北京大学
PEKING UNIVERSITY

Web安全技术分享

2022.08.14 张芊

MORESHI POWERPOINT

目录

H 厂商远程代码执行漏洞

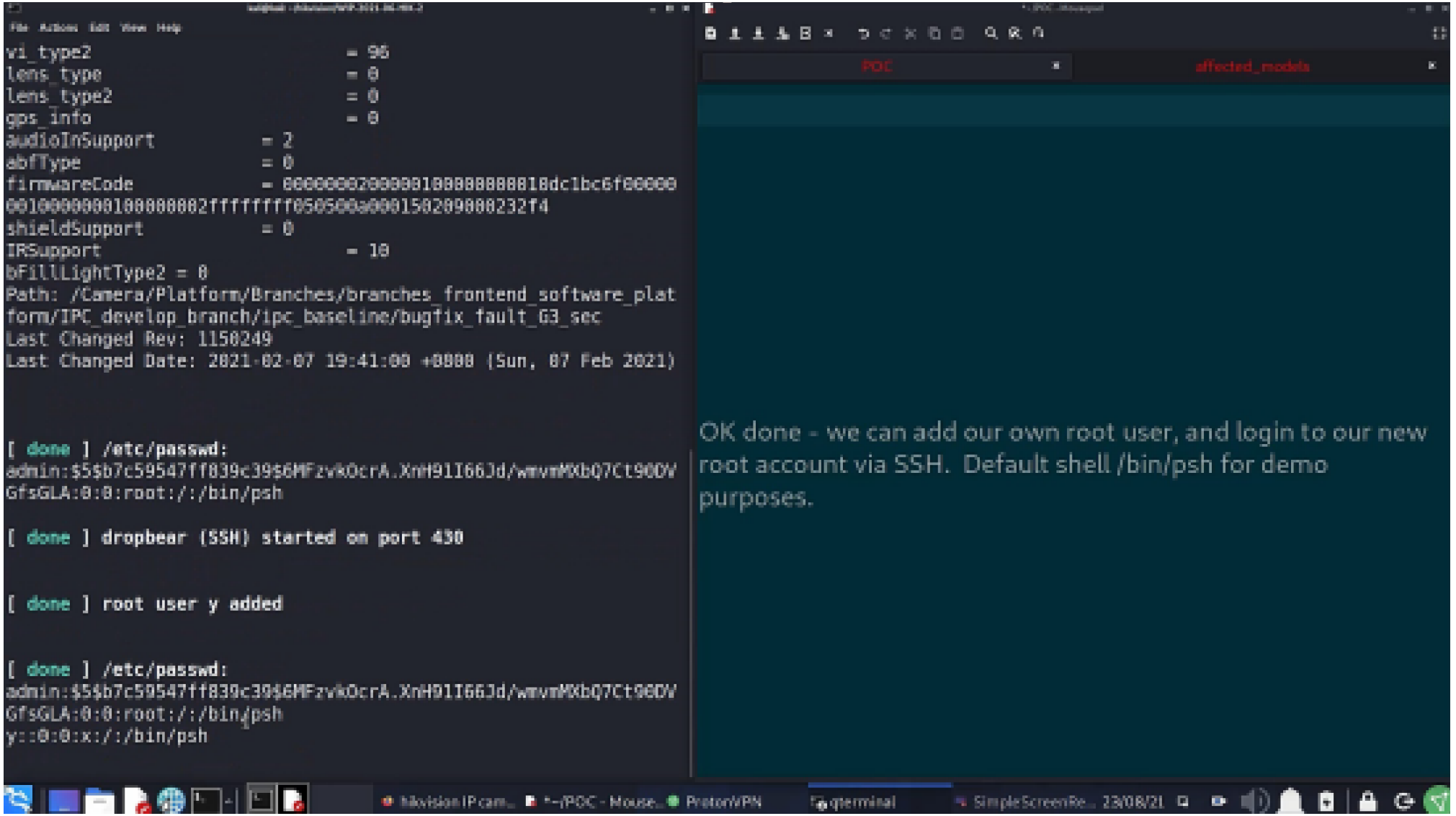
MORESHI POWERPOINT

2021 年 6 月，研究人员在 H 厂商 IP 摄像机设备固件中发现了一个未认证的远程代码执行漏洞，漏洞编号为 CVE-2021-36260。2021 年 8 月，H 厂商开始发布补丁和相关受影响的设备和固件。

该漏洞的利用很简单，不需要用户的交互，攻击者只需要访问 HTTP 或 HTTPS 服务器的80 或者 443 端口就可利用该漏洞，而且也不需要登录时的用户名，密码或者其他任何操作，摄像头本身也不会检测到任何登录信息，在获取设备信息之后，添加一个 root 账户，并进行登录。一旦攻击成功，攻击者可以读取和更改用户数据，而且还可以访问和攻击内部网络。

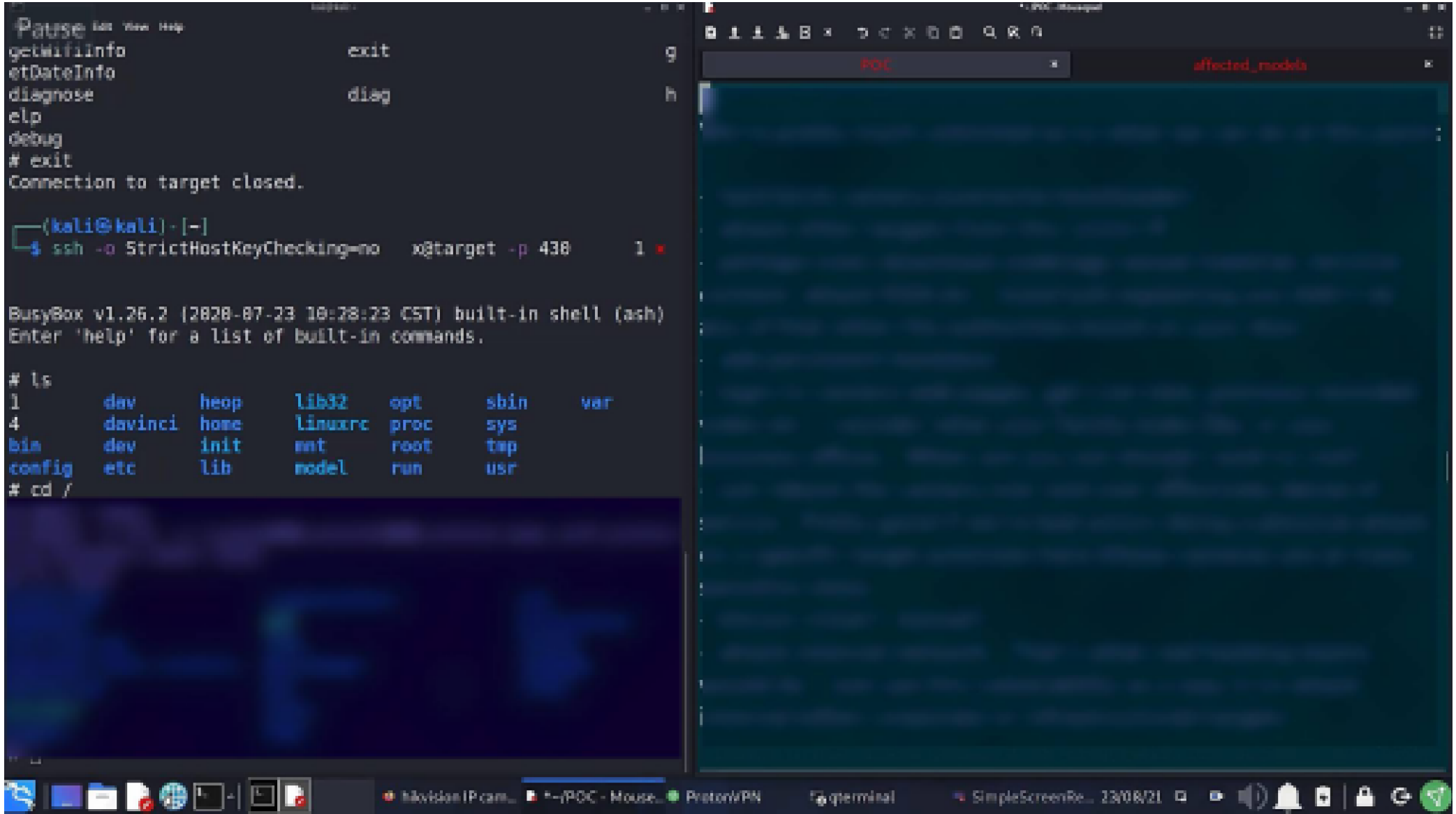
H 厂商远程代码执行漏洞

MORESHI POWERPOINT



H 厂商远程代码执行漏洞

MORESHI POWERPOINT



H 厂商远程代码执行漏洞

MORESHI POWERPOINT

当下，摄像头已经成为了家庭、企业单位必不可少的监控设备，这些设备的数据大多数情况下都是隐私数据，一旦被入侵或者非法利用后果不堪设想，轻则被用于肉鸡对其他互联网资产进行 DDoS 攻击，重则隐私数据泄露，被攻击者利用从而发起社会工程学攻击，造成财产损失。作为用户来讲，可以从以下两个方面提高设备的安全性：一个是关闭不必要的端口，减小攻击面，另一个是及时打厂商发布的补丁，尽量避免影响

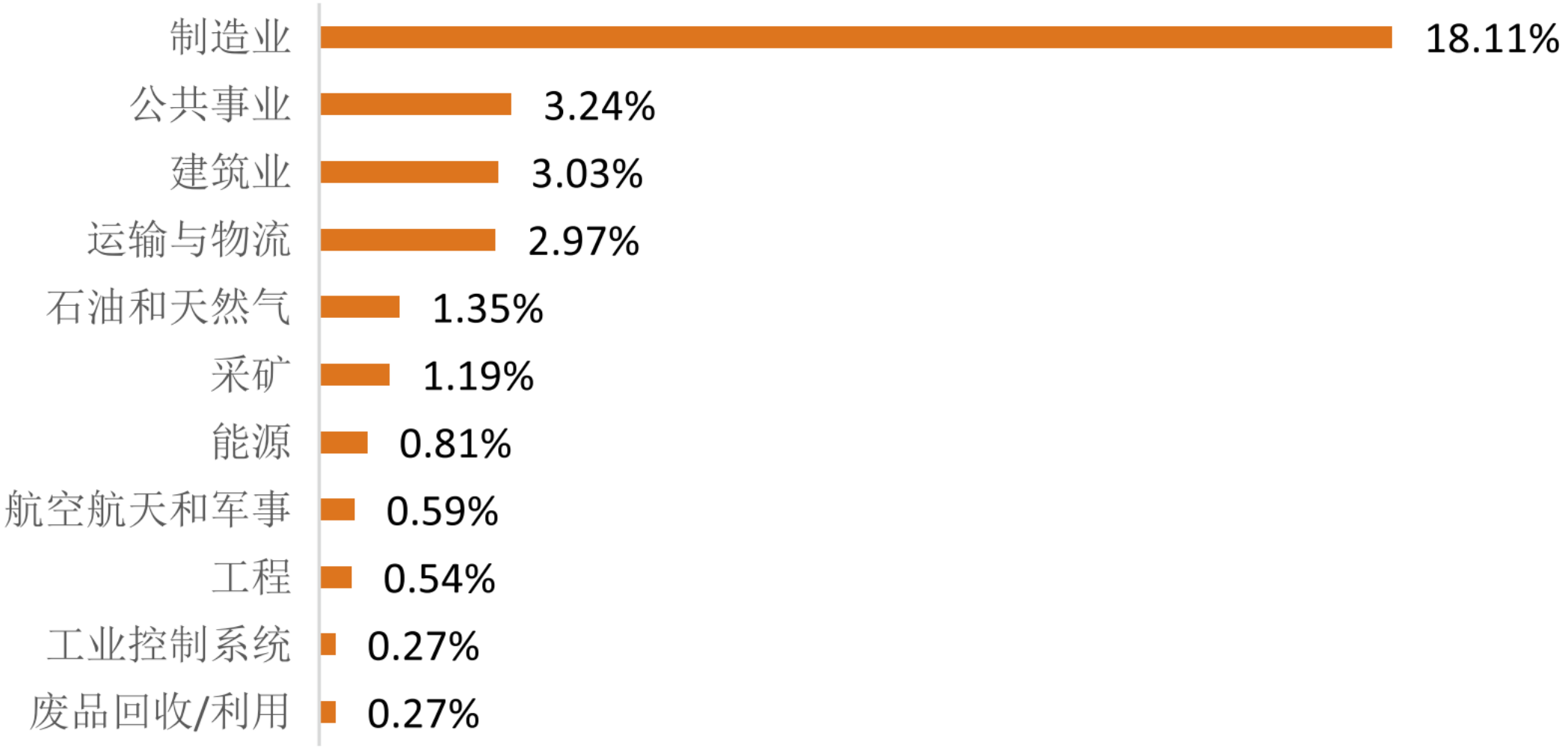
数百家工业组织在 SolarWinds 事件中遭受 Sunburst 攻击

MORESHI POWERPOINT

2019.4-2020.2.1 期间发布的版本中植入了恶意的后门应用程序，受此次供应链攻击影响的客户包括政府、国防、网络公司和关键基础设施提供商等，约有 18000 个用户下载了包含后门的恶意软件。安全研究人员对使用后门版本的 SolarWinds 并成为受害者的工业组织进行研究，他们使用 Sunburst 恶意软件域名生成算法生成的 DNS 名称获得的可用内部域名，其中已解码和可归属的域名将近 2000 个，按照行业划分，工业组织占比最高（32.4%），覆盖工业领域中制造业、运输与物流、建筑业、矿业和能源等（如图 1.5 所示），工业组织的地理分布也几乎覆盖整个世界

数百家工业组织在 SolarWinds 事件中遭受 Sunburst 攻击

MORESHI POWERPOINT



TeamTNT 组织在 2021 年多次针对云计算目标进行攻击

MORESHI POWERPOINT

TeamTNT 组织至少从 2011 年就开始活跃。他们攻击手法多样，近两年来，也多采用云及云原生相关攻击手段实施攻击。据不完全统计，TeamTNT 组织在 2021 年进行了一系列的云相关攻攻击

- 2021 年 2 月，TeamTNT 被曝投放针对 Kubernetes 集群的非法加密挖矿软件
- 2021 年 5 月，TeamTNT 被曝针对 Kubernetes 进行蠕虫式攻击，至少五万个 IP 被感染
- 2021 年 9 月，TeamTNT 被曝发起了针对多个操作系统和应用的攻击行动“Chimaera”

TeamTNT 组织在 2021 年多次针对云计算目标进行攻击

MORESHI POWERPOINT

- 2021 年 10 月，TeamTNT 被曝在 Docker Hub 上投放恶意镜像
- 2021 年 11 月，TeamTNT 被曝通过存在未授权访问漏洞的 Docker 控制服务器执行挖矿等恶意操作

随着容器及云原生技术逐渐成熟，云原生化会成为常态。在初始渗透阶段，TeamTNT 并未利用高级的攻击手段，仅仅是目标主机的错误配置，就可以导致上万台主机失陷。上述一次次的事件必须引起我们的重视，加强云和云原生环境的基本配置核查、加固，避免给攻击者可乘之机。

小结

MORESHI POWERPOINT

从物联网 SDK 的安全事件可以看出，对类似物联网产业这种软硬件产业结构复杂的产业，厂商不仅管理好自身安全的同时还需要关注供应链安全。

工业控制系统领域主要是勒索软件青睐对象，因为对于多数受害企业来说，他们无法承受系统或生产线停止服务的代价，所以攻击者会有更大概率拿到赎金。

近年来越来越多的安全设备漏洞被披露出来，其中远程命令执行、SQL 注入和未授权访问较为常见。“新冠”疫情大爆发以来，远程办公成为很多企业主要的工作方式之一，因此企业 VPN 产品相关漏洞受到攻击者的关注。

网络空间资产测绘简介

MORESHI POWERPOINT

网络空间测绘作为一项十分重要的基础性工作，是网络空间国防能力建设的重要部分，是大国博弈背景下，网络主权、网络边疆的重要体现，美国＂智库＂兰德公司也曾断言：工业时代的战略战是核战争，信息时代的战略战主要是网络战。网络空间测绘对推动国民经济和保障国家安全都具有十分重要的理论意义和应用价值。

在网络空间测绘领域的起步阶段，主要集中于理论和概念的研究，结合网络测量技术和地理测绘知识，在资产探测、拓扑测量、IP 定位层面逐步发展 。现阶段更注重的是在海量多源异构数据的基础上进行信息同化和融合分析，根据不同应用场景和需求，应用可视化术，结合人工智能，对所有信息分门别类地进行展示。

网络空间资产测绘简介

MORESHI POWERPOINT

测绘最早来源于地理空间地图的绘制，主要研究测定和推算地面几何位置、地球形状及地球重力场，据此测量地球表面自然物体和人工设施的几何分布，编制各种比例尺地图的理论和技术的学科（维基百科）。网络空间测绘和地理信息测绘的技术路线类似，“测”是对网络空间内一切可获得数据的测量机制的建立，偏向于实现扫描和探测的工程问题；“绘”则是根据对网络空间测量数据分析和关联，包括地址地理、域名、风险脆弱性等信息的关联，目的是绘制出多维的网络空间地图，倾向于对数据的分析和研究。

网络空间资产测绘简介

MORESHI POWERPOINT

相比于地理信息测绘，网络空间测绘存在一些特殊之处。首先从数据维度来讲，地理空间的测绘数据是三维的（经度、纬度、海拔）且连续，而网络空间中将 IP 地址转化为长整形后，地址数据是一维的，并且每个点都是独立存在并不连续。此外，二者还有一个最大的不同之处就是变化频率，地理信息测绘数据一般变化较慢，而且因为是连续的，所以变化趋势相对好预测，比如珠穆朗玛峰的每年都会以一定的高度在增长，但正常情况下一般不会突然升高或下降几十米。而网络空间测绘数据则不同，绝大多数的 IP 地址处于变化是常态。比如存活情况、开放服务、ASN、地理信息、地址所有者等等维度都是处在动态变化中，并且因为网络地址都是离散分布的个体，变化趋势也就更难预测。所以大部分提及的网络空间资产测绘结果，都是基于实时扫描一轮的数据展开的分析，以保证资产测绘的准确性。

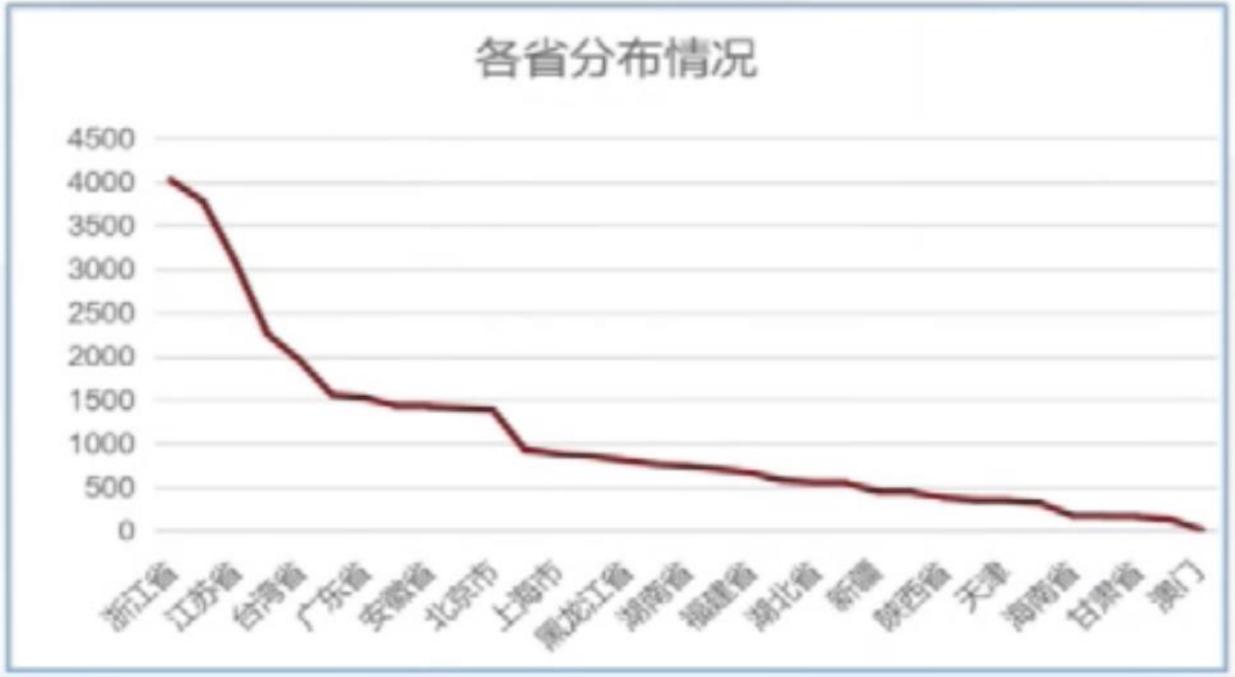
预警实践

MORESHI POWERPOINT

价值点:落实《关于开展摄像头偷
类专项检查;按照中央网信办、工业和信
展摄像头偷窥等黑产集中治理的公告》
展摄像头偷窥黑产集中治理，全面清理
险利用、破解工具售卖，偷拍设备改装
关违法有害信息。

2020年1月20日境外黑客组织A
号密码，数据去重后共计**全球97735**个

94248（国内） IP+Port 数量（个）		35051 独立 IP 数量（个）
10 行业分布（个）	4613 开放端口（个）	1959 开放服务（个）
77 操作系统（种）	410 组件（个）	23 设备类型（种）
1101 漏洞（个）	1 弱口令数量（个）	7 PoC（非常危险）



技术痛点

MORESHI POWERPOINT

测

- 存活探测
- IPv6测绘
- CDN网络测绘
- 安防网络设备测绘
- 绕防技术研究
- 穿透扫描机制
- 拓扑测量
- 协议一致性研究
- 高端口漂移技术研究

绘

- 漏洞测绘的窗口期
- 组织结构识别
- 无特征节点推演
- 拓扑绘制
- IP定位技术
- 暗网探测

调度技术

- 法律法规
- 反溯源
- 反测绘
- 蜜罐
- 安防机制

可视化

- 网络地图可视化
- 组织结构形式化表达

应用场景

- 挂图作战
- 可视化侦察打击
- 漏洞传染面分析
- 攻击面管理
- 预警与联防
- 关基组织结构摸底
- ...

技术痛点

MORESHI POWERPOINT

网络空间资产测绘5步法

[痛点]

1. 存活探测结果前后不一致
2. 发包快容易被运营商拉黑
3. 发包慢探测效率低
4. 多次探测把目标网络扫死
5. 端口组合策略效率问题

基于开源的单节点测绘实验

MORESHI POWERPOINT

网段	配置		Zmap	Masscan	Nmap
14x.x.183.0/24	5Mbps 单端口 服务器固定	耗时	11min43s	1 min27s	7 min39s
		命中率	323	315	317
12x.x.212.0/24		耗时	11 min34s	54s	7min53s
		命中率	496	496	496
13x.x.177.0/24 10x.x.227.0/24		耗时	22min34s	1min29s	2h18min47s
		命中率	363	183	368

2级VPS基于开源的单节点测绘实验

MORESHI POWERPOINT

网段	配置		Zmap	Masscan	Nmap
14x.x.183.0/24	5Mbps 单端口 服务器固定	耗时	16min28s	2min45s	11 min66s
		命中率	242	280	293
12x.x.212.0/24		耗时	17min04s	1min32s	11min40s
		命中率	389	395	412
13x.x.177.0/24 10x.x.227.0/24		耗时	32min14s	3min29s	2h58min47
		命中率	302	135	323

开源软件探测存在问题

MORESHI POWERPOINT

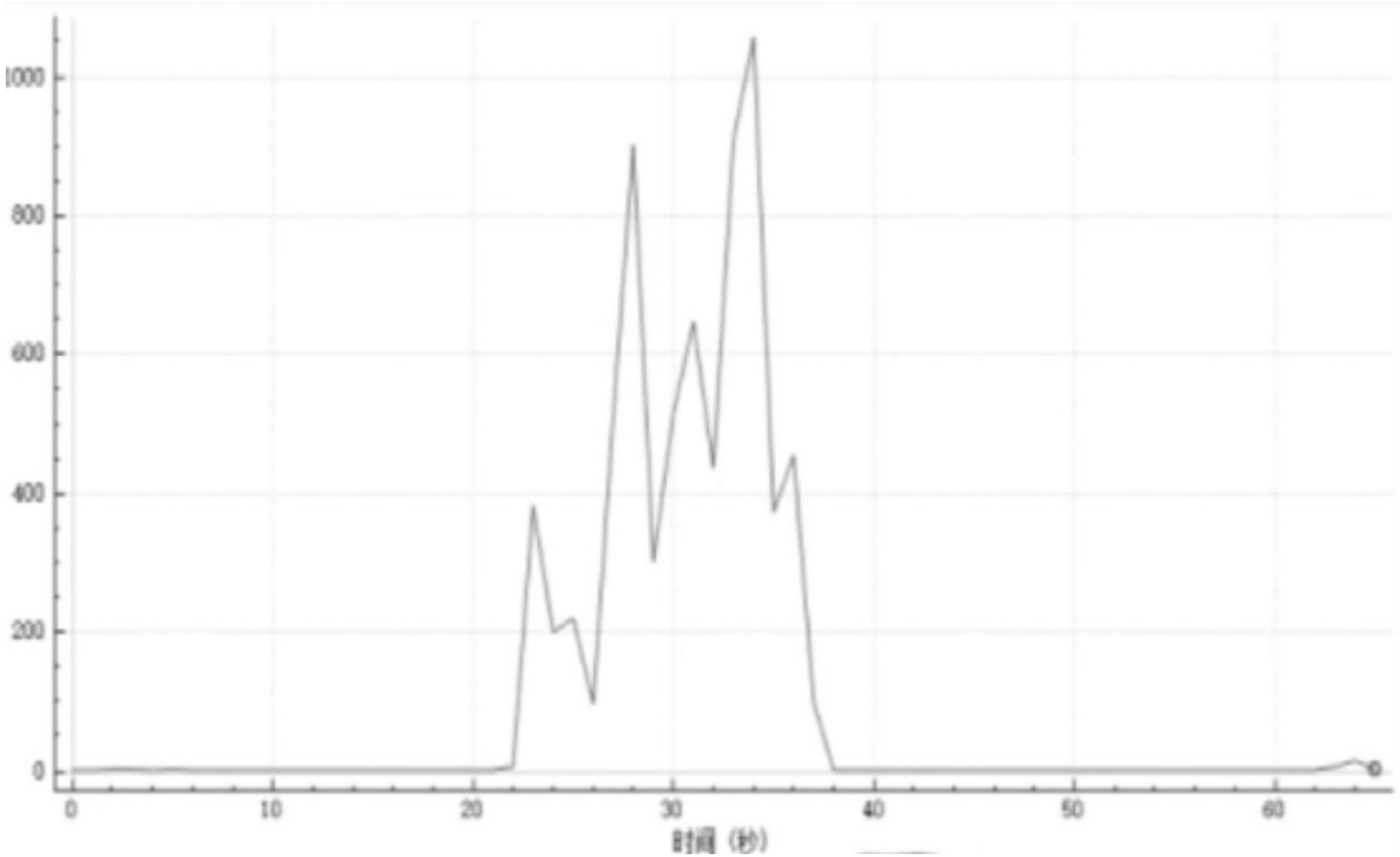
痛点:

①开源软件的探测扫描效率低，网络噪音大;

②开源的探测引擎易被安全设备阻断;

③开源软件容易对扫描目标成致命危害;

④开源软件探测过程不可控;



开源软件探测存在问题

MORESHI POWERPOINT

解决方

①改进

式;

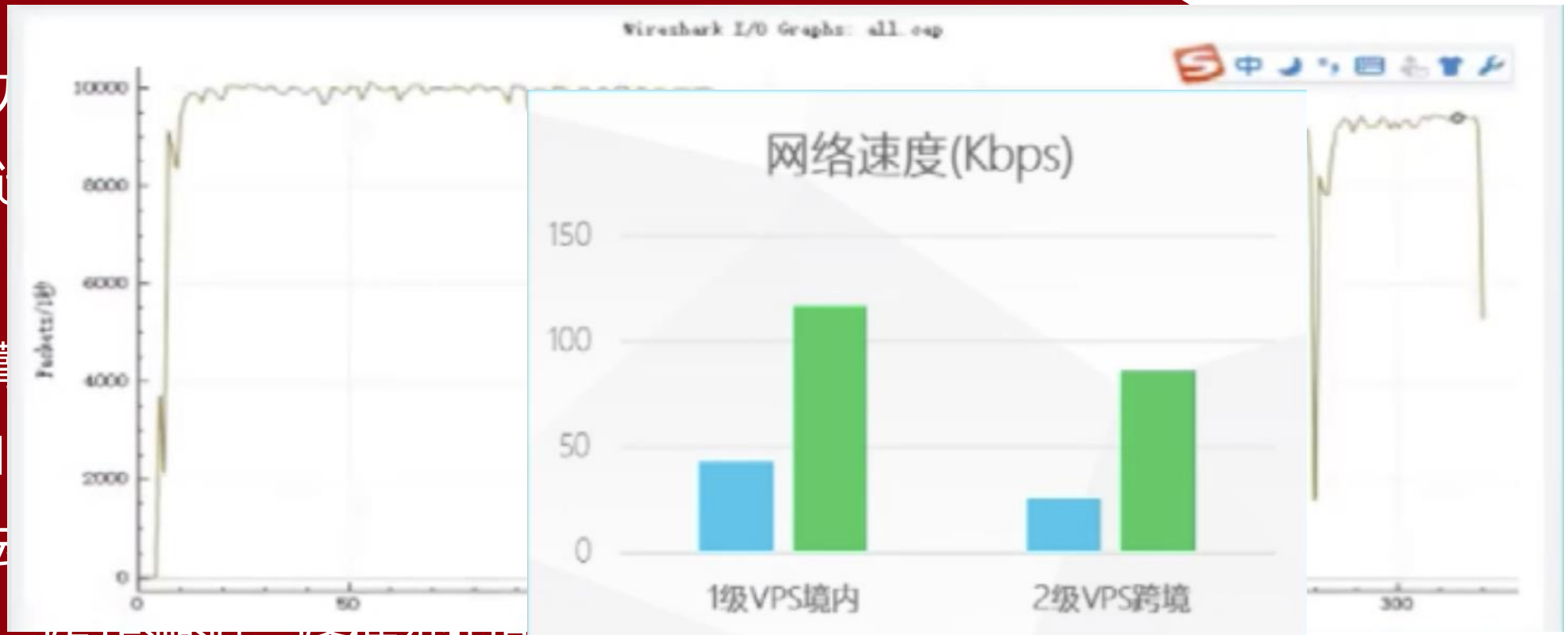
②引擎

③源IP

④自动

发包，防止溯源，降低被识别

⑤节点伪装技术，对抗反扫描



地理术语看地图异同

MORESHI POWERPOINT

Topgrahpic	Chart	Map
地图,地形地貌图	航海图	泛地图

地图的“态”与“势”

MORESHI POWERPOINT

地图的**态**是对环境的描述，以**测绘为手段**，把所有目标分解为**点线面体**。

地图的**势**是基于环境所展现的**活动**，活动**主题不同**，展现的**势**不同，传统地图可以有不同的主题。

网络空间地图，也就是对网络空间的环境进行描述。这是网络空间地图的基础部分，当前以测绘为主要手段，分为主动测绘和被动测绘二种。

网络空间地图

地图的“态”与“势”

MORESHI POWERPOINT

网络空间专题活动

网络空间专题活动

网络空间专题活动

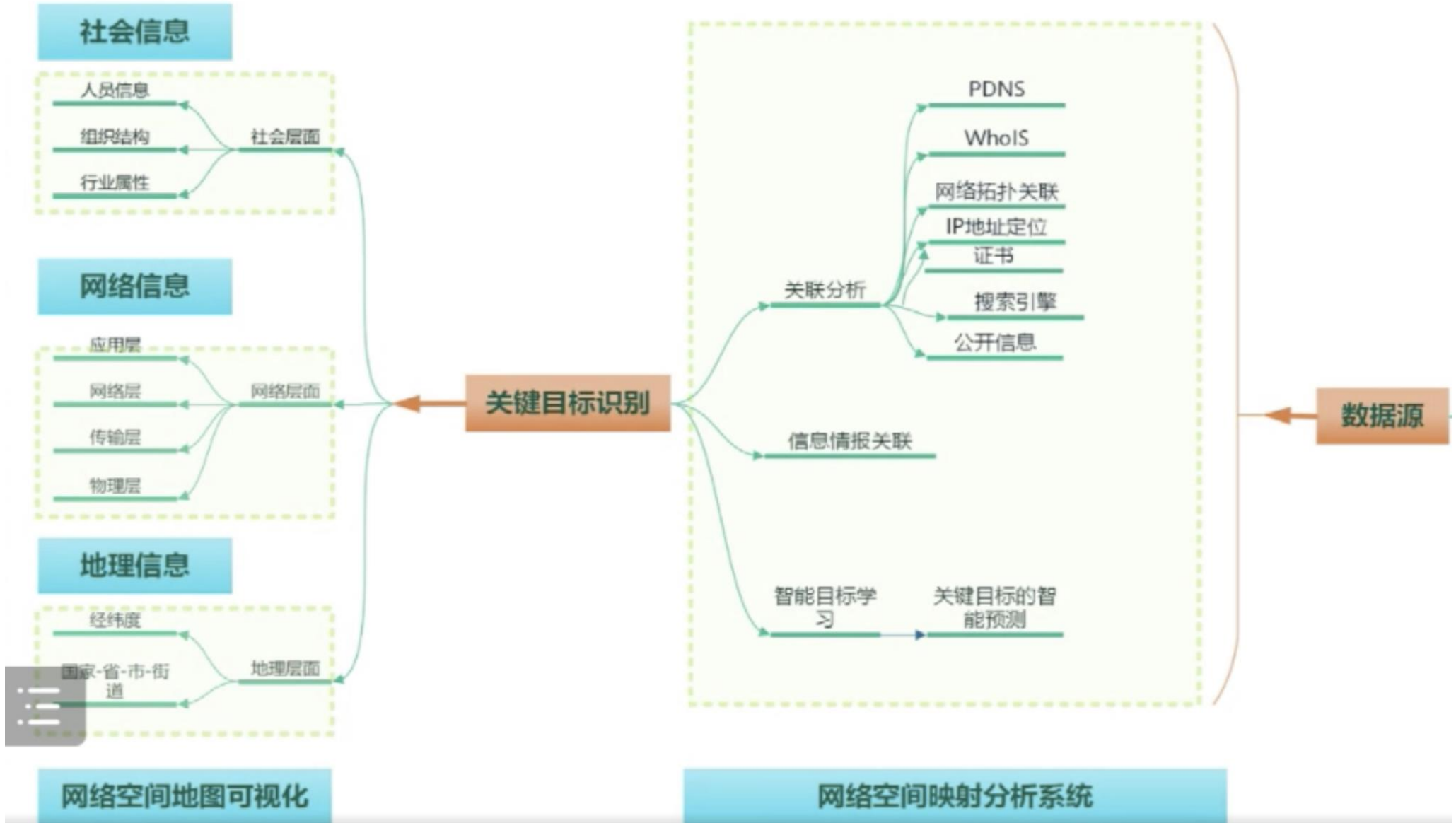
网络空间底图

地理图谱

网络空间地图的核心思想是Mapping

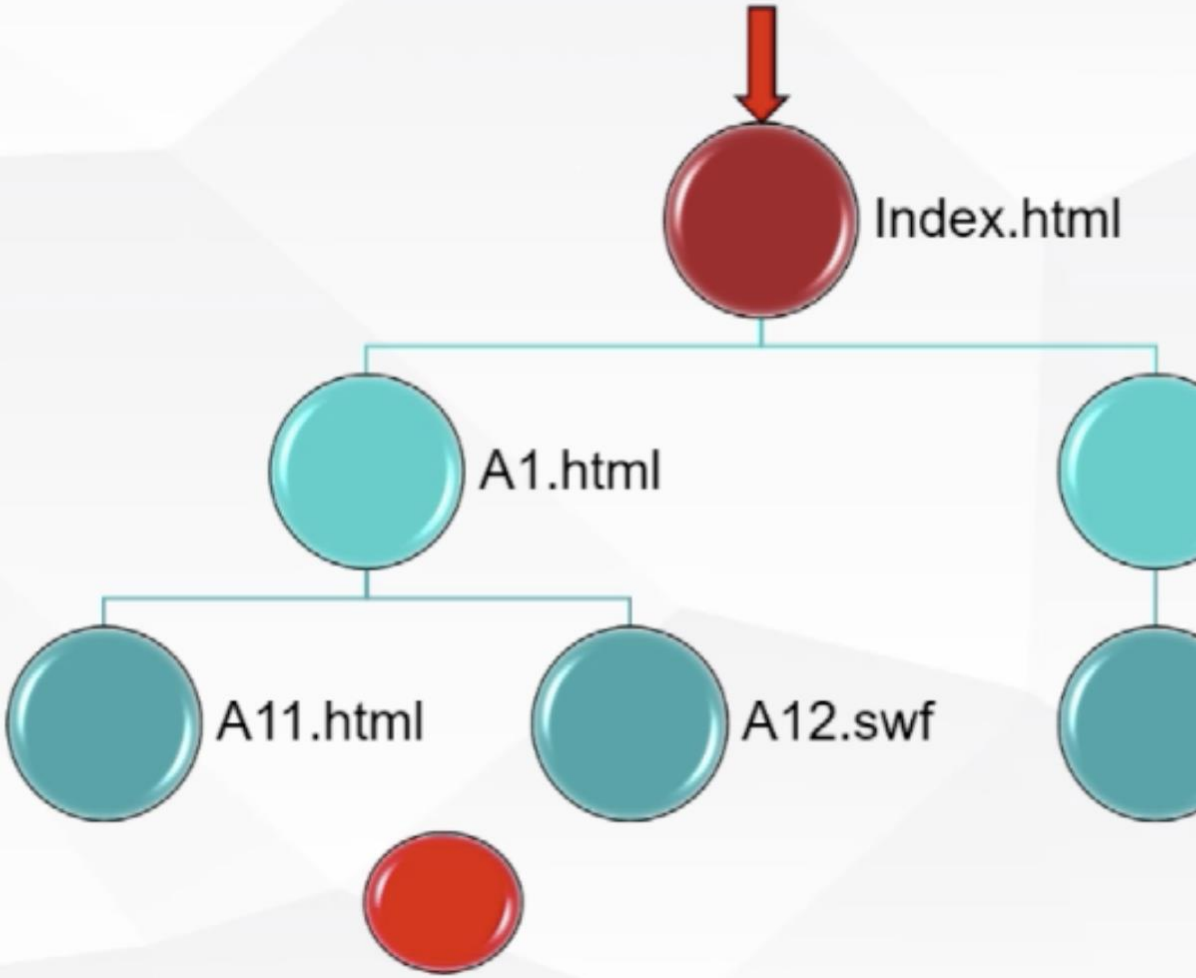
网络空间地图

MORESHI POWERPOINT



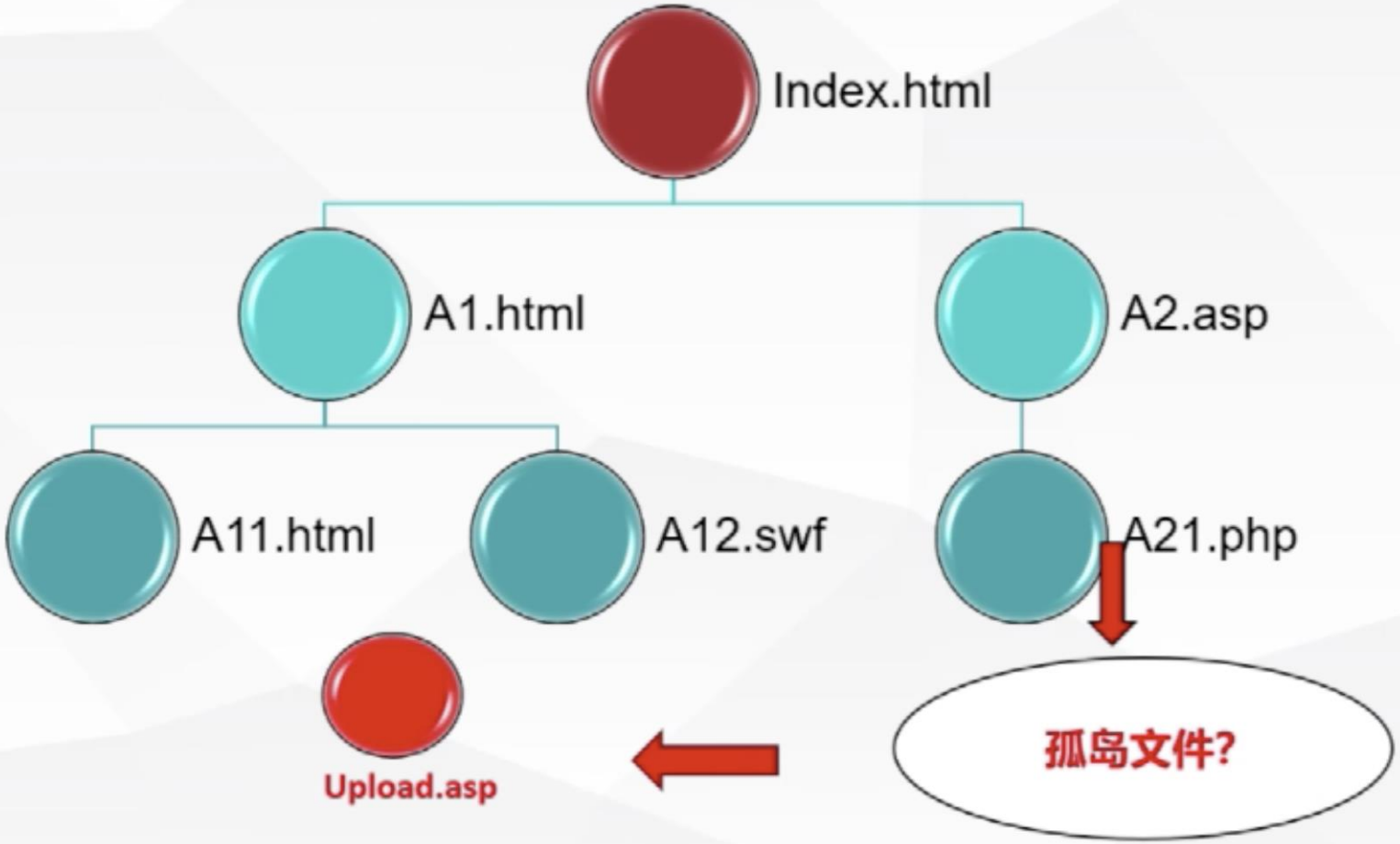
扫描器的谎言

MORESHI POWERPOINT



扫描器的谎言

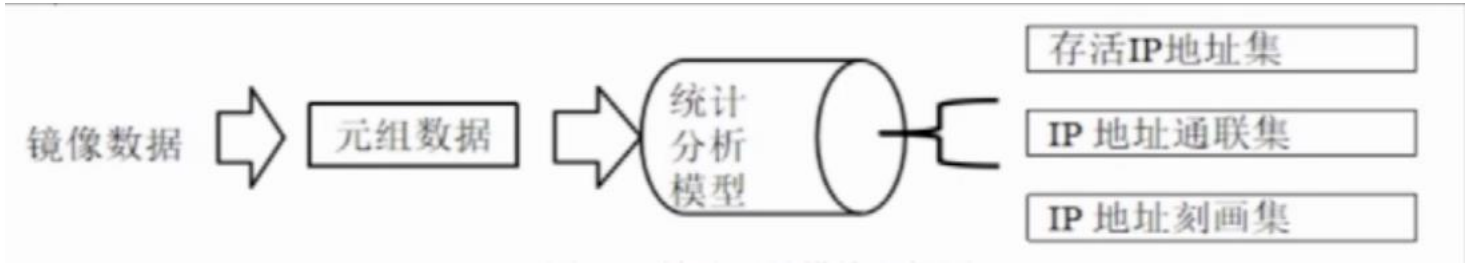
MORESHI POWERPOINT



被动测量-基于流量智能分析技术

MORESHI POWERPOINT

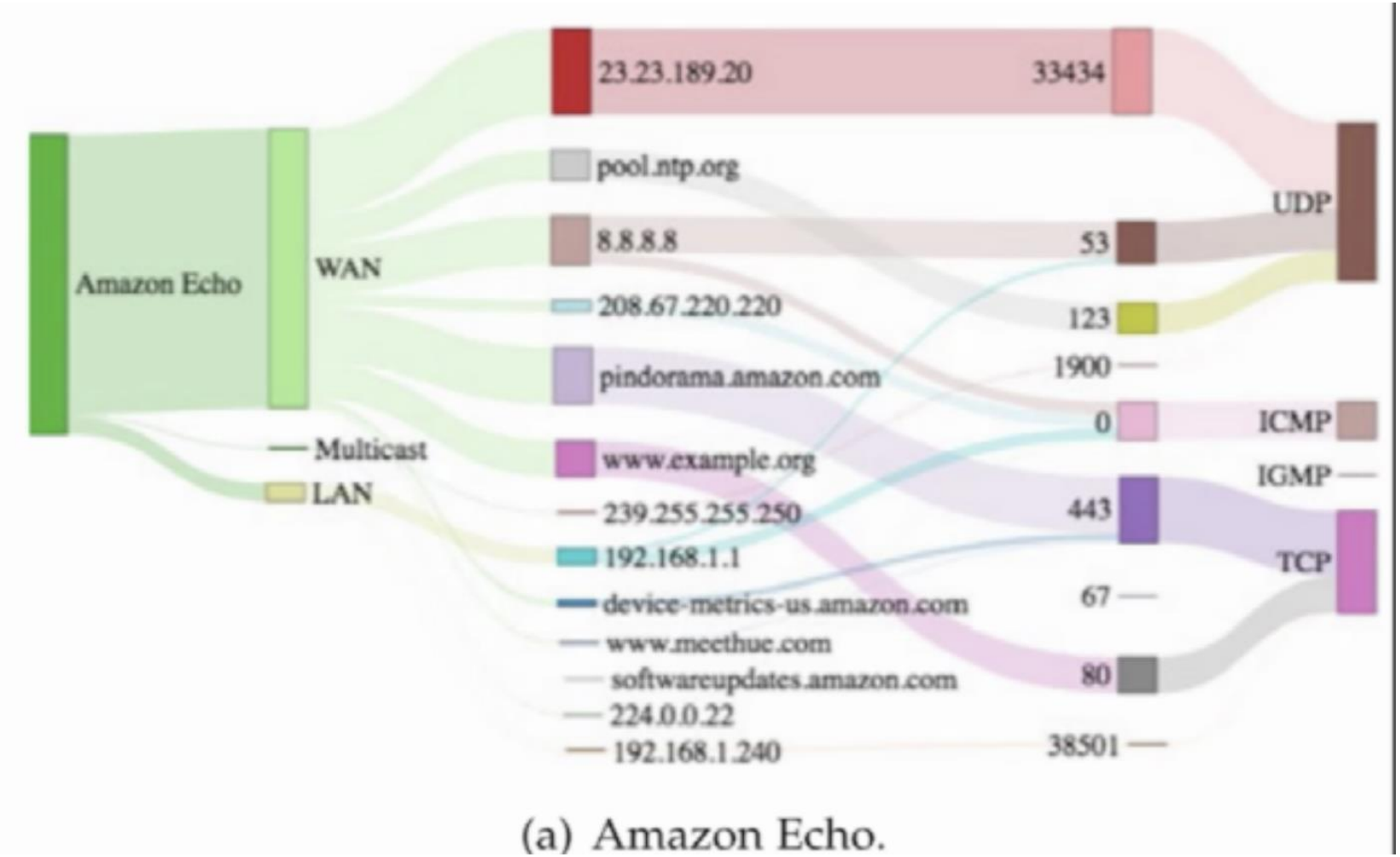
- 资产管理
- 盘点有漏洞的物联网设备
- 防止未授权的网络接入
- 网络资源限制



- 内部网络普遍存在:属于信息"富矿”类网络，与外部物理隔离或逻辑隔离，呈不可见状态。
- 层层设防:网络管控措施多，主动测量包不可达，召回率低;孤岛问题。
- 测量代价:向弱连接的工业互联网内注入大量探测包，可能影响可用性和可靠性。

IoT设备流量指纹识别技术

MORESHI POWERPOINT



(a) Amazon Echo.

方法: DPI >人工或半自动化地从流量中提取指纹信息，类似于IoC提取。

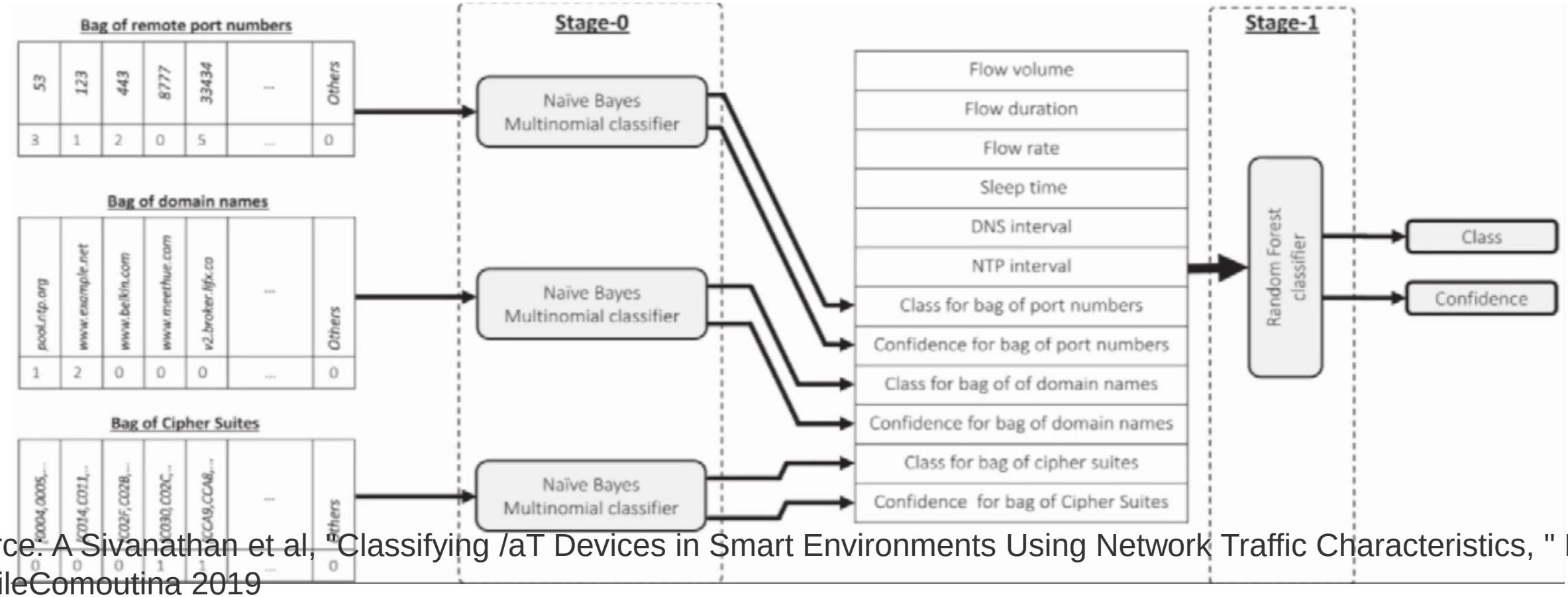
缺点:效率低，指纹不稳定。

Source: A. Sivanathan et al, "Classifying IoT Devices in Smart Environments Using Network Traffic Characteristics, " IEEE Transacticar-MobileComputing, 2019

被动测量-基于流量智能分析技术

MORESHI POWERPOINT

- 方法:流量收集> [人工标签] > L3~L7提取特征>预处理>模型训练>部署
- 难点:标签数据、特征工程、概念漂移、模型在线更新



Source: A Sivanathan et al, "Classifying /aT Devices in Smart Environments Using Network Traffic Characteristics, " IEEE MobileComoutina 2019

流量加密为IoT设备指纹识别带来的挑战

MORESHI POWERPOINT

趋势:

- ➢ 南北向流量加密占比约为81%;
- ➢ TLS 1.3和 DNSSEC 采用率逐步升高;
- ➢主流的 IoT Platform 大多已采用HTTPS加密流量。

挑战:

- ➢ 被动测量识别 IoT 设备无法解析 Payload;
- ➢ 成为很多流量分析类安全产品的天花板

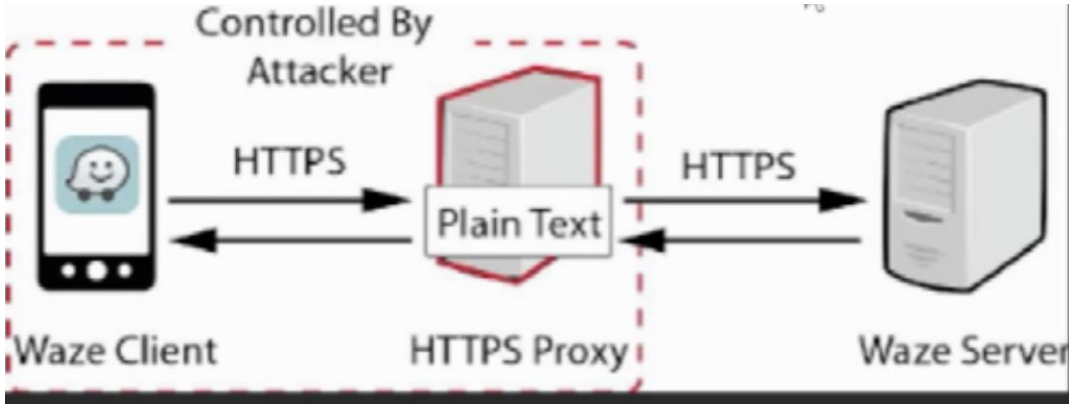
Source: Gigamon, "TLS Adoption Research, ” 2020.
X. He, et. al, "fingerprooting Mainstream 1o1 Platforrs Using Traffic Analysis, " IEEE Internet of Things Journal, 2022

后流量加密时代 IoT 设备指纹识别: MITM

MORESHI POWERPOINT

➢ 安全产品现状:当前应对加密流量的普遍方法是MITM解密;

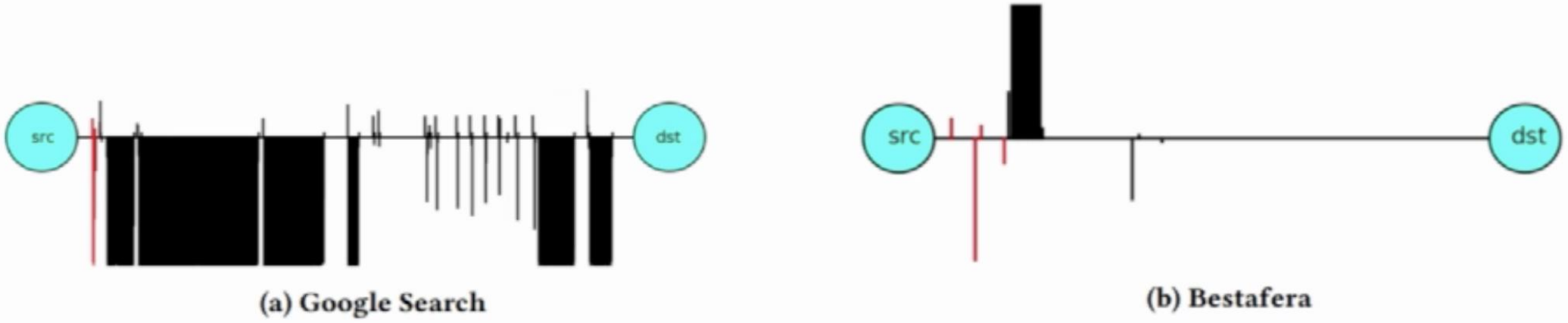
➢ 存在问题:资源消耗高、密钥管理风险、数据泄露风险。



Platform	Difficulty	Fingerprinting Features
SmartThings	Easy	The Value of the Field "Server" is "smartthings".
AWS IoT	Medium	Service Provider is Identified by the Value "Server" of the Field "Server" and the Existence of Several Fields Contains the String "amz". Then IoT and Non-IoT Traffic is Distinguished by a Series of Browser Security Configuration Fields and Their Values.
Mijia	Difficult	Json Contents Follow Specific Formats.
Connect	Difficult	Samsung Connect Use OpenResty as Web Server, and the Contents of the Packets Follow Certain Rules.
Alink	Difficult	Service Provider is Roughly Identified by the Value "Tengine/Aserver" or "Tengine" of the Field "Server" and the Existence of Several Fields Contains the String "am". Service Provider and IoT Traffic are Then Formally Identified by a Series of Browser Security Configuration Fields and Their Values, as well as the Certain Format of Json Contents.

Source: X. He, et. al, "fingerprinting Mainstream 1oT Platforms Using Traffic Analysis,” IEEE Internet of Things Journal, 2022
Gang w et al Attacks and Defenses in Crowdsourced Mapping Services[J]. Computer ence, 2015.

后流量加密时代 IoT 设备指纹识别: Non-Payload+ML



(a) Google Search

(b) Bestafera

Figure 2: The TLS packet lengths and inter-arrival times for a typical Google search and malicious data exfiltration from bestafera. Upward and downward lines represent the sizes of packets being sent from client → server and server → client, respectively. The x-axis represents time.

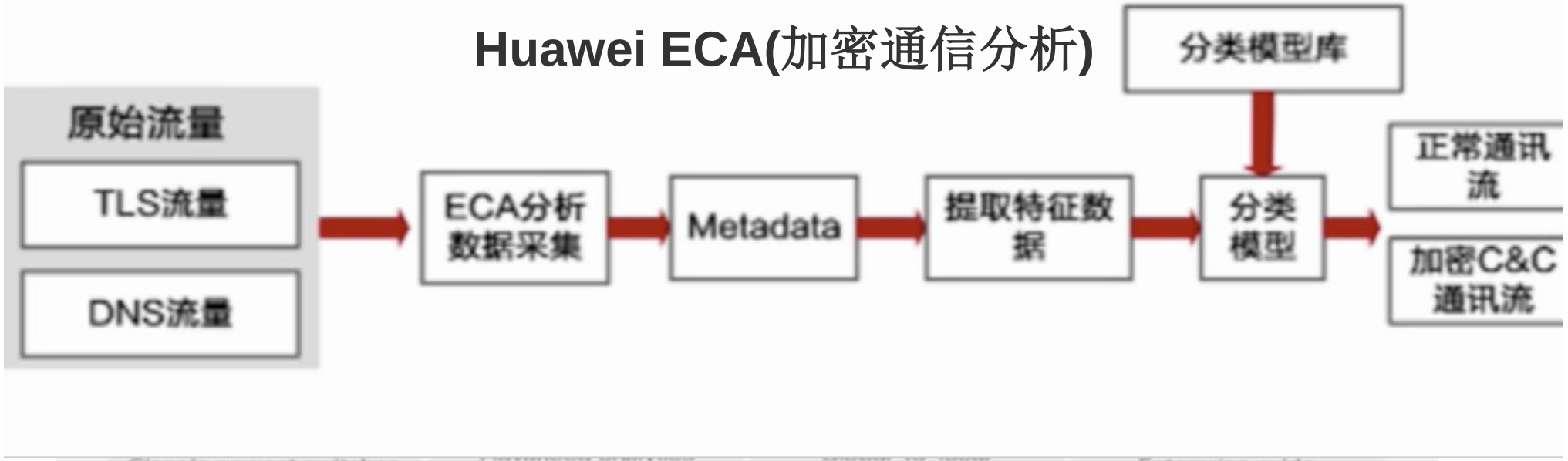
除Payload, PDU头域及时空统计数值蕴含了丰富的特征信息

Source: B. Anderson el. al, "Machine learning for encrypted malware traffic classification: Accounting for noisy labels and non-sta vrity"ACM SIGKDD, 2017.

后流量加密时代 IoT 设备指纹识别: Non-Payload+ML

Cisco ETA (Encrypted Traffic Analytics)

Finding malicious activity in encrypted traffic



Cisco's newest switches and routers

and machine learning

knowledge correlation

Enterprise-wide compliance

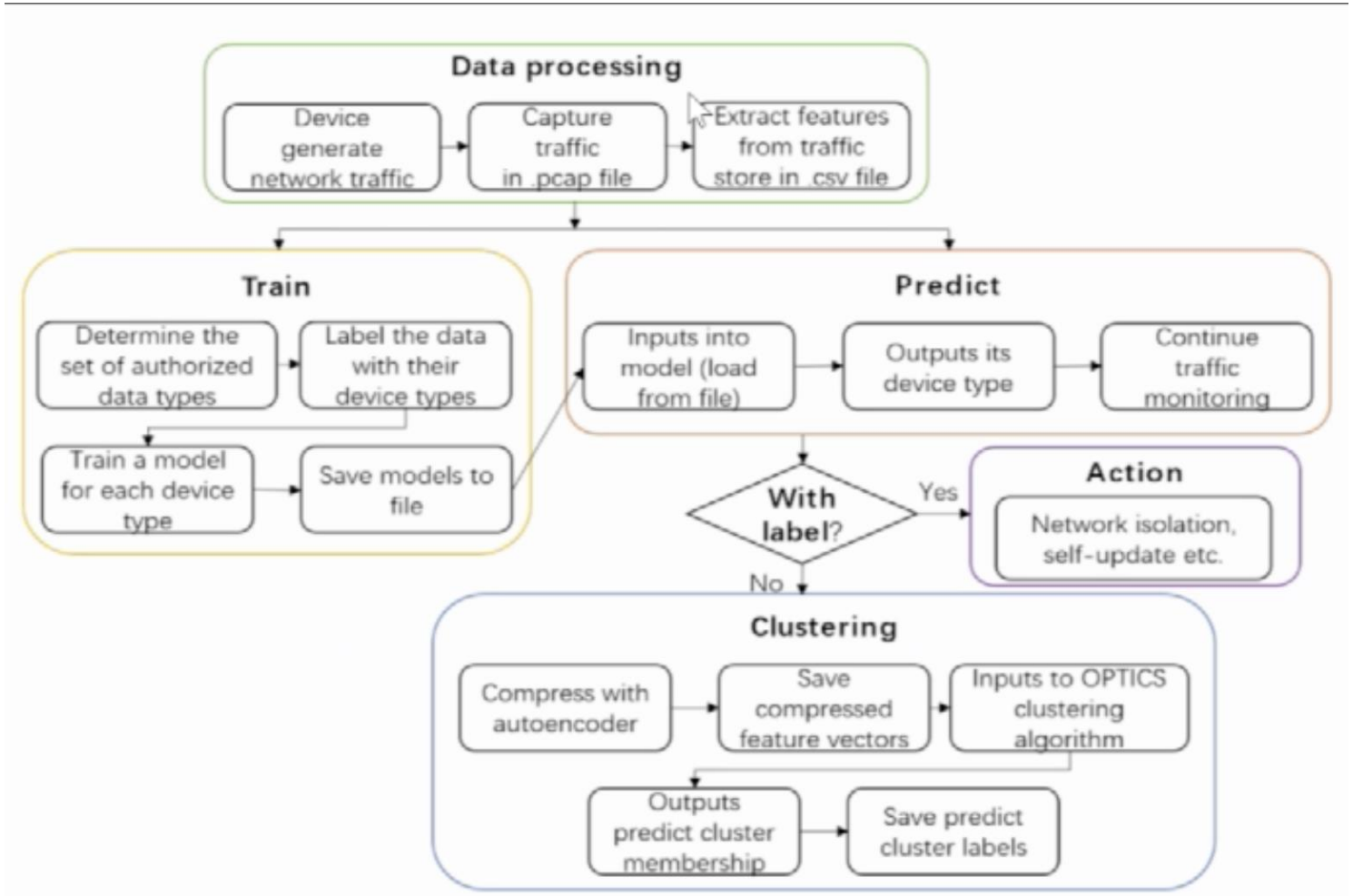
Source: B. Anderson el. al, "Machine learning for encrypted malware traffic classification: Accounting for noisy labels and non-sta vrity"ACM SIGKDD, 2017.

后流量加密时代 IoT 设备指纹识别: Non-Payload+ML

➢ 用于识别IoT设备的Non-payload特征+ ML的方法在学术研究和产品落地都已非常成熟。

➢ 已公开的Non-payload特征达数百种。

➢ 特征并非越多越好，有些受网络环境影响，易概念漂移。



Source: . Bao, B. Hamdaoui, and W. -K. Wong, "oT Device Type Identification Using Hybrid Deep Learning Approach for Increase Security" IWCMC 2020.a

后流量加密时代 IoT 设备指纹识别: 特征空间扩展

MORESHI POWERPOINT

从时域到频域

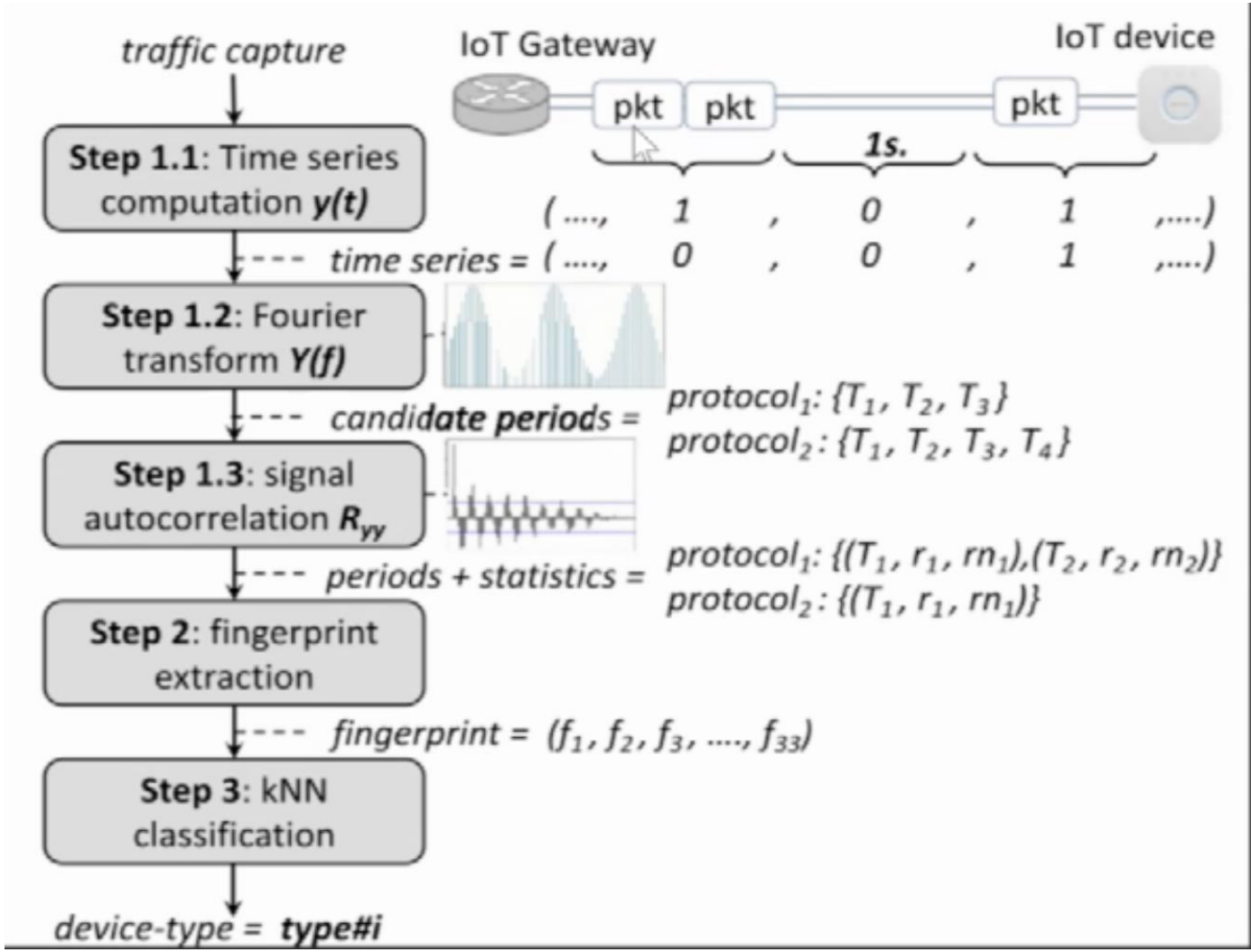
①以Flow为单位拆分IoT设备流量

②以秒为单位抽样时间序列

③运用DFT (离散傅里叶算法)分析识别周期性信号

④对周期性信号进行时间窗统计

⑤运用机器学习方法聚类



Source: s. Marchal et. al, "AuDl: Toward Autonomous /oT Device Type ldentificatian Using Periodic Communication, " IEEE Journ' 1 SelectedAreas in Communications, 2019.

基于流量变形的反识别技术

MORESHI POWERPOINT

为什么研究IoT设备流量指纹反识别技术?

- 保护IoT设备:不暴露IoT资产，不引起攻击者的注意，例如工业设备的启停。
- 保护隐私: IoT设备流量模式可能引发发隐私泄露,例如健康设备数据传输模式。
- 测试指纹识别技术的有效性:是否能够轻易绕过，导致无法正确识别设备。

基于流量变形的反识别技术

MORESHI POWERPOINT

- 目的:绕过IoT流量指纹分析模型
- 本质:流量操纵扭曲特征向量
- 约束:不改变流量承载功能
- 建模:最优化问题

$$\begin{aligned} \arg\min_{m \in \mathcal{M}} \quad & Fingerprint(\mathcal{F}(m(traffic))), \\ \text{s.t.} \quad & \mathcal{B}(m(traffic)) = \mathcal{B}(traffic), \\ & Fingerprint(\mathcal{F}(traffic)) = 1 \end{aligned}$$

传统流量变形技术: 利用系统差异

MORESHI POWERPOINT

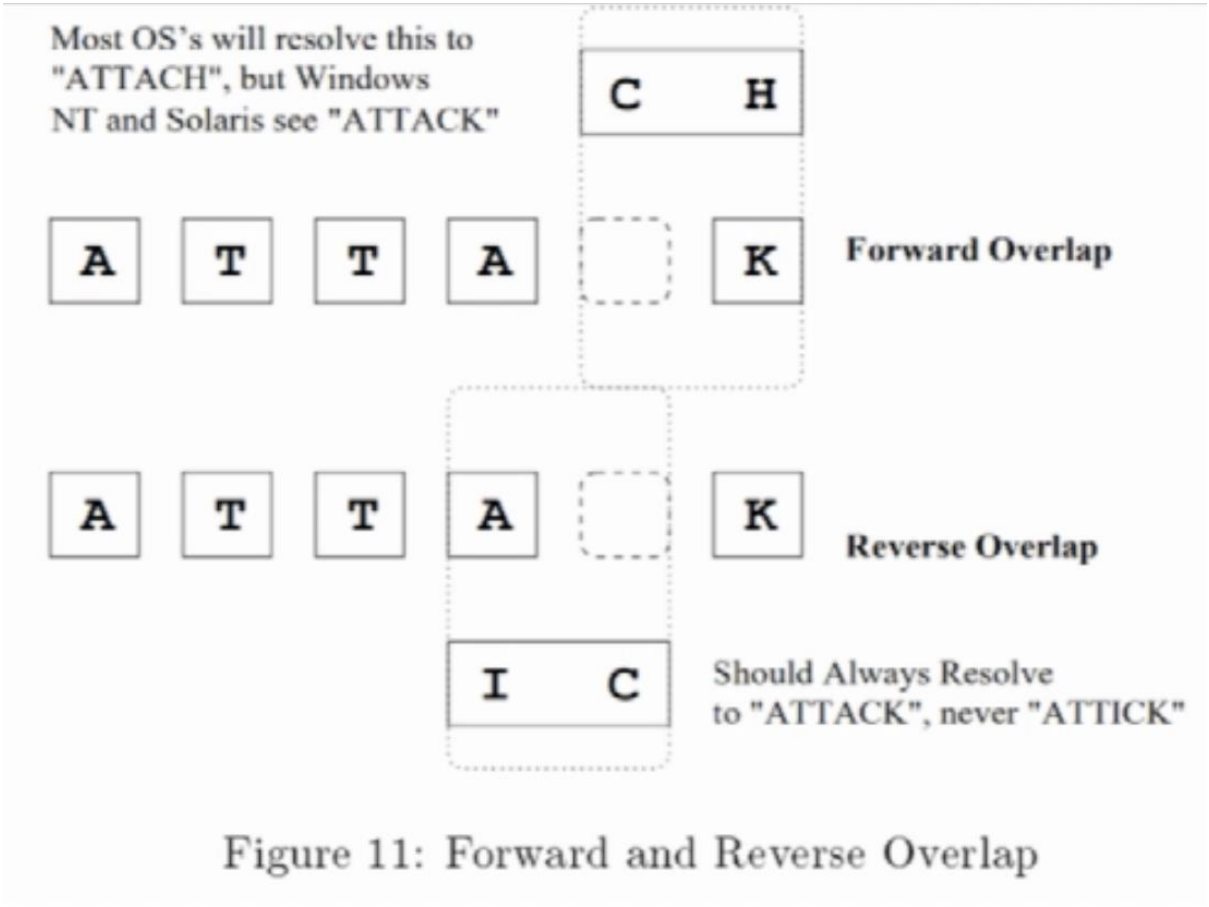


Figure 11: Forward and Reverse Overlap

ICS 35.030
CCS L 80

中华人民共和国国家标准

GB/T 20275—2021

代替 GB/T 20275—2013

信息安全技术 网络入侵检测系统技术要求和测试评价方法

7.4.1.2.x 防躲避能力

a) 测试方法：利用入侵检测躲避工具进行攻击，测试系统是否对攻击进行报警。

b) 预期结果：

1) 系统能够检测出经过分片、乱序、变形等之后的安全事件；

2) 系统能够正确地报出经过规避的扫描 HTTP 事件；

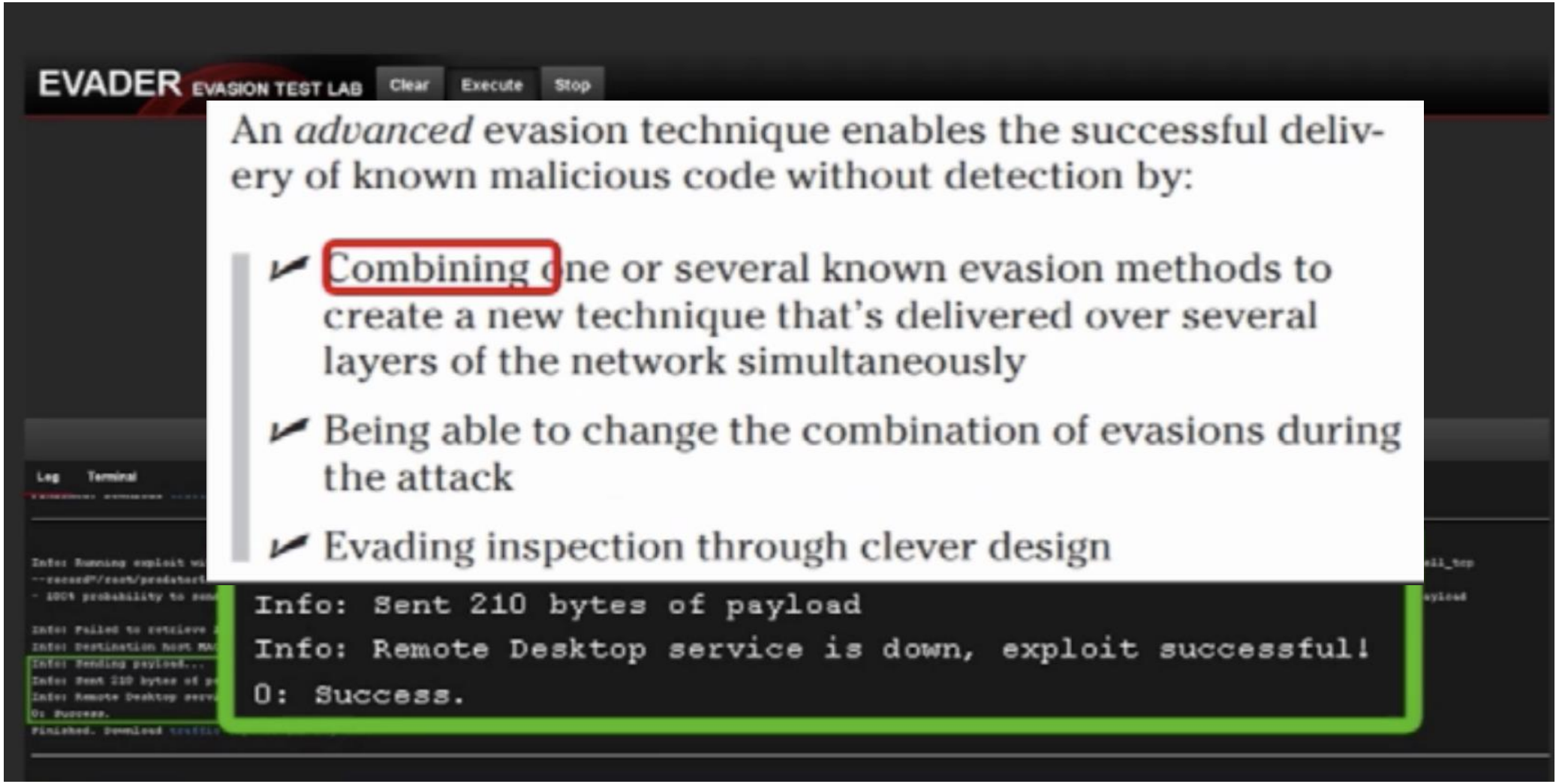
3) 系统能够正确地报出经过协议端口重定向的安全事件。

c) 结果判定：

上述预期结果均满足判定为符合，其他情况判定为不符合。

Source: T. Ptacek and T. Newsham, "Insertion, Evasion, and Denial of Service: Eluding Network Intrusion Detection, " in CEUR Work Proceedingse o.

传统的流量变形技术: AET (Advanced Evasion Tech.)



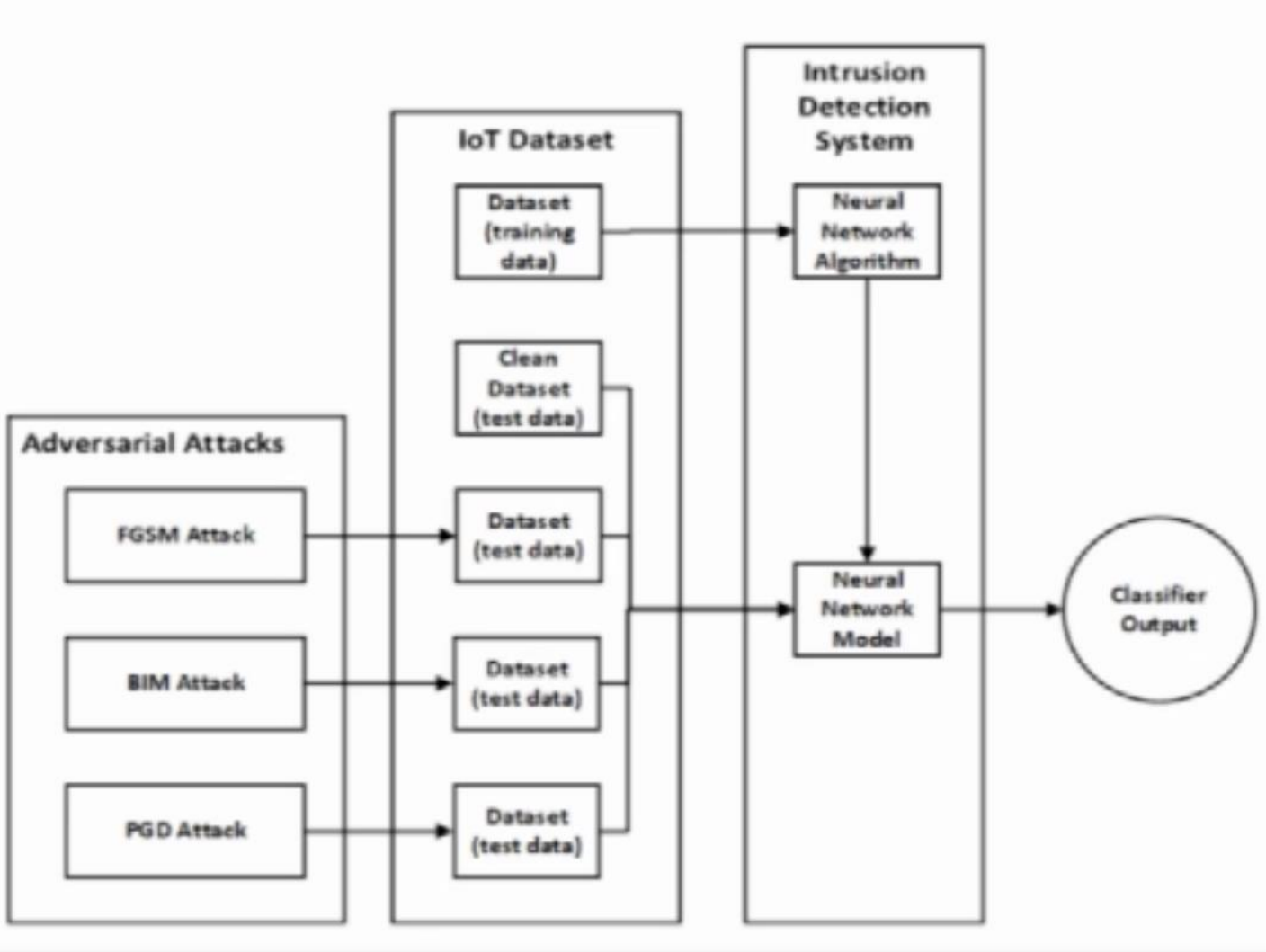
Sourece: K. Majewski, Advanced Evasion Techniques for Dummies. Forcepoint, 2015.

流量智能变形技术:反向优化

MORESHI POWERPOINT

➢ 方法:假设识别模型为白盒，特征定义已知，通过对目模型损失函数反方向优化，生成误导模型的对抗样本。

➢ 主流算法: FGSM/JSMA/CW



Source: o. lbitoye et. al, Analyzing Adversarial Attacks against Deep Learning for Intrusion Detection in GLOBECOM

流量智能变形技术: GAN (Generative Adversarial Network)

方法一:利用GAN的生成器对原始IoT流量特征进行复杂运算，最终使鉴别器(IoT识别模型)不能正确识别。

方法二:通过对原始流量加噪后的复杂运算(Generator)实现流量变形

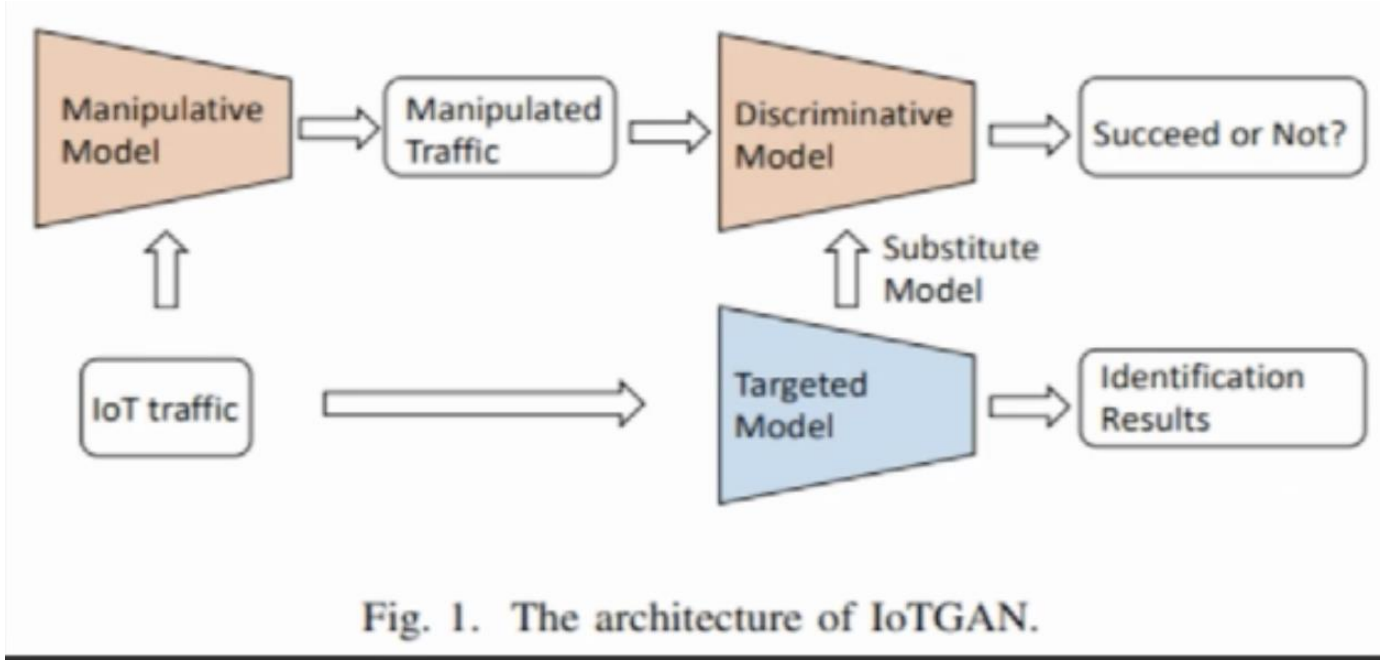


Fig. 1. The architecture of IoTGAN.

Source: T. Hou et.al, ": GAN Powered Camouflage Against Machine Learning BasedT 'ev be ootaos , " DySPAN

流量智能变形技术:从特征空间到流量空间

MORESHI POWERPOINT

Building IoT fingerprinting models and morphing Tools

Extract features

Preprocess

Raw dataset

Classifier	Test set	Accuracy	Recall	F1-score
AdaBoost	original	0.978	0.958	0.979
	mutated	0.776	0.567	0.724
Multilayer Perceptron	original	0.979	0.960	0.979
	mutated	0.507	0.049	0.093
Decision Tree	original	0.979	0.960	0.980
	mutated	0.487	0.010	0.020
Logistic Regression	original	0.950	0.941	0.949
	mutated	0.501	0.079	0.141
Random Forest	original	0.979	0.959	0.979
	mutated	0.646	0.316	0.481
Support Vector Machine	original	0.887	0.923	0.894
	mutated	0.802	0.760	0.799

道高一尺魔高一丈:高维空间争夺战

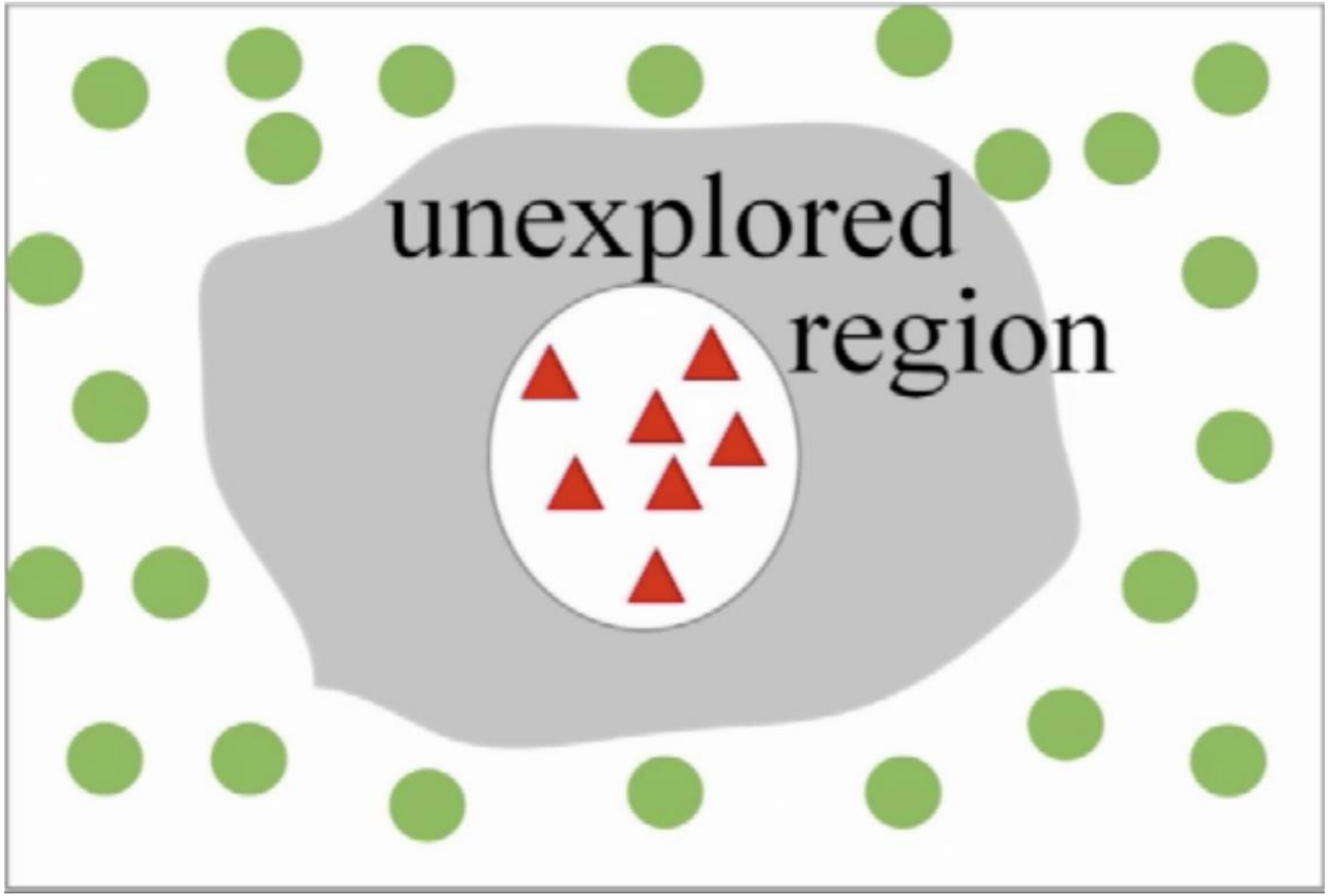
MORESHI POWERPOINT

Classifier	Decision Threshold	Accuracy	Recall	F1-score
AdaBoost	default	0.776	0.567	0.724
	optimal	0.991	1.000	0.992
Multilayer Perceptron	default	0.507	0.049	0.093
	optimal	0.666	0.411	0.561
Decision Tree	default	0.487	0.010	0.020
	optimal	0.487	0.010	0.020
Logistic Regression	default	0.501	0.079	0.141
	optimal	0.498	0.284	0.370
Random Forest	default	0.646	0.316	0.481
	optimal	0.999	1.000	0.999

基于自适应阈值的流量识别防御算法

道高一尺魔高一丈:高维空间争夺战

MORESHI POWERPOINT



保证设备识别性能的前提下占领稀疏空间

北京大学
PEKING UNIVERSITY

谢谢观看

2022.08.14