网安加社区 | 网安加·百家讲坛

www.cwasp.cn

UEBA的
基础知识与实际应用

演讲人：王诗涵　2024.11.21

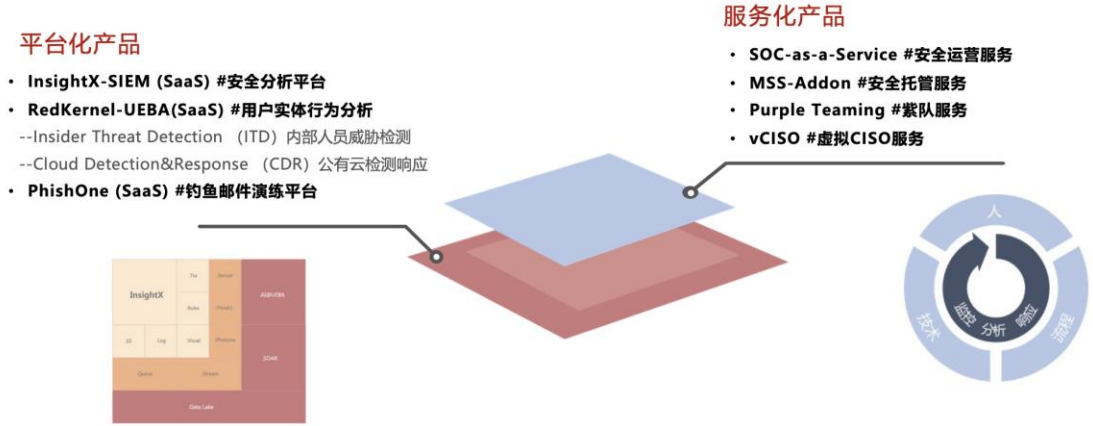
云纷科技
joy.wang@cloudfall.cn

面向全球的下一代安全托管服务商

网安加社区 网安加·百家讲坛

服务：自成立以来云纷科技就把成为交付“监测-分析-响应”完整能力闭环的下一代MSSP作为使命。将“安全运营中心（SOC）”分层分解，以更灵活的模块化方式交付“下一代安全运营服务”；高效整合人、技术和流程。

技术：基于自研的云原生和人工智能的——InsightX SIEM 和 RedKernel Cloud UEBA平台，为不同的企业客户提供符合现状和中远期发展的安全运营能力。在为客户建立运营体系的同时，以“实践+数据+AI”重塑安全运营链路，最终帮助用户清晰定位威胁、降低风险，提高整体安全水平。



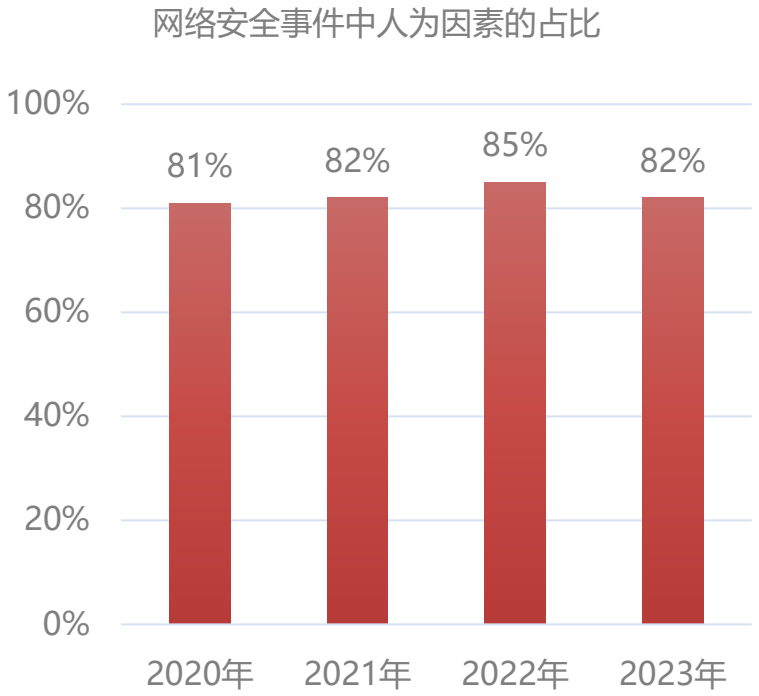
CONTENTS

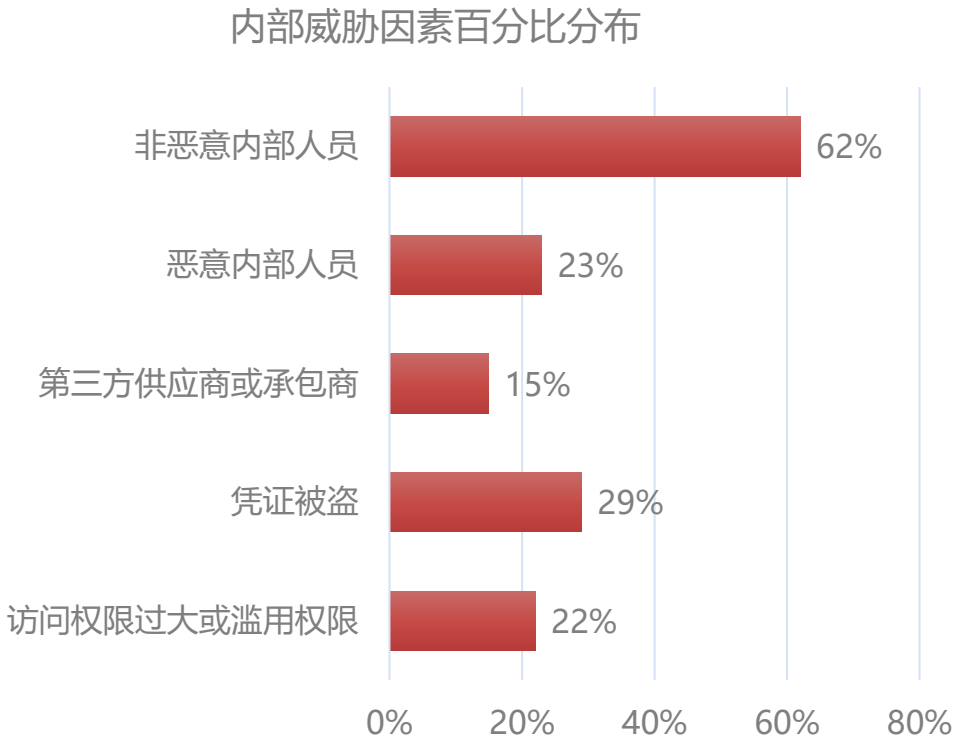
目录

01

企业内部安全威胁的定义

内部威胁的比例





Cyber Magazine 2024 《82% of all cyberattacks involve the human element》
Verizon 《2023 data breach investigations report-DBIR》

凭证被盗：外部攻击者通过获取内部人员的合法凭证伪装成他们，进行非法操作。

内部威胁的定义

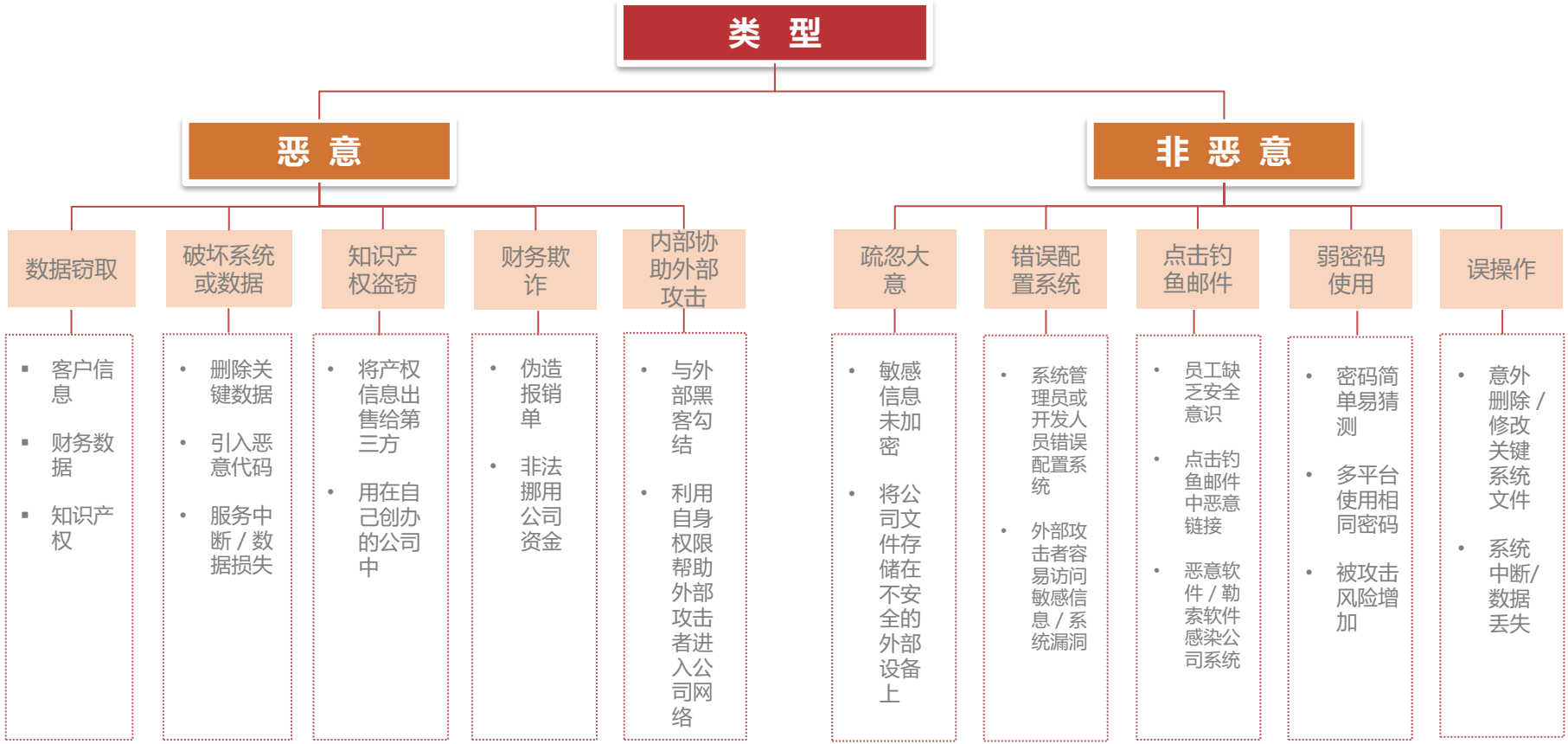
定义:

- 由具有内部访问权限的人员（如员工、合同工、供应商等）对组织的安全、系统、数据或业务运营造成的威胁。
- 这类威胁可能来自恶意行为（有意破坏、泄露数据等）或无意的错误（如操作失误或疏忽）。
- 与外部攻击者不同，内部威胁者具有合法的访问权限，并且对组织的系统和数据有一定的了解，因此他们能够更容易地避开常规的安全防护措施。

影响:

数据泄露	财务损失	系统中断	知识产权/商业机密泄露	法律和合规性风险	声誉损害
•身份信息（PII） •财务信息 •商业机密 •其他敏感信息	•财务欺诈 •挪用资金 •盗窃公司资产 •破坏系统	•删除关键数据 •关闭系统或 •植入恶意软件 •系统宕机 •业务中断	•专有技术 •产品设计 •生产流程 •企业失去市场竞争力 •长期发展受阻	•数据保护法、 •行业监管要求 •其他法律法规 •公司面临诉讼、巨额罚款和合规性问题	•客户信任下降 •股价下跌 •公众形象严重损害

内部威胁的类型


类 型
恶 意
非 恶 意
数据窃取
破坏系统或数据
知识产权盗窃
财务欺诈
内部协助外部攻击
疏忽大意
错误配置系统
点击钓鱼邮件
弱密码使用
误操作
客户信息
财务数据
知识产权
删除关键数据
引入恶意代码
服务中断 / 数据损失
将产权信息出售给第三方
用在自己创办的公司中
伪造报销单
非法挪用公司资金
与外部黑客勾结
利用自身权限帮助外部攻击者进入公司网络
敏感信息未加密
将公司文件存储在不安全的外部设备上
系统管理员或开发人员错误配置系统
外部攻击者容易访问敏感信息 / 系统漏洞
员工缺乏安全意识
点击钓鱼邮件中恶意链接
恶意软件 / 勒索软件感染公司系统
密码简单易猜测
多平台使用相同密码
被攻击风险增加
意外删除 / 修改关键系统文件
系统中断/数据丢失

内部威胁方法论

网安加社区 网安加·百家讲坛

MITRE Insider Threat

Insider Threat TTP Knowledge Base

MITRE ATT&CK – UEBA by Cloudfall

ATT&CK	Techniques	Descriptions
Initial Access	Valid Accounts	检测异常的账户使用行为，如在非常规时间、从未知设备或异常地理位置进行的访问
	External Remote Services	检测对外部远程服务的异常访问，如VPN、RDP等
Execution	Command and Scripting Interpreter	检测异常的命令行或脚本执行行为，如普通用户突然频繁运行PowerShell或Python脚本
	User Execution	检测异常的用户执行行为，如运行了非常规的程序或文件
Persistence	Account Manipulation	检测异常的账户操作行为，如在非工作时间创建新用户或修改用户权限
Privilege Escalation	Abuse Elevation Control Mechanism	检测异常的提权行为，如普通用户突然开始频繁使用sudo或运行管理工具
	Access Token Manipulation	检测异常的访问令牌使用，如一个用户的令牌突然被用于访问许多不同的资源
Credential Access	Brute Force	检测异常的登录尝试行为，如短时间内大量失败的登录尝试
	Credentials from Password Stores	检测异常的密码库访问行为，如一个用户突然访问了许多其他用户的凭据
Discovery	Account Discovery	检测异常的账户枚举行为，如一个用户突然开始查询大量的Active Directory账户信息
	Permission Groups Discovery	检测异常的用户或组权限查询行为，如一个普通用户突然开始查询管理员组的成员
	Query Registry	检测异常的注册表访问行为，如一个用户突然开始查询关键的系统配置信息
Lateral Movement	Remote Services	检测异常的远程服务使用行为，如一个用户突然频繁使用SSH或RDP在不同系统之间移动
	Taint Shared Content	检测异常的文件共享行为，如一个用户突然修改了许多共享文件
	Remote Desktop Protocol	检测异常的RDP使用行为，如在非工作时间或从未知设备发起的RDP连接
	Windows Admin Shares	检测异常的管理共享访问行为,如一个普通用户突然开始访问其他系统上的共享
Collection	Data from Local System	检测异常的数据收集行为，如一个用户突然开始访问和收集大量本地文件
	Data from Network Shared Drive	检测异常的网络驱动器访问行为，如一个用户突然开始下载大量文件
Exfiltration	Automated Exfiltration	检测异常的数据传输行为，如在非常规时间或通过异常协议自动传输大量数据
	Data Transfer Size Limits	检测异常的大文件传输行为，如一个用户突然开始传输超大型文件
	Scheduled Transfer	检测异常的数据传输调度行为，如在非工作时间或以固定间隔自动传输数据
	Transfer Data to Cloud Account	检测异常的云存储使用行为，如用户突然开始将大量数据上传到个人云存储账户
	Exfiltration Over Alternative Protocol	检测异常的网络协议使用行为，如使用非标准端口或协议传输大量数据
More	.....	.....

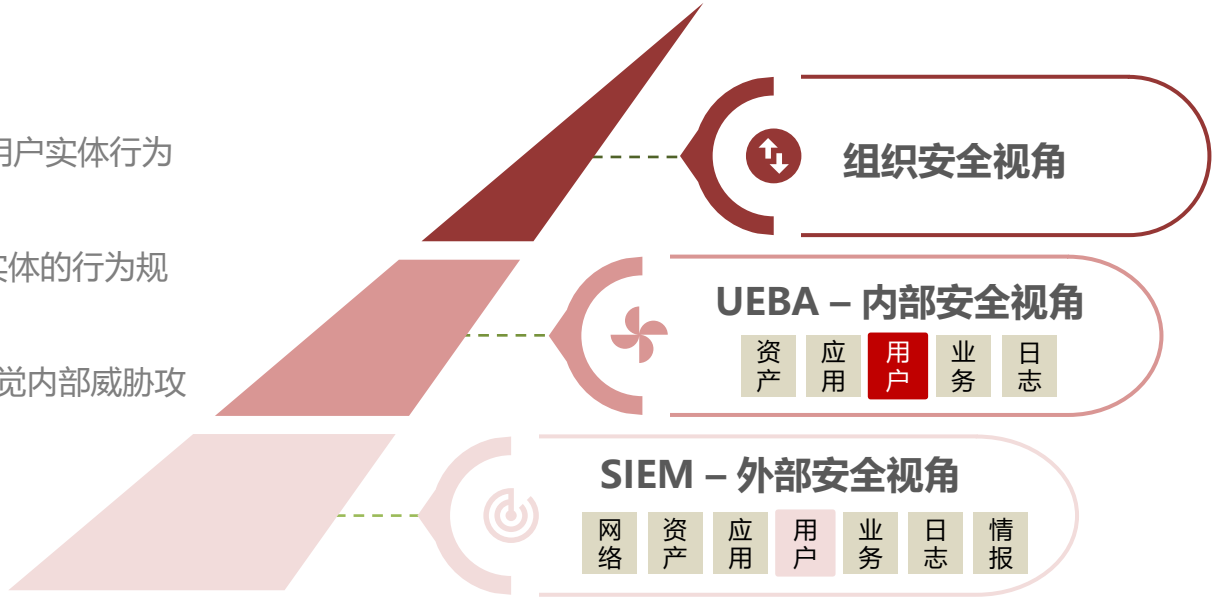
02

认识UEBA

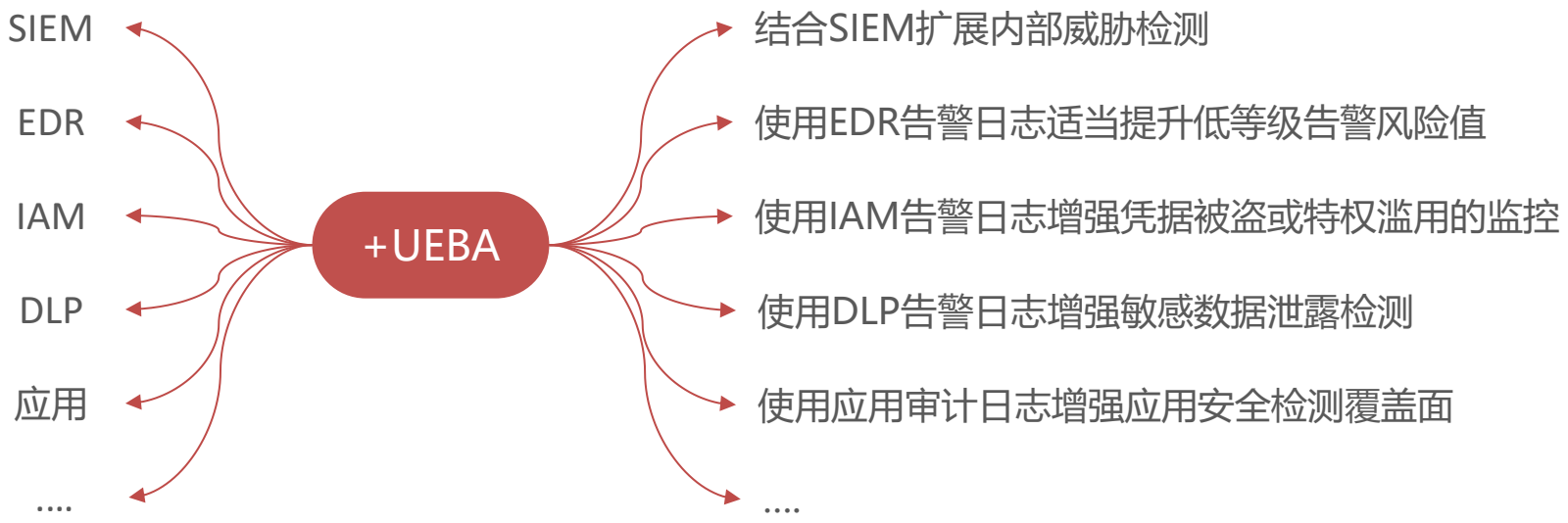
WHAT IS UEBA?

定义：

- User and Entity Behavior Analytics 用户实体行为分析
- 通过机器学习，统计学方法寻找人类/实体的行为规律
- 提取可能导致威胁的异常行为，从而发觉内部威胁攻击等



UEBA vs SIEM及其他安全工具



- UEBA可以与SIEM, EDR, IAM等产品结合，达到企业抗网络威胁能力的最大化
- 补充员工风险监控能力
- 但不是他们的替代品

HOW UEBA WORKS?

数据收集

收集各种日志，如身份认证日志、DLP日志、数据库审计日志、应用程序日志等。日志中需包含 时间、用户和行为的详细信息。

数据预处理

对收集的数据进行预处理，包括清洗、解析和标准化，确保数据的一致性和准确性。以便后续分析使用。

特征提取

数据预处理后，从数据中提取各种特征，这些特征可以用来描述用户和实体的行为模式。如用户登录次数、是否为工作日、访问的设备类型、地点等。

建立基线

使用机器学习和统计分析等技术，学习历史数据，识别正常行为模式和异常行为模式。基于历史数据训练出模型，建立基线。

异常检测

将新观察到的行为输入模型进行比较，以检测异常行为。

事件响应&分析

检测出异常行为 累加相应用户/实体的风险值并排序。

Email Activity

+10

Logon Activity

+20

File Activity

UEBA常见应用场景

登录地点、频率、时间异常

文件下载、删除、外发频率异常

账号共享/凭据泄露

操作数据库行为异常

……

安装渗透工具、使用高危命令

账号安全

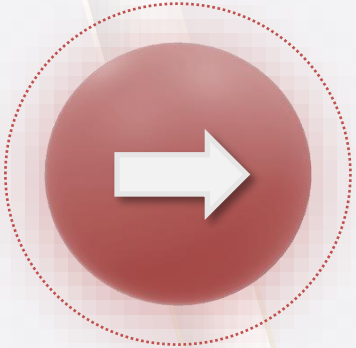
数据安全

设备安全

场景（一）：恶意员工导致的数据泄露

日志类型

- DLP日志
- 文件服务器日志
- 网盘日志
- OA日志
- 上网行为日志
- 应用日志



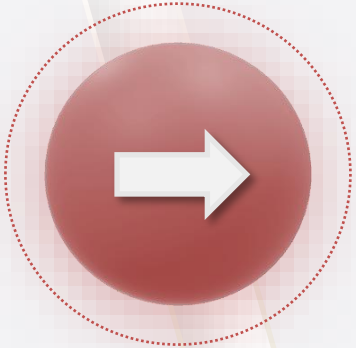
异常行为

- 用户访问文件时间偏离部门基线文件服务器日志
- 用户访问文件频次偏离部门基线
- 文件下载数量异常
- 邮件、打印、USB文件数量异常
- 邮件、打印、USB等多通道外发同一文件

场景（二）：账号被盗导致的系统安全

日志类型

- PAM
- 堡垒机
- DB审计
- …



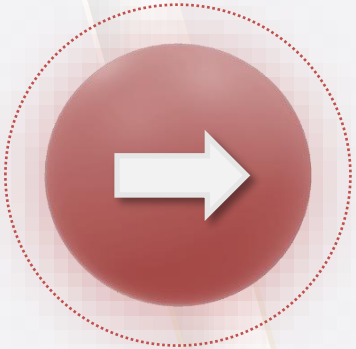
异常行为

- 用户登录服务器时间偏离个人、部门基线
- 用户安装、使用渗透工具
- 用户执行高危命令
- 用户操作数据库行为偏离个人基线
- …

场景（三）：离职员工引起的数据泄露

日志类型

- DLP
- EDR
- 上网管理日志
- 网盘日志
- 文件服务器日志
- OA日志
- 应用日志
- …



异常行为

- 工作时间异常
- 访问多个招聘类网站
- 频繁访问大量文件
- 文件下载数量异常
- 打印、USB外发文件数量异常
- …

03

安全案例与技术实现

UEBA落地的挑战

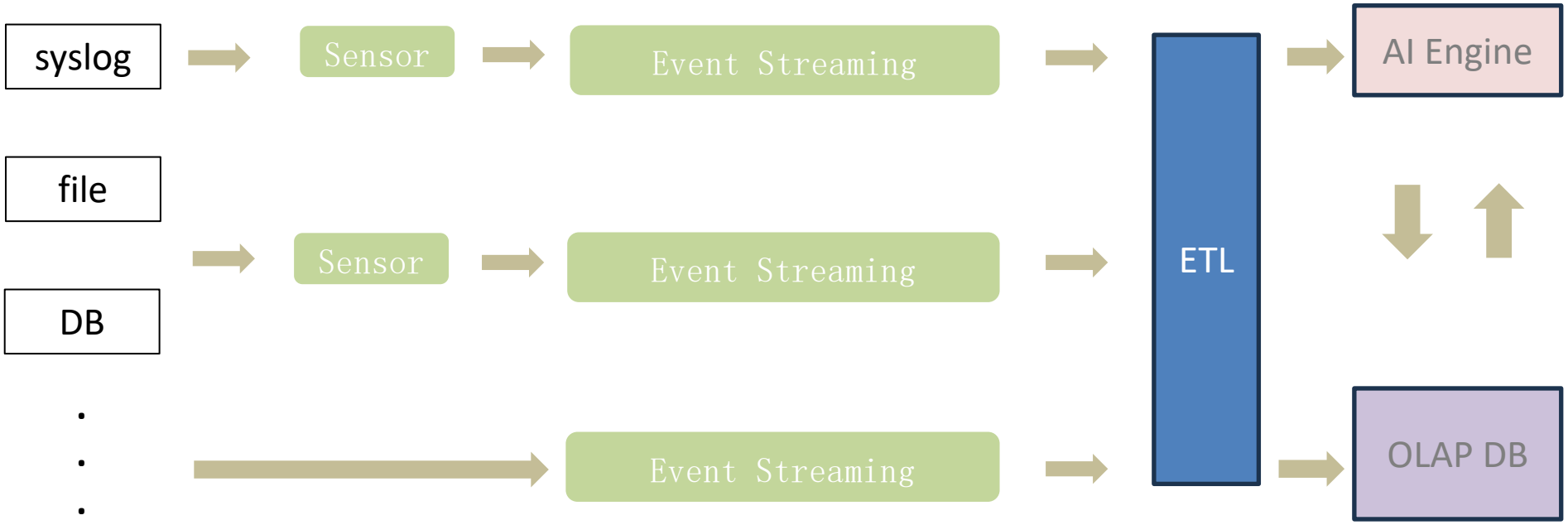
- 大多数企业并没有很好的使用和运营SIEM解决方案，缺乏数据基础、检测体系、专家知识等相关条件
- 企业用户可能发现，即使对于单一场景和用例，UEBA部署也可能比厂商承诺的时间和劳动强度更大。 此外，添加自定义用例可能是一个艰巨的过程，需要诸如数据科学和数据分析之类的专业知识。
- UEBA主要依靠诸如机器学习之类的高级分析方法，但是产品厂商的炒作和诸如“人工智能”之类的术语的使用使企业很难有效地评估供应商的技术和能力。

内部威胁数据源

用户行为视角

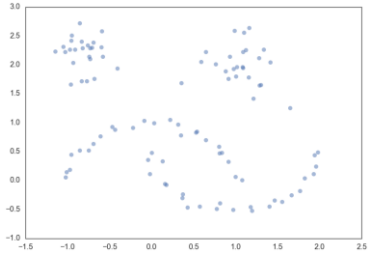
大数据与算法能力

大数据与算法


syslog
file
DB
Sensor
Sensor
Event Streaming
Event Streaming
Event Streaming
ETL
AI Engine
OLAP DB

大数据与算法

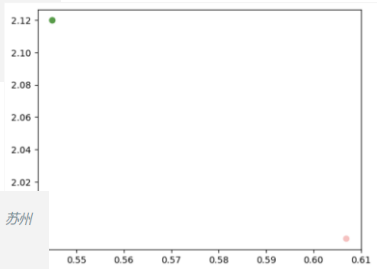
登录地点偏离个人基线

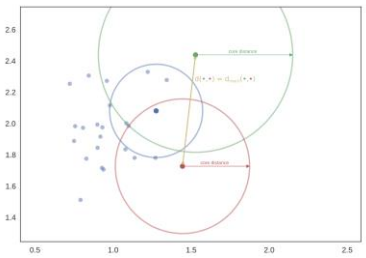


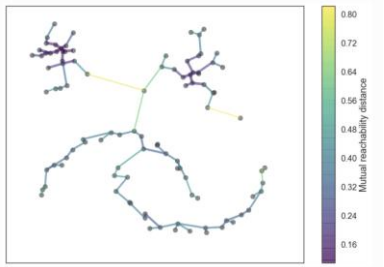
Mutual reachability distance

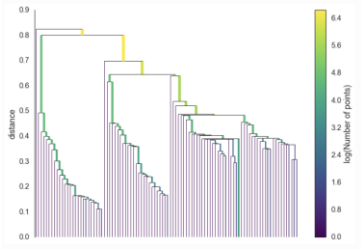
$$d_{\text{mreach}-k}(a,b) = \max\{\text{core}_k(a), \text{core}_k(b), d(a,b)\}$$

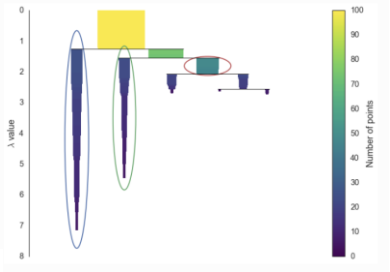
```
# 示例经纬度数据（角度表示）
latitude = np.array([31.2222,31.2222,31.2222,31.2222,31.2222,34.7732])   # 上海 上海 上海 上海 上海 深圳
longitude = np.array([121.4581,121.4581,121.4581,121.4581,121.4581,113.722])

# 将角度转换为弧度
latitude_rad = np.radians(latitude)
longitude_rad = np.radians(longitude)

# 将经纬度组合成一个二维数组，供HDBSCAN使用
coords = np.vstack((latitude_rad, longitude_rad)).T

# 使用Haversine度量进行HDBSCAN聚类
clusterer = HDBSCAN(metric='haversine', min_cluster_size=2, allow_single_cluster=True).fit(coords)
```



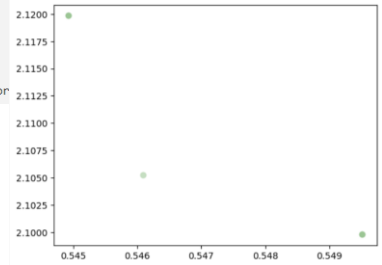






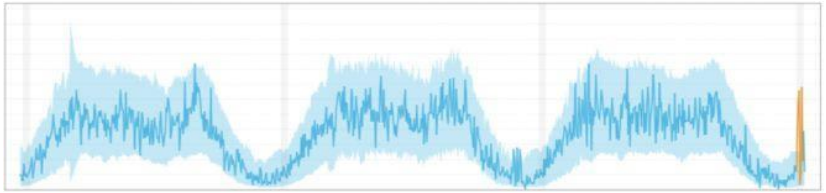


```
# 示例经纬度数据（角度表示）
latitude = np.array([31.484881, 31.2222,31.2222,31.484881, 31.289013])   # 无锡 上海 上海 无锡 苏州
longitude = np.array([120.308480,121.4581,121.4581,120.308480, 120.620109])

# 将角度转换为弧度
latitude_rad = np.radians(latitude)
longitude_rad = np.radians(longitude)

# 将经纬度组合成一个二维数组，供HDBSCAN使用
coords = np.vstack((latitude_rad, longitude_rad)).T

# 使用Haversine metric 进行HDBSCAN聚类
clusterer = HDBSCAN(metric='haversine', min_cluster_size=2, allow_single_cluster=True).fit(coor
```



大数据与算法

文件下载、删除、外发频率异常



逻辑配置

算法*

Frequency Anomaly Individual

算法参数*

```
{
  "window_size": {
    "num": 1,
    "interval": "Hour"
  },
  "predict_config": {
    "predict_type": "upper",
    "max_yhat_lower": 1,
    "min_yhat_lower": 1,
    "lower_threshold": 0.5,
    "upper_threshold": 1.5
  }
}
```

算法参数模板

请输入 算法参数模板 名称查询

参数预览

Frequency Anomaly Individual AI entities y yhat yhat_lower yhat_upper freq event_time contexts window_end

Algorithm for detecting actions with frequency anomaly for single entity 频次异常检测基线

window_size 选填

参数类型: dict

默认取值:

```
{
  "num": 10,
  "interval": "Hour"
}
```

子参数:

num 必填 参数类型: int 默认取值: 1 子参数: No Data

interval 必填 参数类型: str 默认取值: Hour 子参数: No Data

predict_config 选填

参数类型: dict 默认取值: No Data 子参数:

predict_type 选填 参数类型: str 默认取值: upper 子参数: No Data

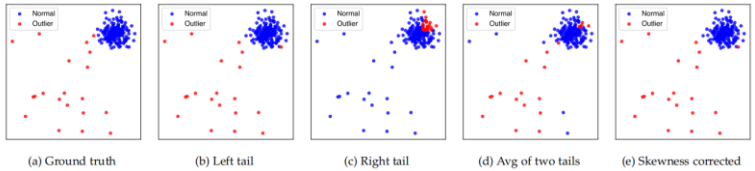
max_yhat_lower 选填 参数类型: int 默认取值: No Data 子参数: No Data

min_yhat_upper 选填 参数类型: int 默认取值: No Data 子参数: No Data

lower_threshold 选填 参数类型: float 默认取值: 0.5 子参数: No Data

upper_threshold 选填 参数类型: float 默认取值: 1.5 子参数: No Data

文件下载、删除、外发时间异常



逻辑配置

算法*

Time Anomaly Individual

算法参数*

```
{
  "max_outlier_ratio": 0.1
}
```

算法参数模板

请输入 算法参数模板 名称查询

参数预览

Time Anomaly Individual AI entities hour_ day_of_week outlier_score max_outlier_ratio cnt event_time contexts window_end

Algorithm for detecting actions with time anomaly for single entity 时间异常检测基线

max_outlier_ratio 必填 参数类型: float 默认取值: 0.1 子参数: No Data

案例（一）：账号安全/凭据泄露事件

事件背景：某科技企业通过VPN为远程员工提供安全的办公网络连接，确保员工在不同城市都能安全访问公司内网。VPN账号的安全性成为该企业信息安全管理的重点。

调查结果：在企业的UEBA建设中，安全团队发现多起VPN账号共享事件。经过调查，确认了共享行为。为应对风险，企业重置了相关账号，要求员工重新设置密码，并加强账号管理。企业还组织了全员信息安全培训，重申VPN账号管理的重要性和一人一账号的原则，警告违规行为的严重后果。

案例（二）：敏感数据泄露的违规行为

事件背景：某全球零售企业处理大量敏感数据（客户信息、供应链数据、财务报表等）。为防止数据泄露，企业部署了DLP系统管理敏感数据外发行为。

异常调查：在UEBA建设中，发现用户违规使用网盘。上线常规策略后，安全团队注意到某用户外发文件频率异常，审查发现该用户频繁上传文件至网盘。由于DLP策略基于字符串匹配，无法及时应对新变种网盘网址，导致部分违规行为未被拦截。

调查结果：确认违规后，企业对该员工进行约谈，并组织全员信息安全培训，强调敏感数据泄露和违规行为的严重后果。同时，调整UEBA策略，加强对数据外发的监控，并优化DLP策略，提升管控效能。

案例（三）：暴力破解事件复盘

事件背景：某大型零售企业发现一起低频率、时间分散的暴力破解事件，攻击模式隐蔽，难以察觉。通过UEBA累积用户风险值，安全团队迅速定位到被暴力破解的账号。

调查结果：经调查，企业不存在名为“test”的账号。安全团队进一步识别了VPN的风险，并推广双因子验证以提升账号安全性，同时封禁了相关恶意IP，确保环境安全。

80分 **test VPN-恶意ip尝试登录**

2024-08-19 05:15:21.777000

VPN账户：test被恶意IP地址：146.19.78.222 Location：New York US PureVoltage Hosting Inc.尝试登录中

80分 **test VPN-恶意ip尝试登录**

2024-07-28 22:19:59.257000

VPN账户：test被恶意IP地址：172.202.177.113 Location：San Antonio US Microsoft尝试登录中

80分 **test VPN-恶意ip尝试登录**

2024-07-04 11:26:25.032000

VPN账户：test被恶意IP地址：172.212.61.226 Location：Boydton US Microsoft尝试登录中

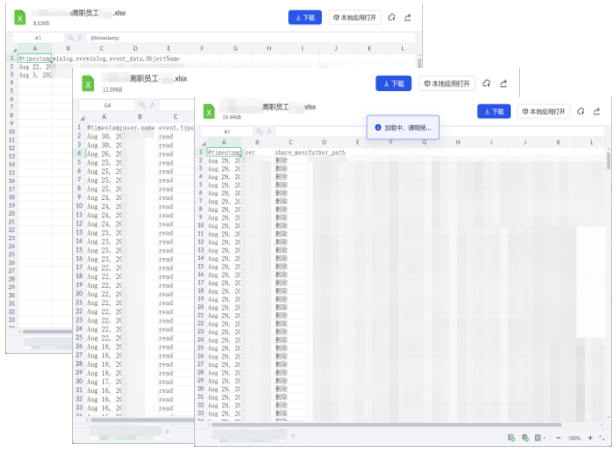
案例（四）：待离职员工审计

网安加社区 网安加·百家讲坛

事件背景：某企业为了保障信息安全，对于已提交离职申请的员工，企业会对其过去一个月的行为进行审查。
确保离职员工在最后的工作阶段没有进行任何违规操作或敏感数据的异常处理。

然而，手动审查离职员工行为复杂且耗时，更早的异常行为无法被洞察。

解决方案：在建设UEBA系统后，只需对该员工进行搜索便可快速查询出该员工的历史全部异常行为，帮助审查人员快速分析是否存在安全风险。



04

最佳实践

明确需求

1. 评估业务需求：确定要解决的具体问题。比如，数据泄露、账号安全。
2. 识别资产和系统：考虑企业的规模和复杂性
3. 评估风险：基于角色的评估，对整个组织的安全风险进行全面的了解和管理。

Asset Register资产登记表						
Organisation 组织		公司名/部门名 ABC/IT				
Asset 资产	Type 类型	Function 功能职能	Owner 责任人	Critical Y/N 关键与否	Strategic Threat actors 战略威胁者	Comments 备注
IT 服务器	物理	支持公司 IT 基础设施功能	Infra head	Y	商业间谍/内部风险	
电脑/电话	物理	办公设备	Infra head	N	商业间谍/内部风险	
NOC平台	IT	支持IT基础的健康程度	Infra head	N-系统操作简单，只需最低限度的培训，	内部风险	
CRM系统	IT	业务系统	Sales Dept	Y-保存了所有客户的详细信息（包括付款时间表/折扣、联系方式）	商业间谍/内部风险	
.....						
.....						

Asset/Team 资产/团队		ABC系统					
Threat actors威胁者		描述：					
Insider activity 内部威胁行为	Insider risk 内部风险	Likelihood (1-5) 可能性	Assumptions 假设	Impact (1-5) 影响度	Assumptions 假设	Identified roles 识别角色	Priority 优先级
1.未授权披露							
	1a)员工将客户数据库发送给竞争对手(国内)	4	数据库访问未划分区域，员工未接受 IT 培训或安全意识，以前发生过但未采取补救措施，当前没有 IT 措施来审核数据库使用情况	2	声誉受损并可能失去客户，但考虑到该产品的营销具有竞争力，因此不会产生持久影响	整个销售团队、公司总监、IT 支持、运营总监	2
	1b)员工将客户数据库发送给竞争对手(海外)	3	同上	3	同上，担心海外市场正在开发自己的流程，然后为他们提供潜在客户信息	海外销售团队、公司和运营主管、IT 支持和研究团队	1
.....	.....						

明确内部威胁数据源

场景	数据源	描述
账号安全	VPN日志	包括VPN连接日志、VPN认证日志等
	特权账号管理系统(PAM)日志	特权账号管理系统(PAM)日志
	堡垒机日志	记录通过堡垒机访问内部系统的操作日志
	活动目录(AD)日志	记录域账号的创建、修改、删除、登录等操作
	LDAP目录服务日志	记录对LDAP的账号查询、认证等请求
	数据库审计日志	记录对数据库系统的账号登录、权限变更等操作
	单点登录(SSO)系统日志	记录用户单点登录各应用的情况
数据安全	DLP(数据泄露防护)系统日志	记录数据传输、复制等行为
	文件系统审计日志	记录文件读写、删除、重命名等操作
	数据库审计日志	记录对数据库数据的查询、修改、导入导出等
	上网行为审计日志	记录用户上网访问的详细情况
	云存储访问日志	记录对云端数据的各种操作
	邮件系统日志	记录邮件发送、接收、附件下载等信息
	应用埋点日志	记录用户在应用上对数据的查询、下载等操作
设备安全	EDR日志	记录终端登录、程序执行、文件访问、网络连接等信息
	防病毒软件告警日志	记录病毒扫描结果、发现的恶意程序信息、查杀和隔离操作记录
	应用日志	记录程序运行情况、办公软件使用情况
	Windows/Linux系统审计日志	记录安全事件、对象访问等
	设备管理工具日志	记录软件安装、卸载等情况
More....	.....	....

选择合适的UEBA解决方案

技术能力

1. 异常检测：机器学习和数据挖掘技术，适应不断变化的用户行为模式
2. 数据处理：快速处理和分析海量的用户行为数据
3. 集成兼容：与现有系统集成，灵活异构

功能特性

1. 风险模型：灵活定义风险模型，算法+规则组合分析
2. 用户视角：以用户为对象累加风险值并排序
3. 可视化：直观的可视化界面展示用户行为和异常情况

运营能力

1. 安全知识：了解网络安全、信息安全、内部风险等方面的知识
2. 数据科学：能够熟练处理和分析大量的用户行为数据，运用机器学习建模
3. 行业经验：具备安全运营和设计内部风险检测模型的实践经验

05

未来趋势

发展方向

智能化

采用更先进的机器学习算法和深度学习技术，自动识别不同用户在不同场景下的行为特征

自动化

系统能自动识别和整合多源数据，减少人工干预

深度融合

与SIEM、零信任架构、SASE等其他安全技术深度融合

更广泛的应用场景

更加深入运用到云安全、工业互联网领域、医疗行业、金融行业等

大模型安全

企业内部大模型访问模式的分析、权限使用检测等

网安加社区

www.cwasp.cn

云纷科技
joy.wang@cloudfall.cn