AI安全实践：探索图模型异常检测

孟雷　斗象科技高级安全研究员

威胁时刻存在


中国境内被植入后门网站数量
6000
5000
4000
3000
2000
1000
0
1月份 2月份 3月份 4月份 5月份 6月份 7月份 8月份 9月份 10月份 11月份
2017年 2018年


CNVD收录可实施远程攻击系统安全漏洞数量
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1月份 2月份 3月份 4月份 5月份 6月份 7月份 8月份 9月份 10月份 11月份
2017年 2018年


中国境内被篡改网站数量
7000
6000
5000
4000
3000
2000
1000
0
1月份 2月份 3月份 4月份 5月份 6月份 7月份 8月份 9月份 10月份 11月份
2017年 2018年

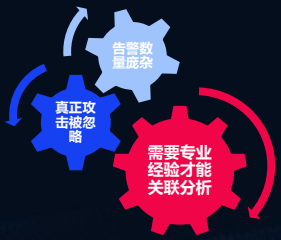

2017年国内DDoS僵尸网络攻击态势
18000000
16000000
14000000
12000000
10000000
8000000
6000000
4000000
2000000
0
1月份 2月份 3月份 4月份 5月份 6月份 7月份 8月份 9月份 10月份 11月份 12月份

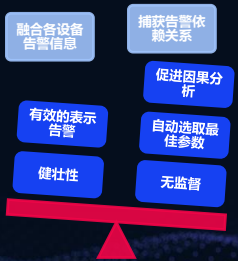

Initial Exploit
Installation
Botnet Monetizatio
Reconnaissance
Lateral Movement
Exfiltration

攻击 VS 保护


WAF
IDS
SIEM
HIDS
SOC

面临的问题

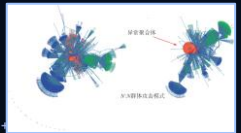

告警数量庞杂
真正攻击被忽略
需要专业经验才能关联分析


融合各设备告警信息
捕获告警依赖关系
促进因果分析
有效的表示告警
自动选取最佳参数
健壮性
无监督

图模型

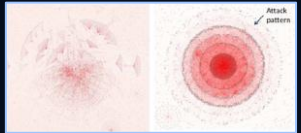
图结构可视化

DDoS攻击



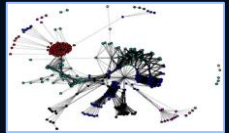
Reference: 叶晓鸣,陈兴蜀,杨力,王文贺,朱毅,邵国林,梁刚.基于图演化+事件的主机群异常检测模型[J].山东大学学报(理学版),2018,53(09):1-1

POSTECH可视化DDoS攻击



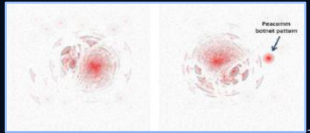
Reference: Le, Do Quoc, et al. "Traffic dispersion graph based anomaly detection." Proceedings of the Second Symposium on Information and Communication Technology. ACM, 2011.

Russia’s APT28攻击



Reference: Pei, Kexin, et al. "Hercule: Attack story reconstruction via community discovery on correlated log graph." Proceedings of the 32Nd Annual Conference on Computer Security Applications. ACM, 2016..

Peacomm僵尸网络



Reference: 정태열, Le Quoc Do, et al. "A Graph-based Detection of Anomalous Network Traffic."

图构建

算法 1. 收集候选主机 N 分钟内告警数据 A_{alerts} 和相应的其他设备告警数据 A_{other}

```
Function   FIND_CANDIDATES(A_alerts, A_other, alert window N)
    C ← { }                                  ▷ Initialize an empty set of (unique) candidates
    For a ∈ A_alerts  do                     ▷ Mine candidates from alert data
        append(C, a_srcIp)
        append(C, a_dstIp)
    End for
    E ← { }                                  ▷ Initialize a map from candidates to event lists
    For c ∈ C do                             ▷ Get last N minutes of other matching candidate
        E[c] ← {a|a ∈ A_other, c_timestamp − a_timestamp ≤ N, a_srcIp = c ∨ a_dstIp = c}
    End for
    Return E
End function
```

算法 2. 构建告警关联图

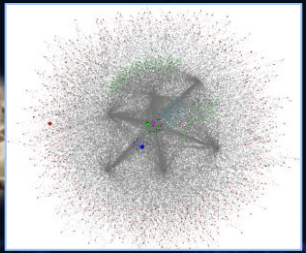
```
Function   MAKE_GRAPH(event list E)
    S ← {a|a.∈ E[ i ], i ∈ 0, . . |E|−1}      ▷ Set of unique attribute-value pairs in E
    G ← (V, E), |V| = |S|                     ▷ Initialize an undirected graph
    H ← { }
    K ← { }
    For a_i ∈ S, i ∈ 0, . . |S|−1  do
        H[a_i] ← i                            ▷ Map attributes to vertices
        K[i] ← a_i                            ▷ Map vertices to attributes
    End for
    For e ∈ E  do                             ▷ Add events to graph
        For a_i ∈ e,  i ∈ 0, . . |e|−1  do
            h ← H(a_i)
            for a'_j ∈ e, j ∈ i+1, . . |e|−1  do   ▷ Link attributes from same eneent
                h' ← H(a'_j)
                Add an edge between V_h and V_h' in G.
            End for
        End for
    End for
    Return G, K
End function
```

HIDS Alert

** AlerT 16783422.345663: syslog. vsftpD. coNnecTioN :AttEmpT 2018 Nov 29 15:06:33 (host) 10.0.81.56 -> /var/log/vsftpd.log Rule: 21502(level 2) -> ‘FTP session opened.’ Src IP: 10.0.81.16 Thurs Nov 29 10:07:46 2018 [pid 16349]

Snort Alert

[1:29-10:22:19.403923 [**] [1:2011487:2] ET POLICY Suspicious inbound to PostgreSQL port 5432 [**] [Potentially Bad Traffic] [Priority: 2] {TCP} 10.0.81.16:38989 -> 10.0.81.73:3512



方法流程

特征抽取：图节点角色模型

异常检测：时序动态角色模型、多目标回归模型

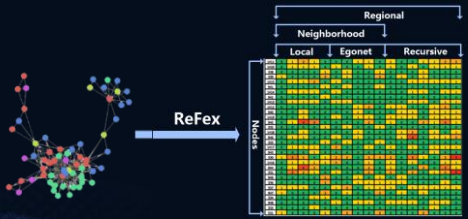
告警真实度预测：随机森林

图节点角色模型

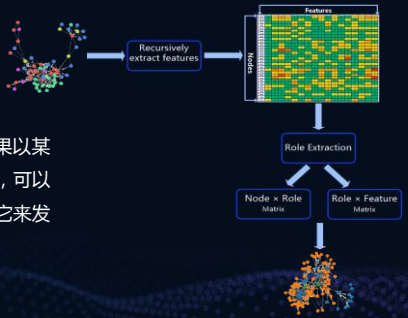
从多个设备*告警日志*中，抽取关联*信息单元*，构成告警**关联图**。根据图方法中的计算指标，对原始告警依赖图做递归特征提取，生成*特征矩阵*。依据前置的*角色度量属性*，对特征矩阵做*非负矩阵分解*，计算每个节点各角色概率分布信息。生成各节点**角色分布图**

图节点角色模型

☛ Recursive Feature eXtraction（ReFeX）由Henderson等人在2011年提出用于对图节点进行递归特征提取，是一种结构图特征提取算法



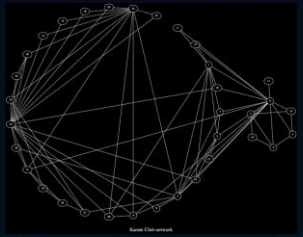
图节点角色模型



- RolX (Role eXtraction)

Henderson等人在2012年提出。RolX的核心思想：如果以某种线性形式（例如一个矩阵）收集关于一个图形的数据，可以使用矩阵分解方法来找到数据中的结构，并且可能使用它来发现图本身中的相应结构

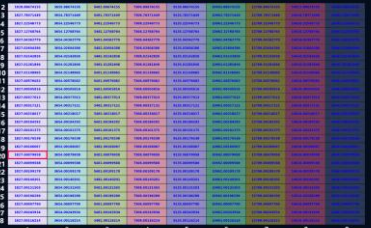
图节点角色模型-效果呈现



- 数据集

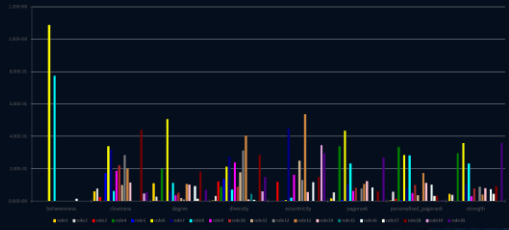
Karate Club network，美国一所大学中空手道俱乐部34名成员间的社会关系

图节点角色模型-效果呈现



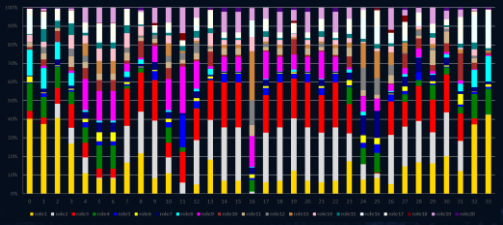
- 计算角色数量Roles和描述长度(bits)b的描述代价L
- 当Roles=20，b=1时，min(L)=1827.00079958

图节点角色模型-效果呈现



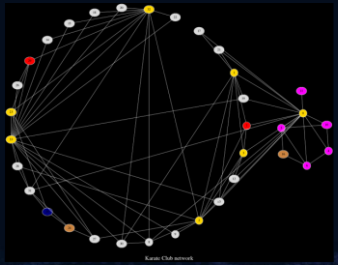
- 构建**20**个role的各角色含义
- 度量值为['betweenness', 'closeness', 'degree', 'diversity', 'eccentricity', 'pagerank', 'personalized_pagerank', 'strength']

图节点角色模型-效果呈现

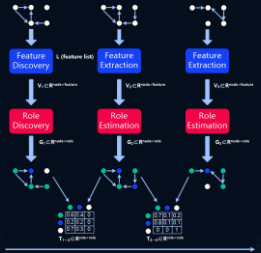


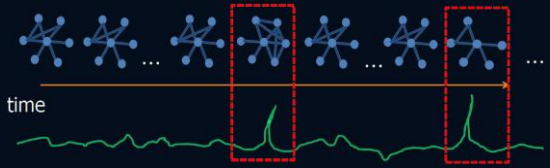
- 根据Rolx算法计算结果，标注出个节点所属角色分布
- 本数据中共有节点34个

图节点角色模型-效果呈现

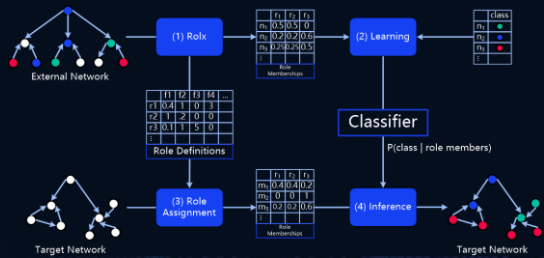


时序动态角色模型


Feature Discovery
Feature Extraction
Feature Extraction
Role Discovery
Role Estimation
Role Estimation


time

多目标回归模型


External Network
(1) RolX
Role Memberships
(2) Learning
class
Classifier
P(class | role members)
Role Definitions
(3) Role Assignment
Role Memberships
(4) Inference
Target Network
Target Network

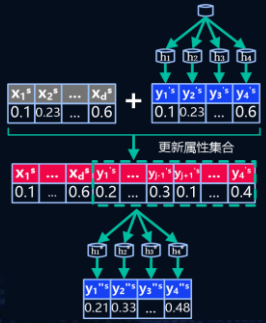
多目标回归模型 learning

一阶模型生成

一阶模型预测

二阶模型生成

多目标回归模型 inference



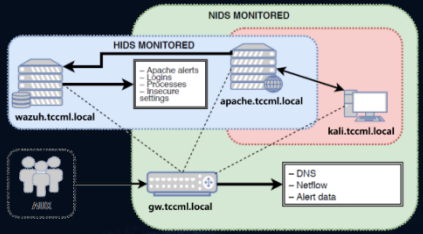
实验案例

攻击场景模拟

- wazuh.tccml.local

wazuh主节点，存储NSM数据

- gw.tccml.local网关、外部网络入口。已安装Suricata提供NIDS告警、dns和netflow数据



- kali.tccml.local

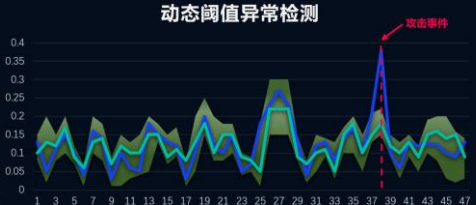
攻击者，内置kali linux

- apache.tccml.local

apache web、wordpress、wazuh-agent

异常检测


静态阈值异常检测
攻击事件
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47


动态阈值异常检测
攻击事件
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47

总结

问题：各检测设备告警数量巨大，多设备检测融合分析

手段： ReFeX、 RolX、时序动态角色模型、多目标回归模型

结果：模拟实验，验证方法可行


FREEBUF | FIT
THANKS