	China Unicom Guangxi Province Network	110.73.52.219	Nan
14:59:19.929	Beaver Valley Intermediate Unit #27	206.127.137.2	Ellw

全息安全


路由器
防火墙
交换机
服务器
中间件
数据库
扫描
资产
IP MAC VPC IDC ORG
身份
1:00 2:00 3:00 4:00 5:00 6:00 7:00
时间
（1点，扫描，防火墙）
入侵
后门
挂马
行为

2015年第二季度中国主要城市拥堵排名TOP10



设备故障应急

基于静态规则的安全

網易数码 网易首页 > 数码频道 > 正文 为3亿人新闻阅读而生的客户端

Windows 3.1故障导致巴黎奥利机场短暂关闭

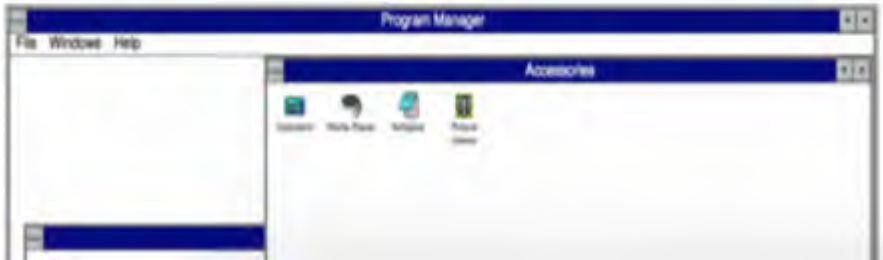
2015-11-15 17:08:40 来源: cnbeta网站(台州)

分享到：

716

上周六，由于一套运行Windows 3.1的关键系统发生故障，迫使巴黎奥利机场因为大雾短暂关闭。Windows 3.1至今已有23年历史。奥利机场运行Windows 3.1的系统叫DECOR，用于在恶劣气候下——比如大雾——向飞行员传送跑道视程信息。它属于必不可少的关键系统，上周六它却停止了工作。

法国空管联盟的 Alexandre Fiacre说，该工具运行在四种古老的操作系统：Windows 3.1、Windows XP和未指定的UNIX系统。这些系统维护不佳，因为它们都有几十年的历史，很难找到合适的人去处理相关问题。



思考 困扰 实践 态度

我们是如何规避困扰的

IDCC

- 集群环境下，硬件故障影响小
- 自行开发，集中维护，维护容易
- 不涉及众多第三方系统
- 我们容易获得系统中的数据，不存在用户不同意上传问题

数据分析

IDCC

- 三岔路口？
- 机器学习？
- 攻击情报？
- 规则？

数据分析

IDCC

- 通过机器学习，发现异常
- 人工标定，分析。产出规则情报
- 规则情报反馈给分析系统，产出更多信息

分析实践

- 发现绕过Web防火墙的攻击行为

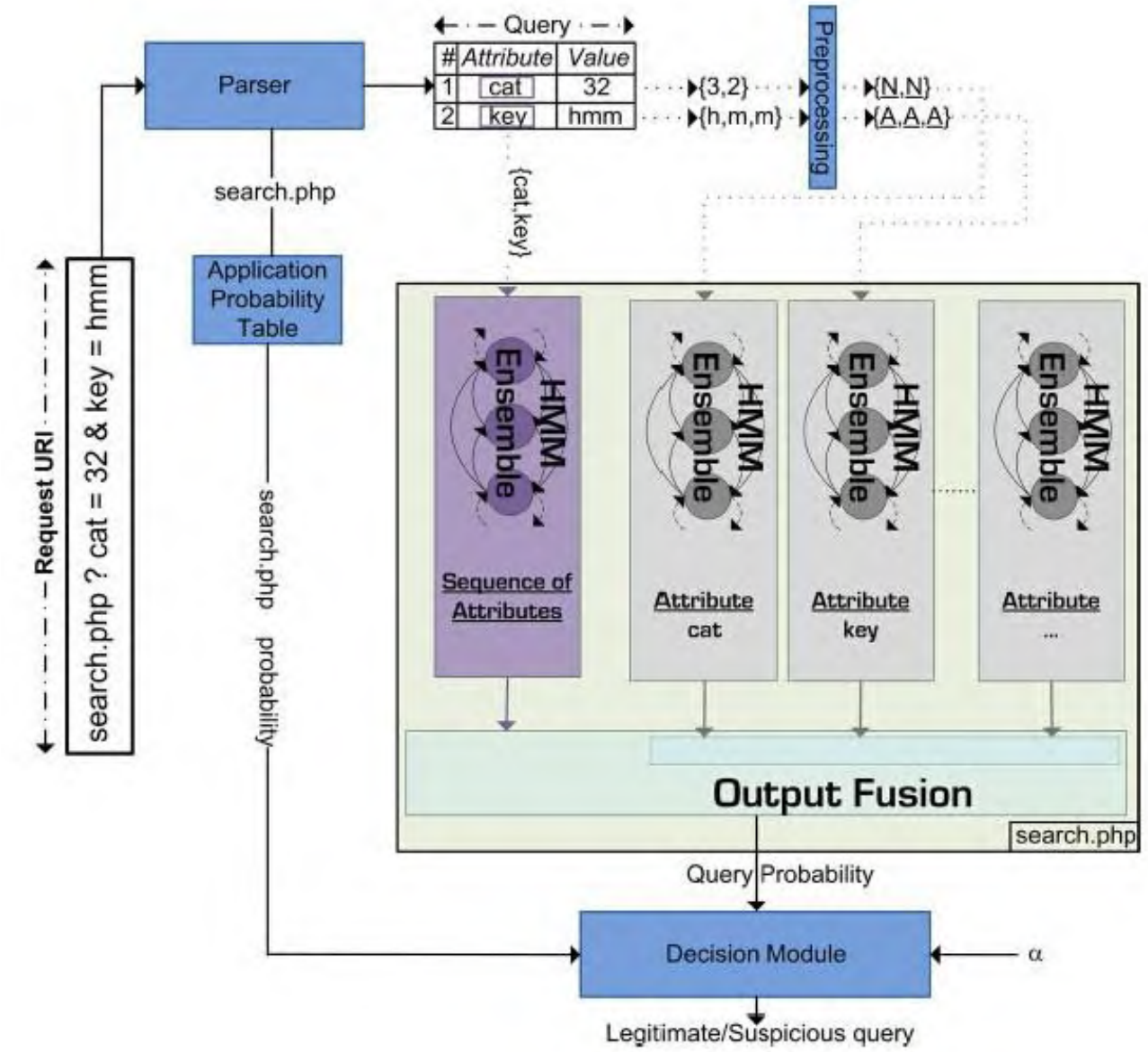
- 提取攻击情报，提取扫描payload供扫描器完善，提取恶意攻击IP供阻断

分析实践

- 分析维度
  HTTP请求的各个角度
  PATH，QUERY，UA，SESSION等

- 分析方法
  基于统计，机器学习，对PATH，QUERY，SESSION等建立模型。包括参数分布，请求频率，SESSION请求宽度，404比例等

Query模型


Query
Attribute Value
1 cat 32
2 key hmm
{3,2}
{h,m,m}
Preprocessing
{N,N}
{A,A,A}
Parser
search.php
{cat,key}
Application Probability Table
Request URI
search.php ? cat = 32 & key = hmm
search.php probability
Ensemble
HMM
Sequence of Attributes
Attribute cat
Attribute key
Attribute ...
Output Fusion
search.php
Query Probability
Decision Module
α
Legitimate/Suspicious query

Path模型

- 图模型

通过referer，构建页面访问图
找到图中的孤立点
对孤立点进行进一步判断

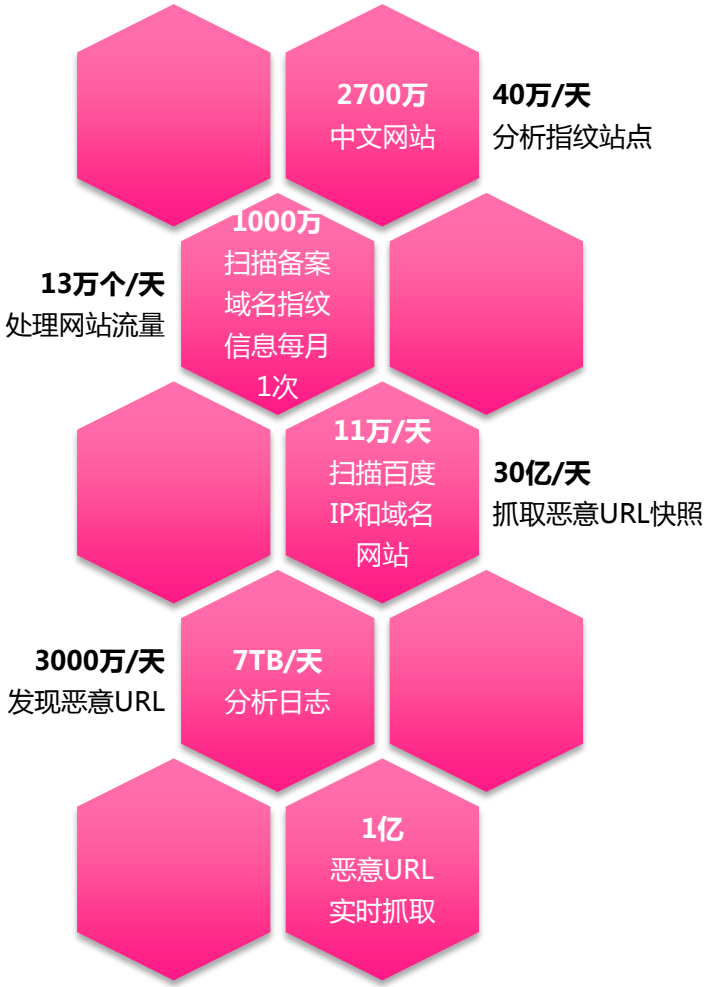
Session模型

- CART（Classification And Regression Tree，分类回归决策树）人机识别模型
  1. Sort & Parser 对原始Web Server Log按时间排序，并进行正则解析
  2. Session 对每个站点的Log进行session重组
  3. Label 对分好的Session进行半自动标记（GoodRobot，BadRobot，Human）
  4. Feature 对分好的Session设计并提取特征
  5. CART Model 对特征训练Tree Model
  6. Classify 使用 Tree Model进行分类

实践效果

运行时间14年10月至今
日输入数据3T左右
产出异常十几兆
发现众多绕过WAF的webshell
增加waf规则10+条
完善waf规则10+条
发现很多有趣的payload（ˇ_ˇ）

如何识别风险



➢ 海量安全数据和情报收集， 处理

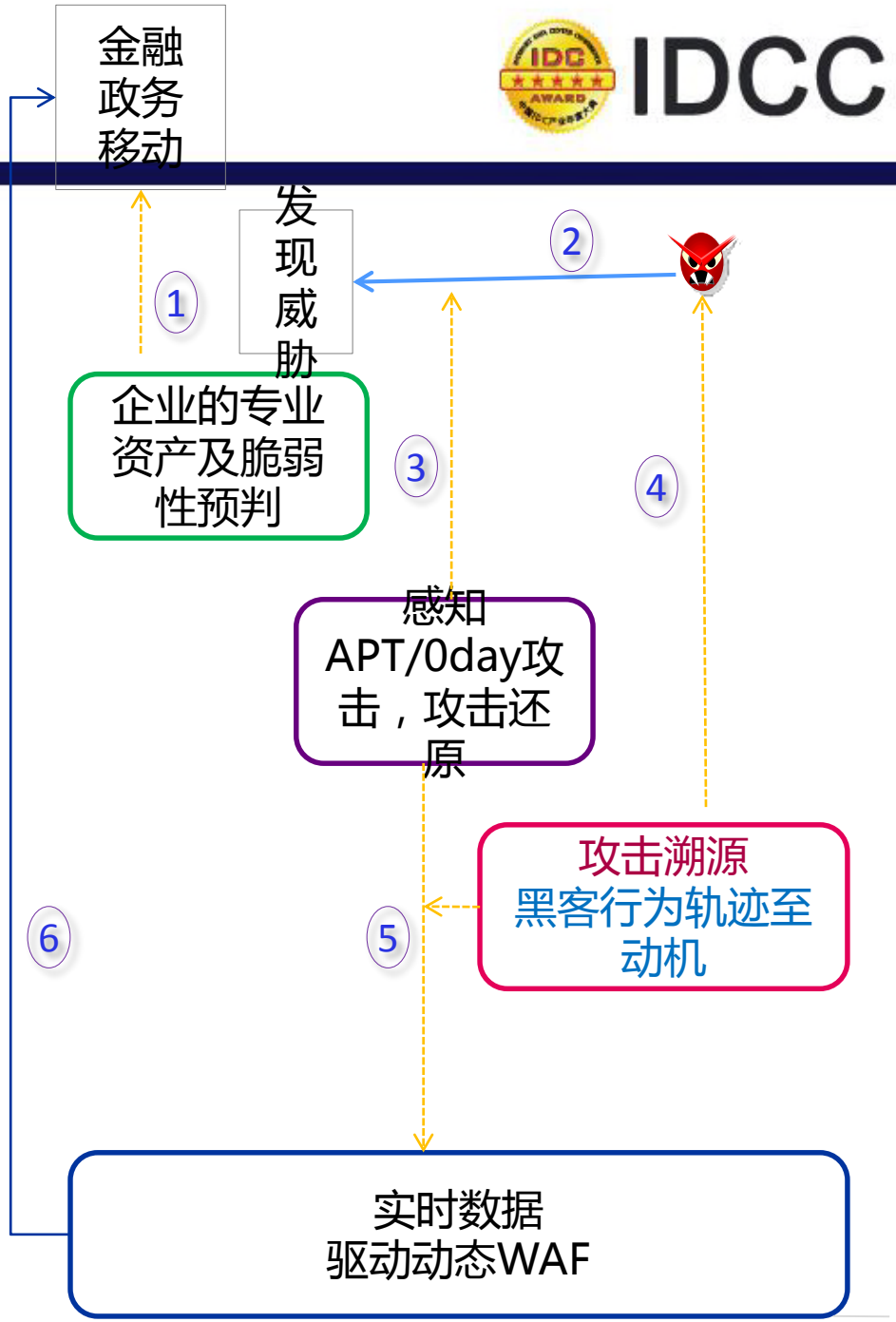
- 高效节能的数据中心： 年PUE低于1.28，日交付1万台能力，年规模采购数万定制服务器，
- 高效专业运维团队：运维数十万台服务器集群，支持百度20+用户过亿的产品规模PB级安全大数据分析系统，支持常见的web、os、数据库20种日志类型
- 规模TB级数据毫秒级响应

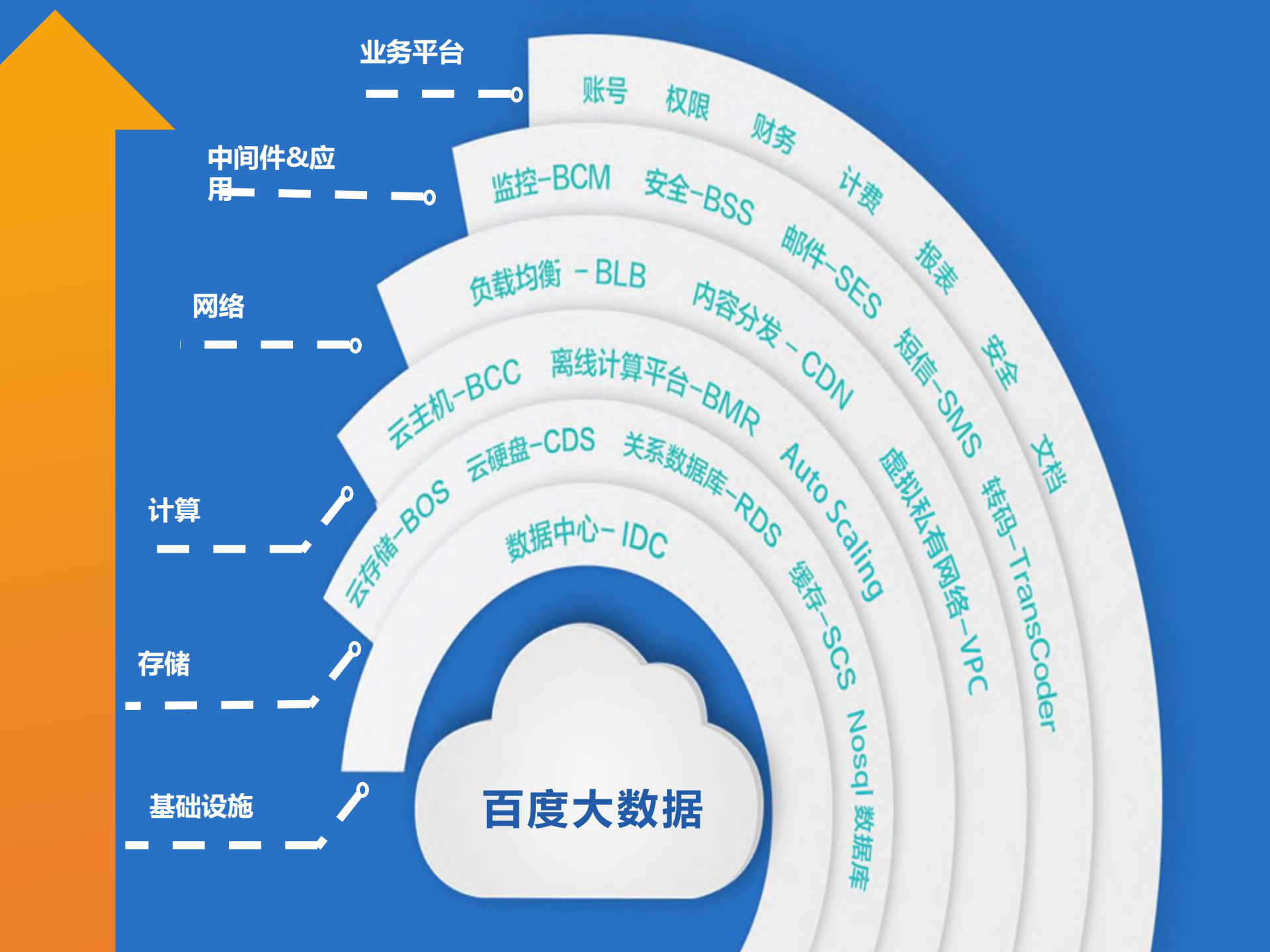
➢ 深度学习和异常检测算法

- 一流的人机识别算法
- 访问行为分析， 异常数据检测 (如 HMM), webshell检测率达90%

全息安全


报警列表
攻击过程
攻击详情分析
2015
发现来自 中国 的攻击者，USERAGENT是Mozill
2015-10-16
12:51:00
攻击者访问了 124.251.46.31 的/UserFiles/Image/
进行了2次 WebShell/菜刀 该攻击状态是 成功 的
2015-10-16
13:36:06
攻击者访问了 124.251.46.31 的/UserFiles/Image/
进行了1次 WebShell/菜刀 该攻击状态是 成功 的


金融
政务
移动
发现威胁
1
2
企业的专业资产及脆弱性预判
3
4
感知APT/0day攻击，攻击还原
攻击溯源
黑客行为轨迹至动机
5
6
实时数据
驱动动态WAF


业务平台
中间件&应用
网络
计算
存储
基础设施
账号
权限
财务
计费
报表
安全
文档
监控-BCM
安全-BSS
邮件-SES
短信-SMS
转码-TransCoder
负载均衡 - BLB
内容分发 - CDN
虚拟私有网络-VPC
云主机-BCC
离线计算平台-BMR
Auto Scaling
云存储-BOS
云硬盘-CDS
关系数据库-RDS
缓存-SCS
Nosql 数据库
数据中心- IDC
百度大数据

IDC与CDN分布

大

数十个IDC，近50万服务器大规模集群部署

快

遍布全国的CDN分发网络，可同时支持动态和静态的全方位加速

智能

不同层级网络出口服务+百度智能流量调度+智能运维

大数据分析解决方案

面对大量的数据，如何高效处理，挖掘其中的业务价值？

通过**数据分析**，发现业务机会
通过**机器学习**，挖掘数据价值

离线数据分析与挖掘

BMR

国内首个云端全托管的Hadoop/Spark服务

机器学习

BML

机器学习算法
超完备的特征库
深度学习

在线数据分析

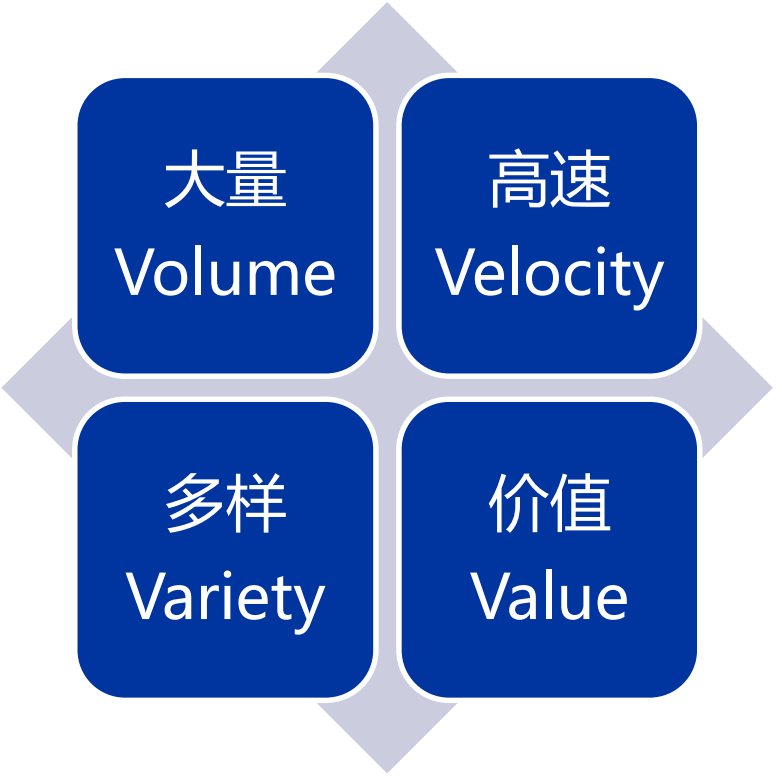
PALO

具备超高性价比，
PB级数据仓库
毫秒级延迟

大数据

IDCC

- 服务网站数近100万
- 日均请求数40亿次
- 日防CC近1亿次
- 对DDOS压制能力达1T
- 全国网站平均提速450%



我们有更多的数据源

审计

IDCC

今日安全状况
396
0
2
历史攻击类型
今日
昨日
7天
30天
应用程序漏洞 17%
文件访问控制 8%
SQL注入攻击 53%
其它 1%
攻击来源Top5（今日）
116.194.208.138 山东 4
221.1.90.154 山东 4
115.199.237.169 浙江 1
218.18.120.250 广东 1
Web攻击拦截详情
攻击IP
攻击URL
攻击类型
111.37.7.69 2014-02-12 18:00:00 详情 1 Remote Database File Access ... 已阻断
115.199.237.169 2014-02-12 18:00:00 详情 1 SQL Injection Attack 已阻断
218.18.120.250 2014-02-12 15:00:00 详情 1 Web Struts2 Vulnerabilities 已阻断
218.94.82.115 2014-02-12 13:00:00 详情 1 DedeCms 已阻断
60.219.138.82 2014-02-12 10:00:00 /favicon.ico 详情 9 File Upload 已阻断
119.120.182.81 2014-02-12 07:00:00 详情 1 File Upload 已阻断
114.241.24.186 2014-02-12 01:00:00 / 详情 7 File Upload 已阻断
54.219.42.81 2014-02-12 00:00:00 详情 2 XSS Attack 已阻断

大数据威胁处理—网站后门识别

IDCC

百度云分析 xi.baidu.com

仪表盘 搜索 报警 日志源 返回旧版 [redacted] 个人中心

数据统计

0个	0个	0个	0个	0个	0个
黑客数	攻击成功	攻击未知事件	日志数量	攻击次数	后门数量

大数据威胁处理

搜索攻击源IP 全部 失败 成功

攻击源IP	攻击目标IP	攻击类型	攻击次数	攻击起始时间	攻击结果	深度分析
120.24.156.38	[redacted]	WebShell/菜刀	3次	开始：2015-10-16 12:38:40 结束：2015-10-16 13:35:09	成功	攻击溯源
112.74.13.199	[redacted]	WebShell/菜刀	1次	开始：2015-10-16 12:21:52 结束：2015-10-16 12:21:52	成功	攻击溯源
223.223.176.194	[redacted]	WebShell/菜刀	2次	开始：2015-10-16 12:58:35 结束：2015-10-16 12:58:39	成功	攻击溯源
61.146.155.145	[redacted]	WebShell/菜刀	10次	开始：2015-10-16 11:20:47 结束：2015-10-16 11:29:37	成功	攻击溯源

攻击过程分析

IDCC

百度云分析
xi.baidu.com

仪表盘 搜索 报警 日志源

报警列表 | 攻击过程 | 攻击详情分析

2015

发现来自 中国 的攻击者，USERAGENT是Mozilla/5.0 (Windows; Windows NT 5.1; en-US) Firefox/3.5.0

2015-10-16 12:51:00

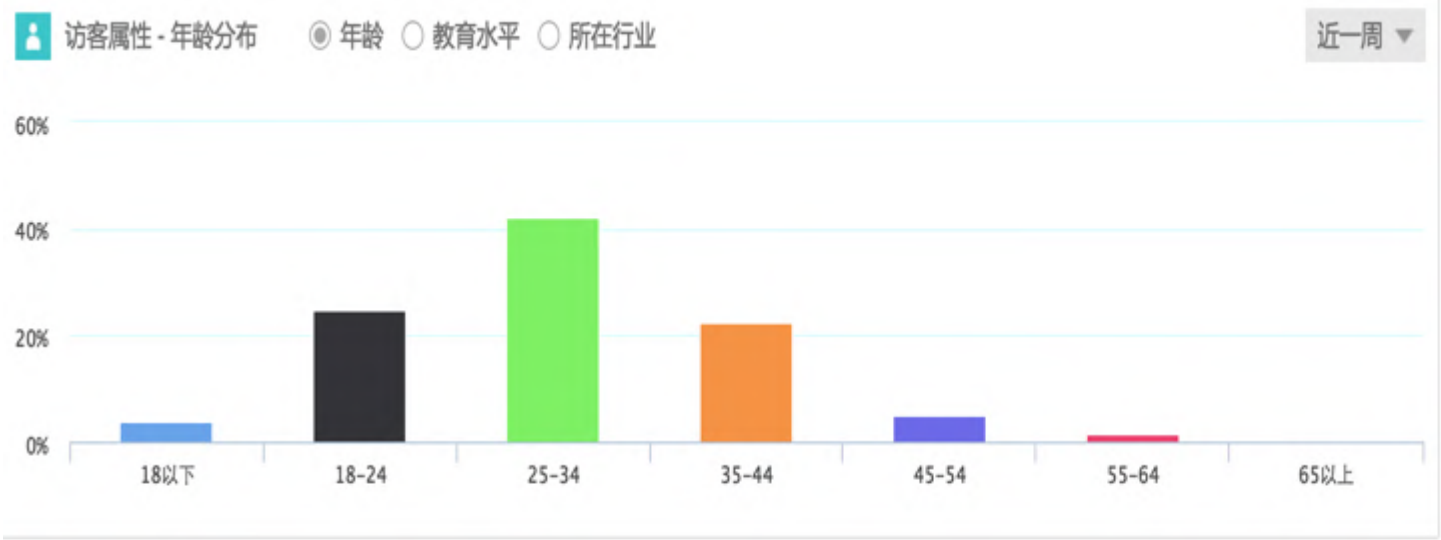
攻击者访问了 124.251.46.31 的//UserFiles/Image/ali.php页面
进行了2次 WebShell/菜刀 该攻击状态是 成功 的

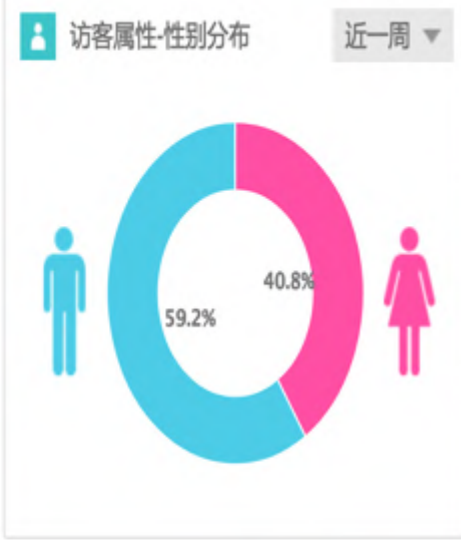
2015-10-16 13:36:06

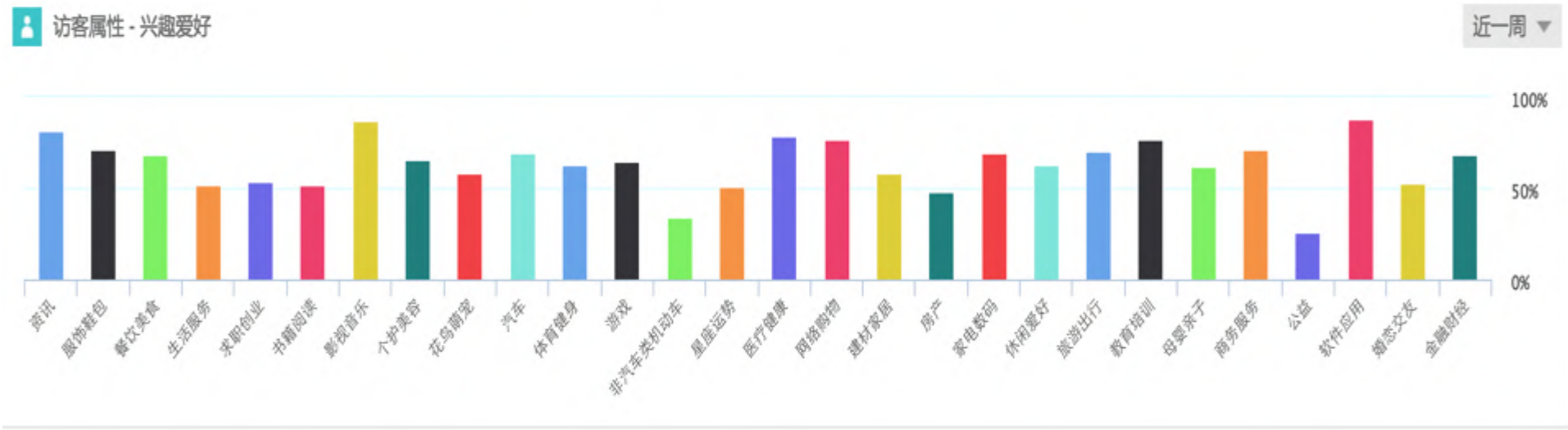
攻击者访问了 124.251.46.31 的//UserFiles/Image/ali.php页面
进行了1次 WebShell/菜刀 该攻击状态是 成功 的


大数据
用户画像
人口属性
性别
年龄
人生阶段
消费水平
受教育程度
行业&职业
车产状况
手机设备
更多
手机品牌
手机型号
手机使用时段
兴趣&关注
汽车
房产
教育培训
医疗健康
游戏
更多
* 等28个垂直行业
细分万级标签
消费场景
需求场景
决策阶段
预测时间
预测位置
旅游
餐饮
汽车
培训
更多
国内游
出国游
港澳台
意向
调研
比价
购买
时间
地点＋概率
信用评分
身份特质
履约能力
行为偏好
人脉关系
信用历史
App行为
卸载过的App-list
安装的App-list
App使用行为
下载、卸载时间
首、末次使用时间
总使用时长
总使用次数
实时位置
详细地址
常驻位置
工作地
工作时段
居住地
在家时段
常访地

业务增值-常规用户画像


访客属性 - 年龄分布
年龄
教育水平
所在行业
近一周
60%
40%
20%
0%
18以下
18-24
25-34
35-44
45-54
55-64
65以上


访客属性-性别分布
近一周
40.8%
59.2%


访客属性 - 兴趣爱好
近一周
100%
50%
0%
资讯
服饰鞋包
餐饮美食
生活服务
求职创业
书籍阅读
影视音乐
个护美容
花鸟萌宠
汽车
体育健身
游戏
非汽车类机动车
星座运势
医疗健康
网络购物
建材家居
房产
家电数码
休闲爱好
旅游出行
教育培训
母婴亲子
商务服务
公益
软件应用
婚恋交友
金融财经

思考 困扰 实践 态度

我是谁？

IDCC

我们是建设者还是使用者？
我需要建立威胁情报系统么？

高大上的方案不一定适合你

人与工具的平衡：

- 自行研发安全工具
- 使用开源安全工具
- 购买商业安全工具
- 采用云服务化工具

云和大数据不解决一切问题

IDCC

云是否适合所有场景？
非大数据的信息富矿我们挖了多少？

高难度的技术不一定带来高价值

IDCC

技术人员的鄙视链
操作系统漏洞挖掘vs安全配置

解决问题最重要

IDCC

人呐，最终要的是开心；
工程师呐，最重要的是解决问题。

Thank you!