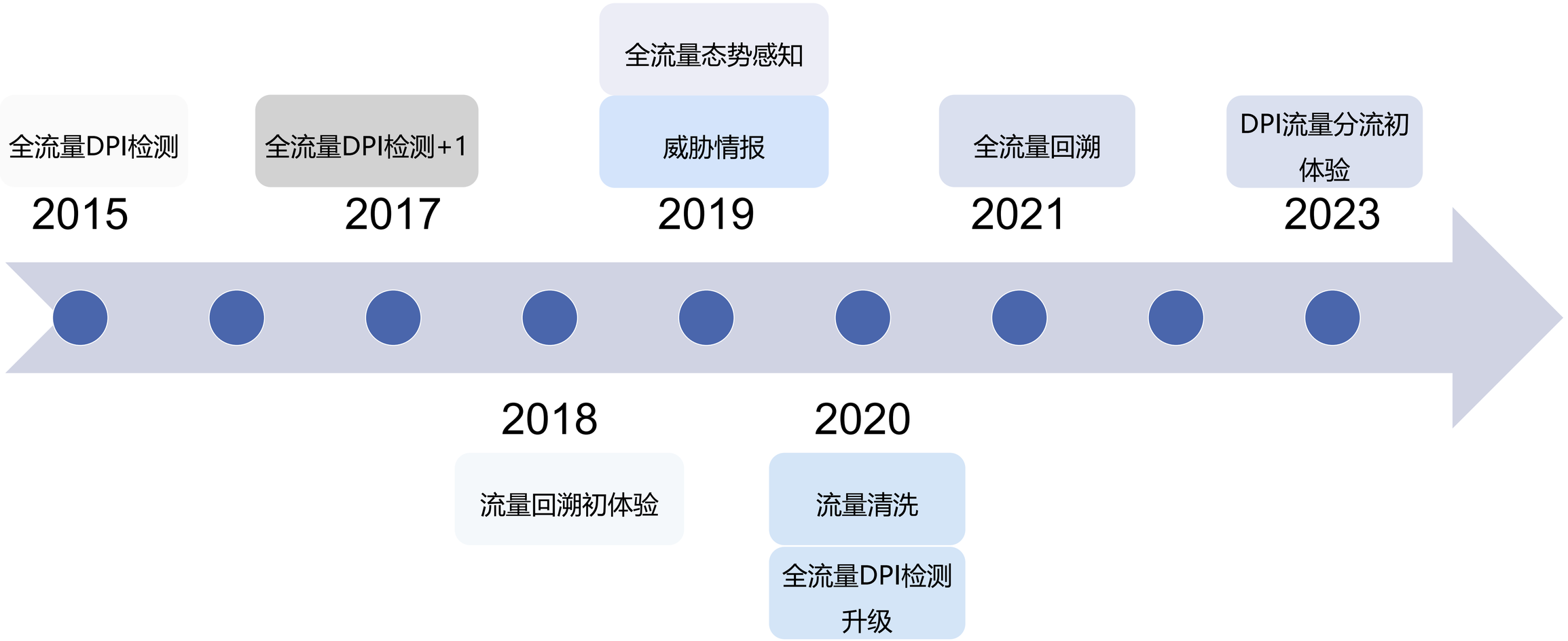
全流量回溯结合APT建模分析技术

姚星昆
清华大学

工作的演进


全流量DPI检测
2015
全流量DPI检测+1
2017
全流量态势感知
威胁情报
2019
全流量回溯
2021
DPI流量分流初体验
2023
流量回溯初体验
2018
流量清洗
全流量DPI检测升级
2020

网络全流量溯源及分析主要功能

存储容量

- 原始流量存储周期：14天-21天
  - 攻防演练
- 会话日志及协议元数据存储周期：6个月
  - 网络安全法

功能

网络流量采集

- 实时采集及**离线数据包导入**
- 支持IPv4和IPv6协议

网络流量存储

- **全流量存储**（原始流量、各类详单）
- **数据包检索**
- 存储策略（全量存储、截断存储、不存储）
- 数据包压缩、数据包加密

网络流量分析

- **对IPv4、IPv6流量分类分项统计**(会话、应用、协议、访问关系等)
- 应用层协议元数据解析 (HTTP、DNS、IMAP、ICMPv6、DHCPv6等)

网络安全分析

- **攻击特征检测及研判**（攻击尝试、攻击成功、主机失陷外联 ）
- **规则检测及威胁情报匹配**(域名、IP、URI、MD5等)

网络流量溯源

- 数据包**内容匹配**(ASCII值、16进制、中文、正则)
- **一站式IP回溯**
- **关联检索及下钻分析**（会话日志、协议元数据、数据包、访问关系）

网络流量取证

- **数据包下载**
- **历史流量快速回放**
- 会话日志、协议元数据外发

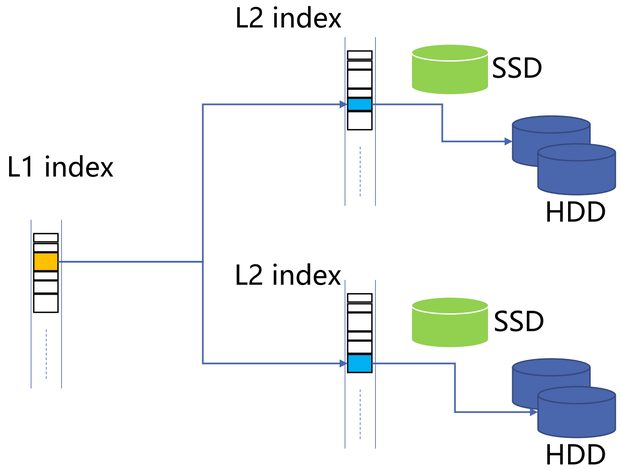
技术难点--数据包的存储与检索

数据包存储

针对数据包特点，采用磁盘预分配、顺序写、大块文件管理、按时间线回收的落盘处理方式，可极大的降低磁盘IOPS。同时带来三个优势：

- 读写性能大幅提升
- 降低机械硬盘失效率
- 不会越用越慢

数据包检索



300TB/s极速索引：

- L1索引快速锁定数据块
- L2索引精确定位数据包
- 列式存储

高速读取：

- 大块落盘
- 流聚合

- 有效容量为500TB+，假设数据包平均大小为500个字节，则设备可容纳1万亿个数据包，遍历这些数据包为秒级别
- 数据包索引维度：时间、五元组、DPI
- 性能与存储容量解耦

主流APT防御技术分析

现有主流高级威胁防御架构及应对策略分析

◆**基于威胁情报的检测方法**的有效性很大程度上依赖于IP地址、 域名、URL、 样本 HASH 等 IOC 情报的及时性和准确性. 在 APT攻击中攻击者往往同时注册许多基础设施并经常变化对 IOC 时效性要求很高. 实际应用中用户会面临海量的误报很难从中筛选出真正有分析价值的攻击线索。

◆**基于通信数据包规则的检测方法**同样面临误报率过高的问题。由于该检测方法是对某个具体的通信数据包制定的规则大流量环境中会存在许多与制定规则偶合的数据包产生大量误报。实时流量检测方式必须在通信会话处于活跃的短时间内产生告警结果无法完成流程复杂和计算量高的分析动作这使得目前基于通信数据包规则的检测方法的有效性大打折扣。

APT攻击流量检测机制

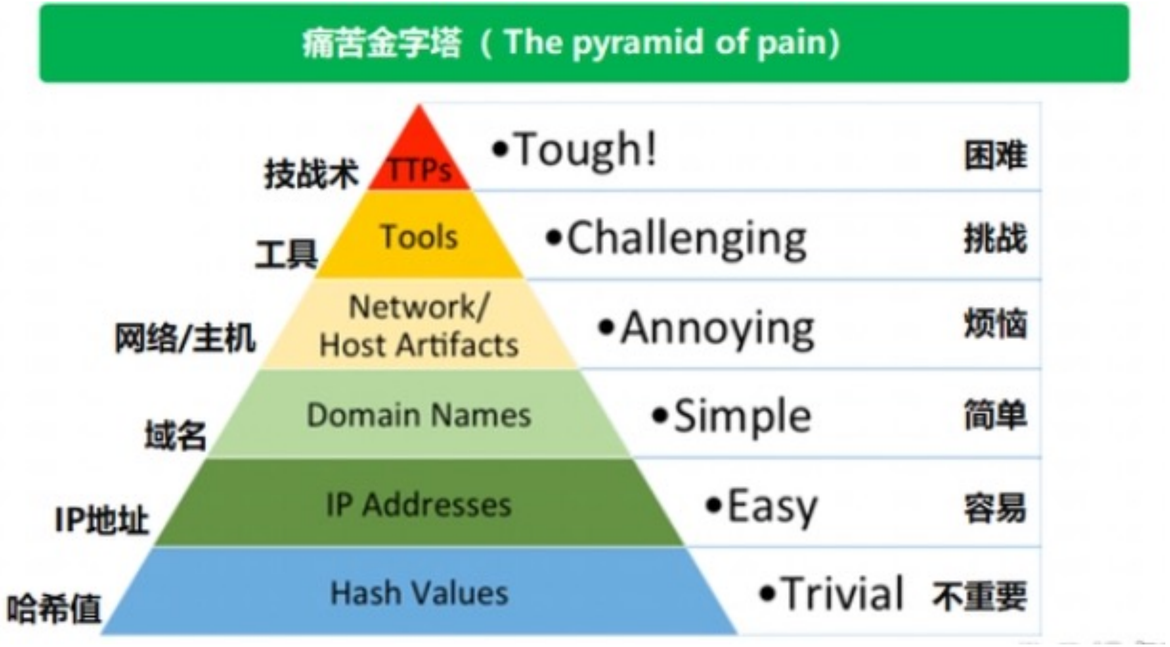
APT攻击都具备“*高度隐蔽和定向持续*”攻击特性。流量检测必须具备全流量全要素采集和多维度分析能力。

全流量全要素采集 → 多维度精准检测 → 高级威胁预警与溯源

- 全流量全要素流量数据采集分析与存储，方便后续数据分析、核查取证及攻击溯源；
- 通过精准APT威胁情报及多种攻击行为模型分析，形成对恶意攻击流量的发现及预警；
- 通过对APT攻击行为的发现预警、核查取证及追踪溯源，提高APT攻击防御能力；

攻击武器流量检测技术

为了躲避传统的基于样本和情报的检测手段，越来越多的APT组织使用0Day资源、无文件攻击技术、供应链攻击方法等等，导致攻击检测效率非常低效。基于***IOC+TTP的新一代流量综合检测技术***，本质是结合了情报和数据包规则两种检测方法各自的优势，并通过引入情报关联挖掘和预判技术、面向会话的TTP模型深度分析(***攻击武器密流量处理***) ，有效解决当期APT攻击检测存在误报率高、检出率低等难题。



基于流量元数据的单一TTP规则匹配

- 根据TTP构造的包检测规则
- URI模式
- DGA模式
- 包长/包结构/包载荷
- SSL/SSH指纹特征
- ……

基于历史流量的复杂TTP行为模型匹配

- 根据TTP构造的会话检测插件
- 数据加密特征
- 控制命令交互特征
- 心跳激活交互特征
- ……

TTP模型

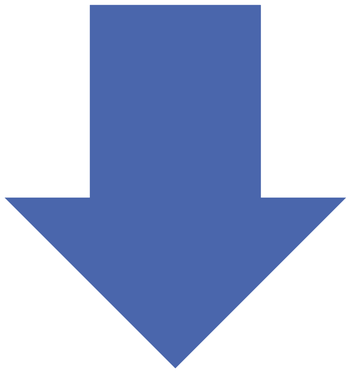
TTP规则

IOC规则

实时情报匹配

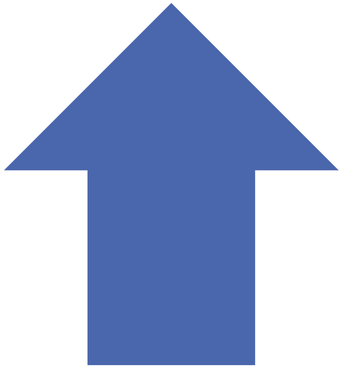
- 核心IOC+拓展IOC
- IP/Domain/Hash/X509/Email
- ……

下一步工作



通过DPI分流设备降噪流量

提高设备使用率



所有安全设备日志信息整合

交叉验证提高准确度

谢谢