流量识别与分析技术

清华大学

诞生前夜：流量规模激增

背景：广域网承载的流量规模持续上升

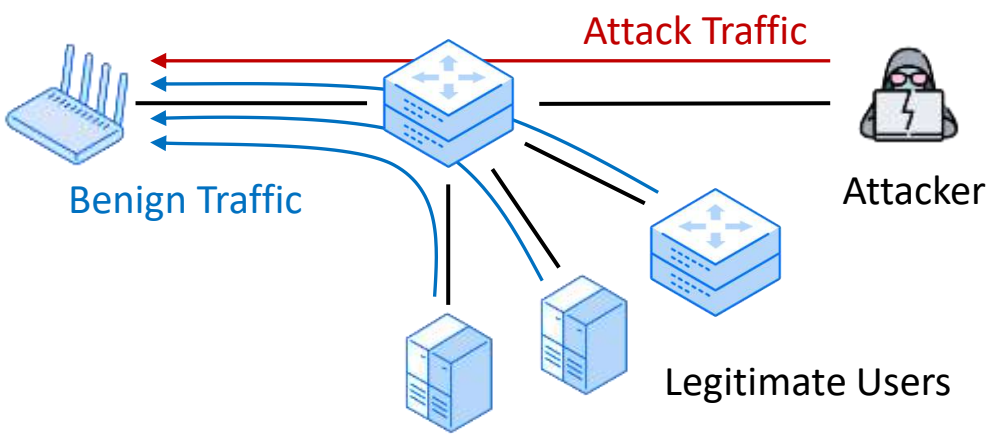
- 广域网（WAN）是互联网之躯干，其连接跨地理区域的局域网，目前广域网流量已突破 **80Tb/s** 规模*
- 广域网转发海量**合法流量**，例如，抖音视频流量
- 广域网也承载了**恶意流量**，例如，DDoS流量，网络扫描，侧信道攻击，和恶意软件通讯流量

互联网流量是网络空间中信息的传递形式

* https://www.de-cix.net, DE-CIX.

广域网核心路由器的地理位置示意图

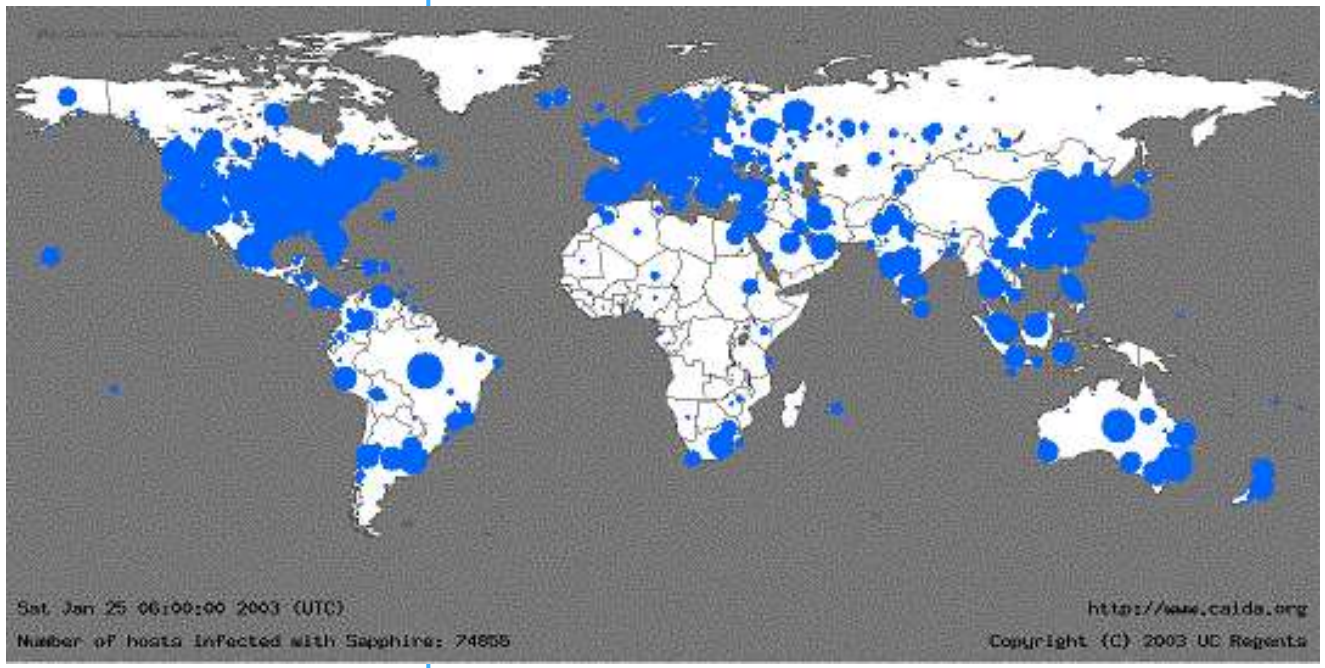
合法流量攻击流量被一同转发



诞生前夜：流量传播计算机病毒

背景：现代计算机病毒依赖网络流量进行传播

复习与回顾： **Slammer** 是一款DDoS恶意程序，它利用SQL Server漏洞，通过1434端口感染SQL Server，通过被感染的SQL Server再大量的散播**携带病毒二进制文件的流量**，使得SQL Server 无法正常作业或宕机



06:00:00 UTC, January 25, 2003
(攻击发生后)

诞生前夜：恶意流量造成经济损失

背景：网络发起的攻击造成严重损失

Kaspersky: 2017年仅DDoS流量造成平均损失达每企业 **$2.3M**

Cisco: 2018年恶意流量造成平均损失达每企业 **$3.86M**

可否通过分析流量、识别流量、从而拦截攻击流量阻断攻击?

诞生前夜：流量传递隐私

背景：用户隐私信息通过加密流量进行传播

- 另一方面，流量不仅仅携带了**恶意数据**还包含了用户的**隐私数据**
- 传统流量携带明文数据，攻击者可以通过捕获数据包进行流量窃听
- 因此现在用户普遍采用数据加密协议进行通讯，目前加密流量已经占据了相当大的比重

May 2019, 94% of all Google web traffic is encrypted.

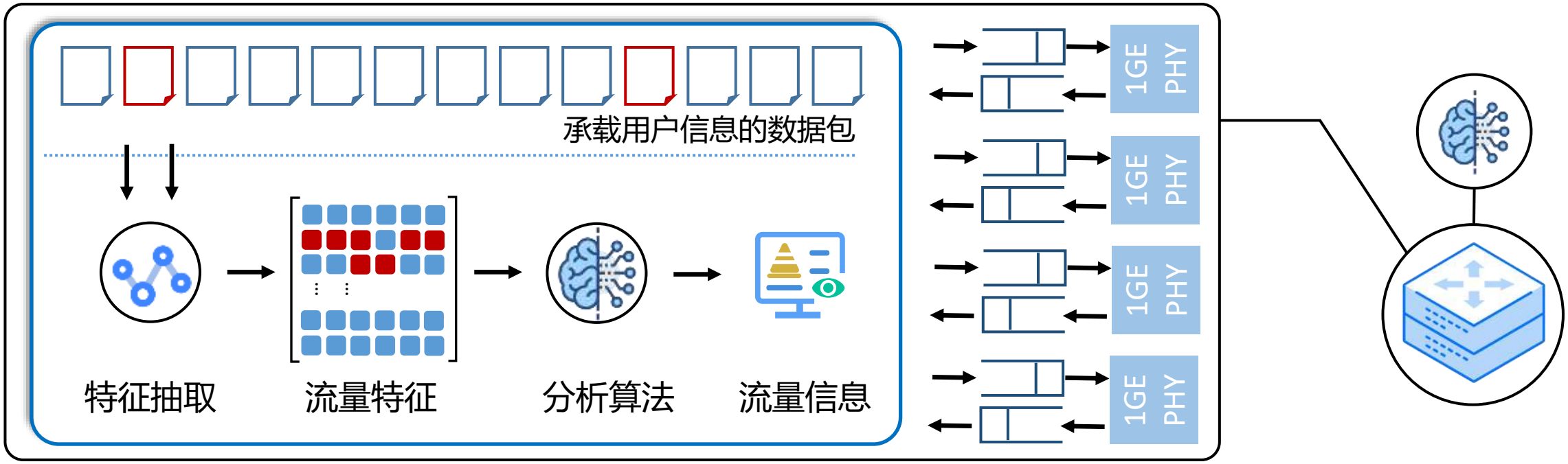
Nearly 80% of web pages loaded by Firefox use HTTPS.

Over 98% Alexa top 1k websites support HTTPS.

流量分析问题定义

科学问题：分析流量的特征，推测通讯相关的信息

- **流量特征定义：**对二进制形式的网络流量进行变换，得到具备语义量化的特征，精准刻画流量的表示
  - 例如流中的包数量，每一个包的长度等等
- **被推测信息是什么？**
  - 用于攻击的信息 + 用于防御的信息



流量分析技术是双刃剑，可以服务于攻击，可以服务于防御

流量分析的系统的架构和三要素和分类法

流量分析系统的三要素：流量特征、分析算法、分析的目标

- **流量特征：**系统采用了何种方法表示流量
  - 负载特征：直接将数据包当中的内容作为特征
  - 统计特征：例如对于数据包头部内容设计统计量

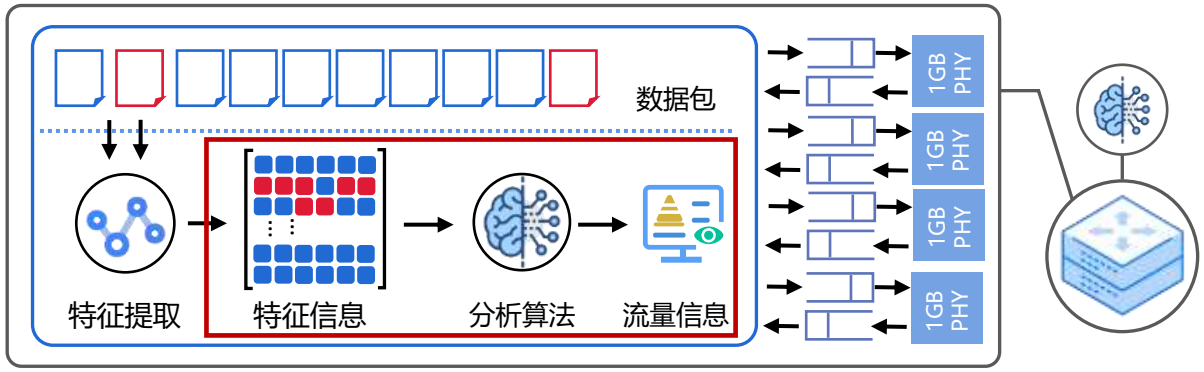
其中，流量特征是影响系统性的关键

- **检测算法：**固定规则和机器学习算法
  - 固定规则：由人类专家设计
  - 机器学习：算法根据已知数据训练
- **分析的目标：**
  - 窃取隐私：流量分析攻击
  - 保护用户：流量的识别和防御

流量分析系统的分类法

分析目的	检测算法	特征	实际应用
保护用户	固定规则	统计特征	攻击流量防御系统 (3.1-3.3)
		负载特征	传统固定规则入侵检测 (1.1)
	机器学习	包头特征	智能入侵检测系统 (2.1 - 2.5)
		负载特征	网站防火墙应用 (1.3)
隐私窃取	固定规则	包头特征	网络侧信道攻击 (第八章)
		负载特征	流量窃听 (第八章)
	机器学习	包头特征	流量分析攻击 (4.1 - 4.3)
		负载特征	------

我们通过三要素模型：
可以将各类流量分析系统放在同一框架下研究



本章的内容组织

攻击流量的检测和防御

从海量的数据包当中抽取流量特征，推断攻击流量是否为正常或者异常，从而实现检测和拦截攻击流量，保护大量互联网用户的安全性

流量特征隐私窃取

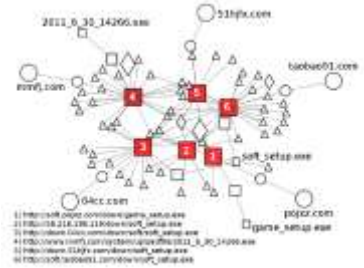
根据加密流量特征，预测流量通讯内容相关的隐私信息

负载特征流量识别

研究内容:

传统NIDS

网站防火墙

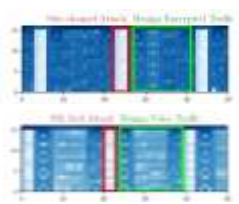


统计特征流量识别

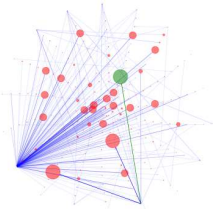
研究内容:

流粒度、包粒度特征

加密流量识别



流量频域特征



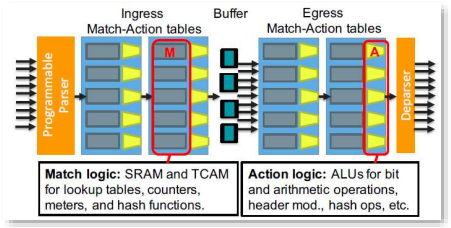
交互特征

攻击流量防御系统

研究内容:

传统匹配过滤机制

可编程数据面防御机制

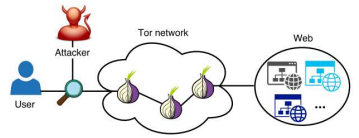


流量分析攻击

研究内容:

网站指纹生成

其他流量分析攻击



网站指纹

APP流量分类

未知攻击

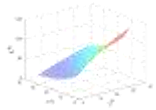
高效信息使用

未知攻击

未知流量模式

流量规模

海量流量匹配



窃取信息

高效信息使用

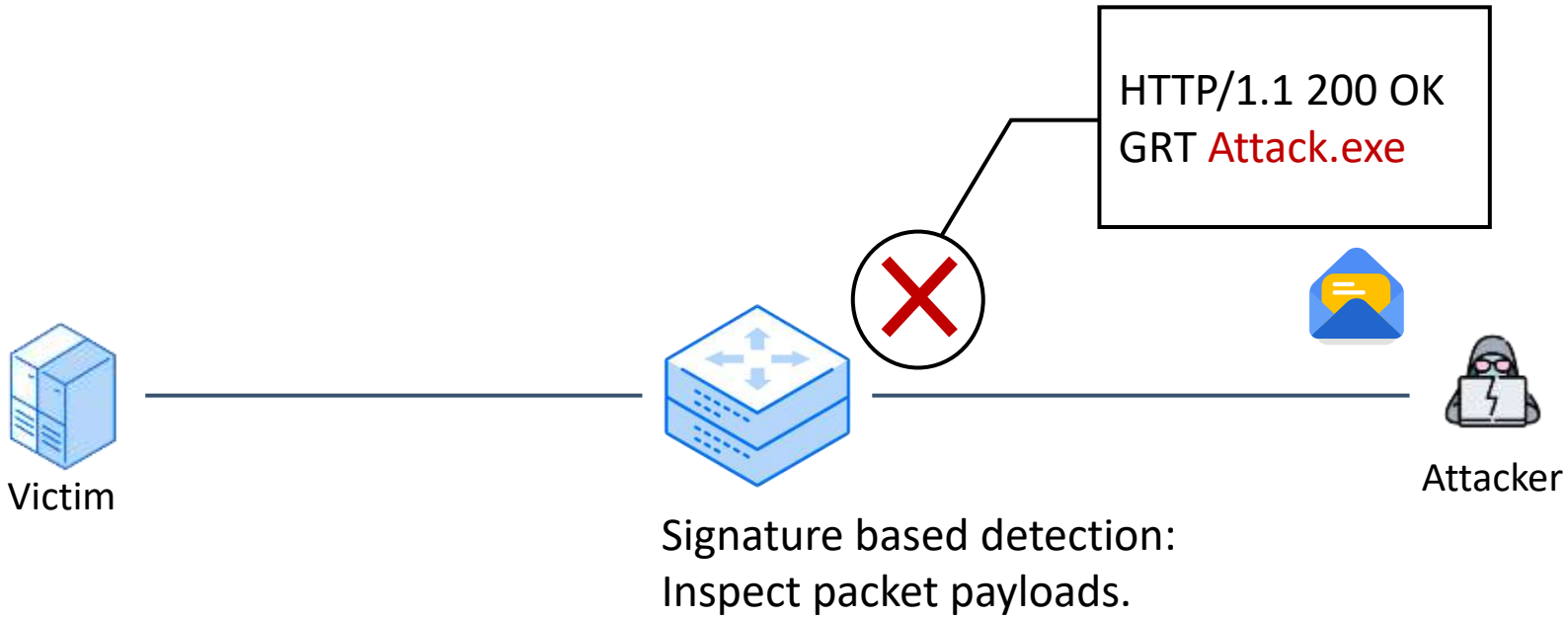
以流量分析系统的三要素为切入点，掌握系统化分析流量识别系统的方法

第1节 负载特征驱动的流量检测

- **基于数据包负载匹配的传统检测方法**
- 基于人工智能的流量检测方法总体框架
- 基于人工智能的负载分析方法：网站防火墙
- 基于人工智能的负载分析方法：恶意软件检测

传统恶意流量识别方法：基于固定规则

利用固定规则对流量的**负载**进行**匹配**



基于固定规则的恶意流量识别应用案例

以Zeek为例，现代固定规则检测系统支持的三大类匹配方法

1. 对负载的统计特征进行匹配

示例：监控和标记异常活跃的主机

以下是一个 Zeek 脚本示例，该脚本监控网络连接，计算特定统计特征，如主机的连接次数，并标记那些异常活跃的主机

- **全局变量定义**

ip_connection_counts 存储了每个 IP 地址的连接次数

- **事件处理**

connection_established 事件在每个新的连接建立时触发。脚本增加该连接源 IP 的计数，并检查其是否超过定义的阈值，如果某个 IP 的连接次数超过阈值，系统将打印一条警告消息，并可以采取进一步的响应措施

```
# 定义一个表来存储每个IP地址的连接计数
global ip_connection_counts: table[addr] of count = {};

# 定义阈值，超过此阈值的主机被认为是异常活跃
const threshold: count = 100;
# 当新的连接被建立时触发此事件
event connection_established(c: connection)
{
    # 对源IP地址的连接计数加1
    ip_connection_counts[c$id$orig_h] += 1;

    # 检查该源IP的连接次数是否超过阈值
    if (ip_connection_counts[c$id$orig_h] > threshold)
    {
        print fmt("High activity detected: %s has %d
connections",           c$id$orig_h,
ip_connection_counts[c$id$orig_h]);
        # 这里可以执行额外的响应措施，如发出警报或进一步调查
    }
}
event zeek_done()
{
    for (ip, count in ip_connection_counts)
    {
        print fmt("%s made %d total connections", ip, count);
    }
}
```

基于固定规则的恶意流量识别方法小结

利用固定规则对流量的负载进行匹配

缺点1：依赖深度包检测技术，对加密流量不可用

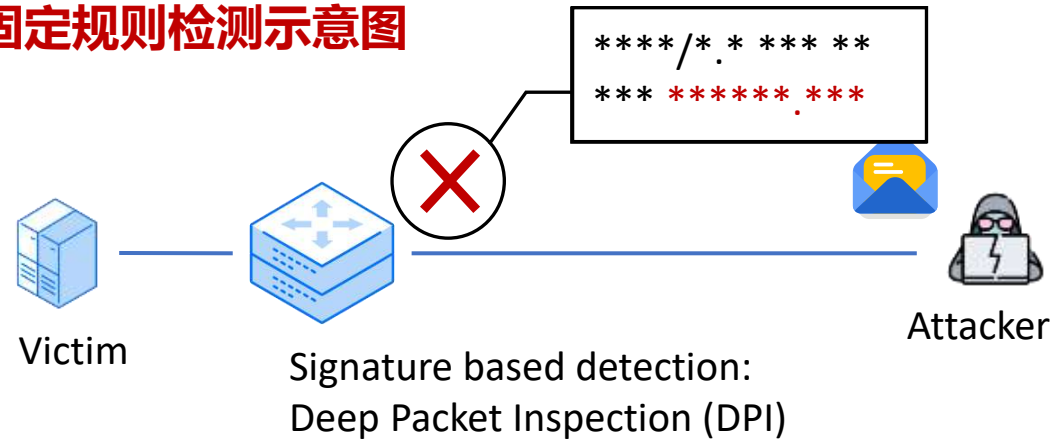
缺点2：大量过滤规则消耗计算资源

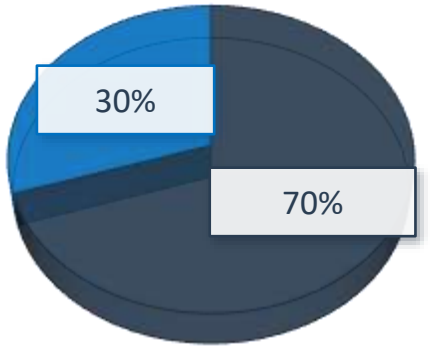
缺点3：逃逸策略躲避检测

缺点4：人工设计的固定规则仅可针对已知流量进行过滤，无法应对未知攻击 (Zero-Day Attacks)

固定规则检测资源开销远大于
检测效果的收益

固定规则检测示意图





Cisco: Over **70%** attacks were constructed by encrypted attack traffic.

第1节 负载特征驱动的流量检测

- 基于数据包负载匹配的传统检测方法
- **基于人工智能的流量检测方法总体框架**
- 基于人工智能的负载分析方法：网站防火墙
- 基于人工智能的负载分析方法：恶意软件检测

基于机器学习的恶意流量识别方法

重大飞跃：基于机器学习的恶意流量识别

› 约十五年前，研究人员提出采用机器学习方法来自动识别恶意流量

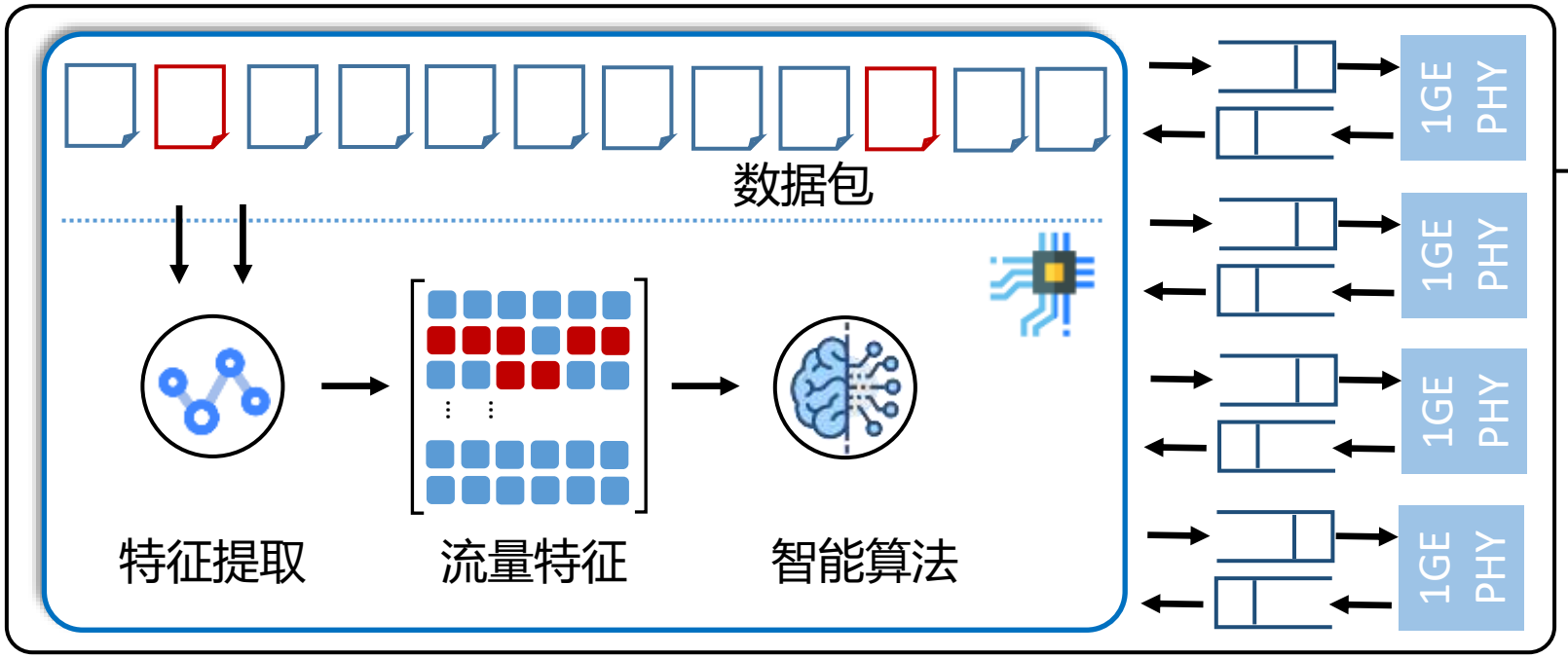
› 基于机器学习的方法主要包括**三个模块**：

» 1. 包解析; 2. 流量特征抽取; 3. 机器学习算法

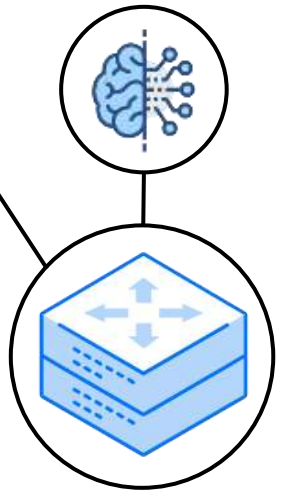
› 相比于固定规则检测，其优势有：

1. 借助机器学习泛化性，检测**未知攻击**
2. 相比于固定规则方案检测**准确度更高**
3. 不依赖人力设定固定规则，**低人工处理开销**

因而，其具有良好的应用前景和价值



智能流量识别系统架构



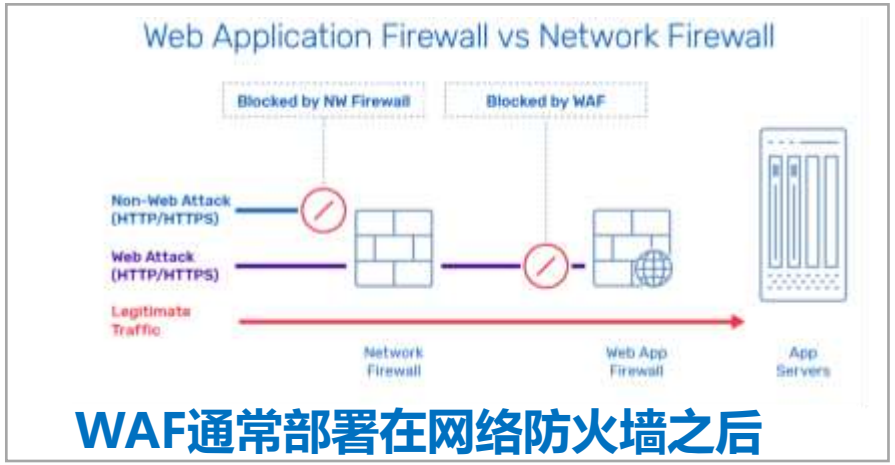
第1节 负载特征驱动的流量检测

- 基于数据包负载匹配的传统检测方法
- 基于人工智能的流量检测方法总体框架
- **基于人工智能的负载分析方法：网站防火墙**
- 基于人工智能的负载分析方法：恶意软件检测

基于人工智能的负载分析方法：网站防火墙

- **网站应用防火墙 (Web application firewall)** 是一种专门用于监控、过滤和阻止传入和传出Web应用程序的恶意流量的安全系统
- 它是在传统网络防火墙的基础上发展起来的，专门针对HTTP应用程序进行优化，从而防止各种在线攻击，如SQL注入、跨站脚本等

- **网络设备：**作为网络中的物理或虚拟设备，直接放置在网络中以保护后端服务器
- **集成组件：**作为特定Web服务器，例如 Apache、Nginx，的插件或模块直接部署
- **云服务：**作为服务提供商提供的托管解决方案，通常称为云WAF或WAF-as-a-Service



WAF通常部署在网络防火墙之后

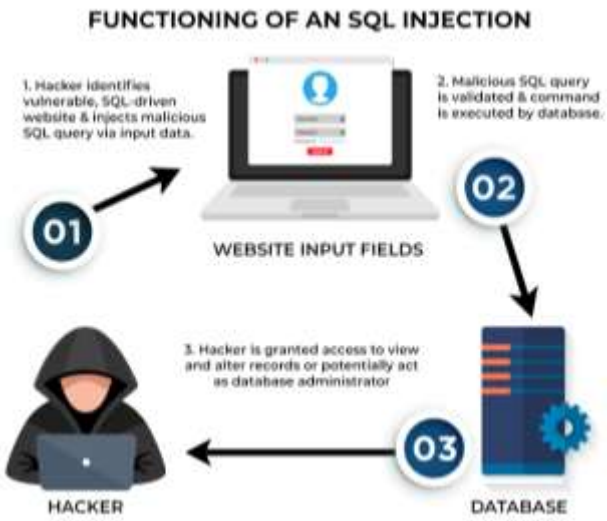
通常WAF可以通过网站密钥对流量解密，直接对明文进行检测

基于人工智能的负载分析方法：网站防火墙

针对的代表性攻击

1. **针对Web站点的控制流劫持**：攻击者通过异常的负载触发各种网站服务器软件，例如Apache/Nginx的内存漏洞，进而控制站点
2. **SQL注入攻击**：攻击者通过恶意操纵SQL查询语句的输入泄露用户数据、删除数据、获取数据库管理员权限等，这一攻击将在十四章详细介绍
3. **机器人用户**：频繁模仿用户行为访问站点，例如为了操纵访问热度

SQL注入攻击流程



技术评价

优势：1. 可以低速检测隐蔽攻击；2.相比于固定规则NIDS，具备一定泛化能力

劣势：1. 深度包检测带来的隐私问题；2. 对加密流量彻底失效；3. 深度包检测的效率不足，导致实时性无法保障

通常WAF通过分析HTTP包当中的负载判定是否未触发浏览器漏洞的流量

第1节 负载特征驱动的流量检测

- 基于数据包负载匹配的传统检测方法
- 基于人工智能的流量检测方法总体框架
- 基于人工智能的负载分析方法：网站防火墙
- **基于人工智能的负载分析方法：恶意软件检测**

基于人工智能的负载分析方法：恶意软件检测

任务背景：恶意软件的模块化设计

恶意软件的下载行为：

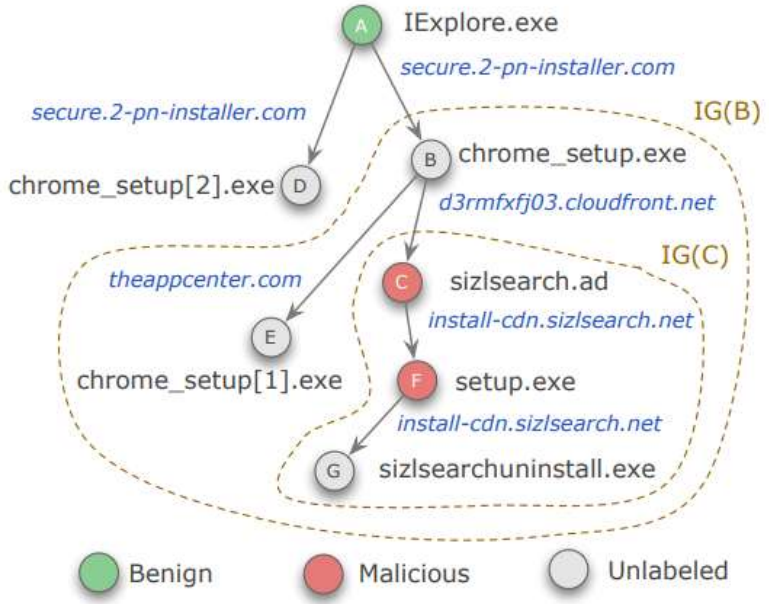
恶意软件在初次感染后，从远程服务器下载额外组件或功能模块的过程；使得恶意软件能够保持较小的初始体积，**降低被安全软件检测的机会**，并根据需要动态地增强其功能

恶意软件的模块化设计：

恶意软件作者常常采用**模块化设计**，使得基本的恶意软件框架可以根据需要下载额外的功能组件；这些组件可能包括键盘记录器、后门、加密模块、窃取特定类型数据的工具等

早期恶意软件检测：

检测恶意软件资源分配服务器，识别恶意软件侵染事件，即对下载流量进行分析



恶意软件进程从多个站点下载资源

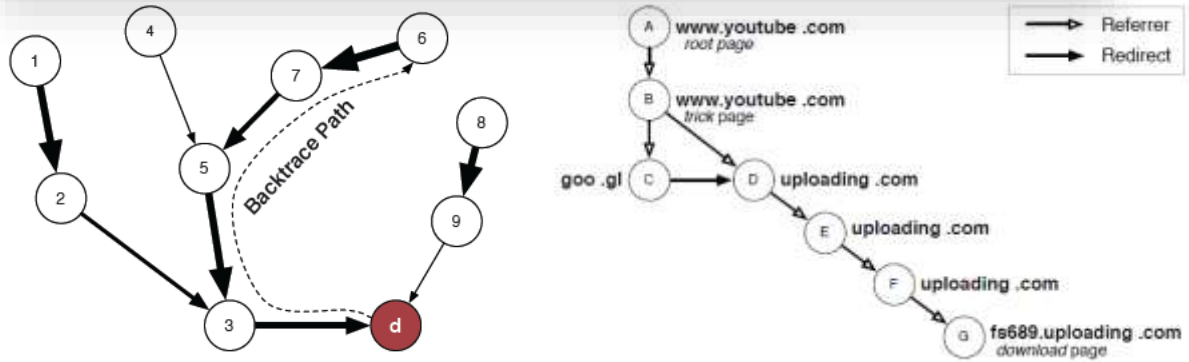
基于人工智能的负载分析方法：恶意软件检测

代表方案: WebWitness

分析目标：通过HTTP请求访问之间的关系，识别恶意软件的下载流量，并追溯潜在产生恶意软件地服务器

流量特征：提取请求中的URL，并将URL表示成为图，进一步地在图上抽取特征，例如跳数关系

分析算法：有监督统计机器学习算法



分析URL，得到资源下载图，检测分发恶意软件资源的服务器

技术优势：较早地实现对于恶意软件站点服务器的演化关系进行研究

有效检测潜在的攻击站点

大量的恶意软件连接隐藏在社交网络上：恶意软件确实通过射箭网络传播

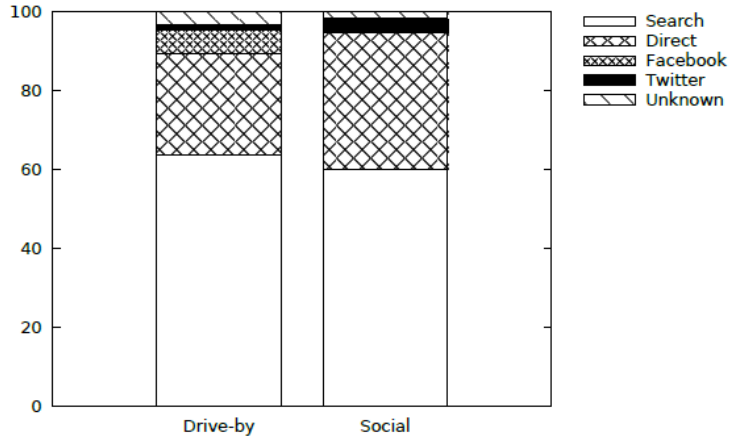


Figure 7: "Root" of malware download paths.

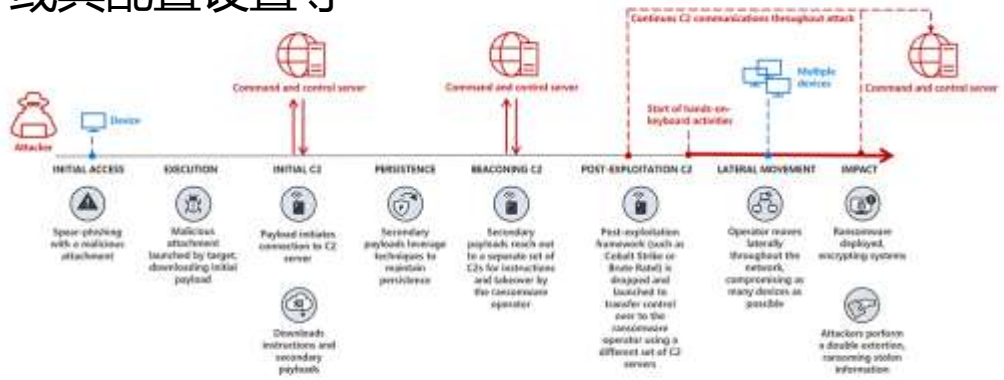
WebWitness: Investigating, Categorizing, and Mitigating Malware Download Paths, USENIX Security, 2015.

基于人工智能的负载分析方法：恶意软件检测

恶意软件在通讯过程当中的迁移性

恶意软件C2

Command and Control，指挥控制是指恶意软件与其操作者（通常是黑客或犯罪分子）之间的通信机制；攻击者远程控制感染了恶意软件的计算机或网络，这些指令可以包括下载和执行文件、收集敏感数据、传播恶意软件到其他设备以及更新恶意软件或其配置设置等



C2服务器参与高级持续性威胁的各种阶段

恶意软件C2的实现：

C2通信可以通过多种方式实现，包括但不限于：

1. **HTTP/HTTPS**：使用常规的网页浏览协议进行通信
2. **DNS**：利用DNS请求来传递命令或数据
3. **P2P**：不通过固定服务器，而是通过网络中的其他感染设备进行通信
4. **社交媒体、电子邮件**：通过这些常用服务来传递指令

注意到C2服务器的位置并非长期不变
恶意软件如何找到C2服务？

基于人工智能的负载分析方法：恶意软件检测

代表方案: BotSniffer

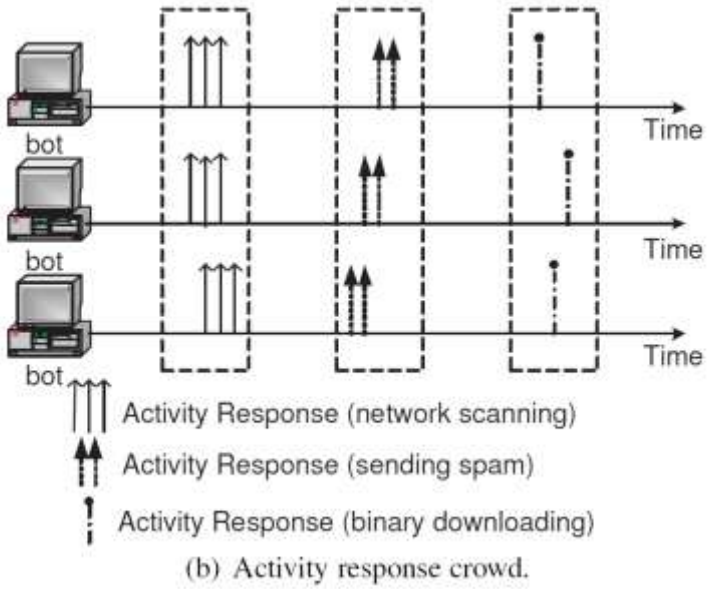
分析目标：通过HTTP流量，检测出C2信道流量，进而定位受感染的主机

流量特征：提取请求中的URL，并将主机之间的访问事件上的特性进行匹配

分析算法：无监督的贝叶斯学习方法，检测相似的访问模式为C2信道的流量

技术优势：较早地考虑恶意软件侵染受害节点之间的流量的相似性，检测C2服务器和受害主机，是早期恶意软件流量检测的代表性工作

恶意软件C2服务器访问模式存在时间上的相似性



由于恶意软件在散播的过程当中采用相似的二进制文件并且C2服务器的源代码也是相似的这导致其产生相似的流量模式

BotSniffer: Detecting Botnet Command and Control Channels in Network Traffic, ISOC NDSS, 2008.

第2节 统计特征驱动的流量识别方案

- **基于数据包粒度特征的检测方法**
- 基于流粒度特征的检测方法
- 编程网络设备上的攻击流量检测
- 加密攻击流量检测

基于包粒度特征的检测方法

技术方法

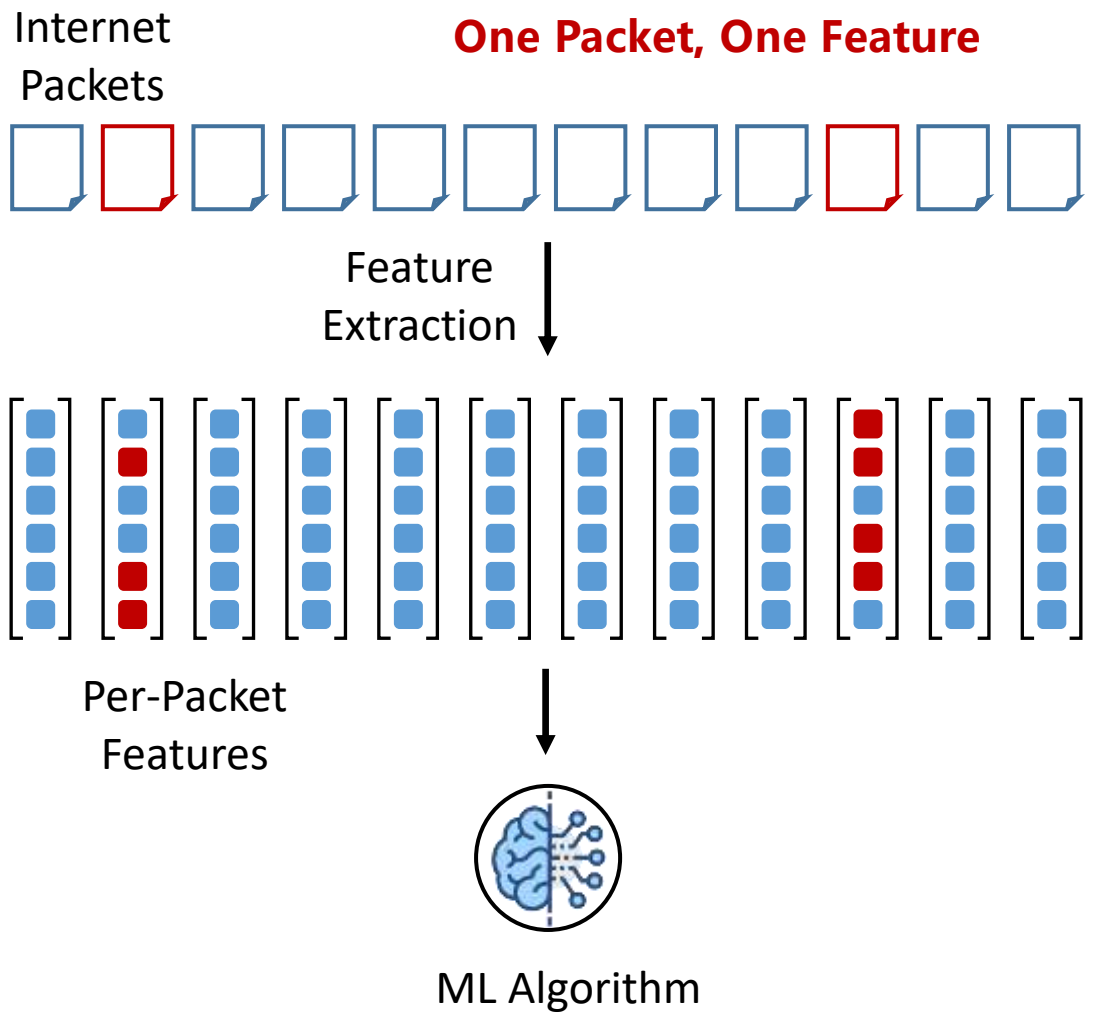
基于数据包负载内容的检测方法

为每一个数据包抽取一个特征向量，作为机器学习的输入，分类每一个数据包是否为正常的数据包

技术评价

优势：1. 不依赖数据包负载进行检测；2. 细粒度分析网络流量

劣势：1. 逐个数据包分析的效率问题，通常离线运行或工作在低负载网络下；2. 鲁棒性不佳，操纵数据包当中字段可以完成逃逸攻击



基于包粒度特征的检测方法

代表方案: Kitsune

分析目标：检测IoT设备的各种攻击流量，包含了洪泛攻击、指令注入、远程漏洞利用等等

流量特征：从数据包头部抽取了115个维度的特征，包含了流级别的统计信息

分析算法：采用聚合的自编码器，本质上是一个无监督模型用于检测未知的攻击策略，以115维特征为输入，以重构误差为输出

技术优势：最早的基于特征的无监督攻击检测方法是恶意流量检测的代表性基线方案

开放了源代码和数据集

具体采用的模型是多个自编码器的集成模型

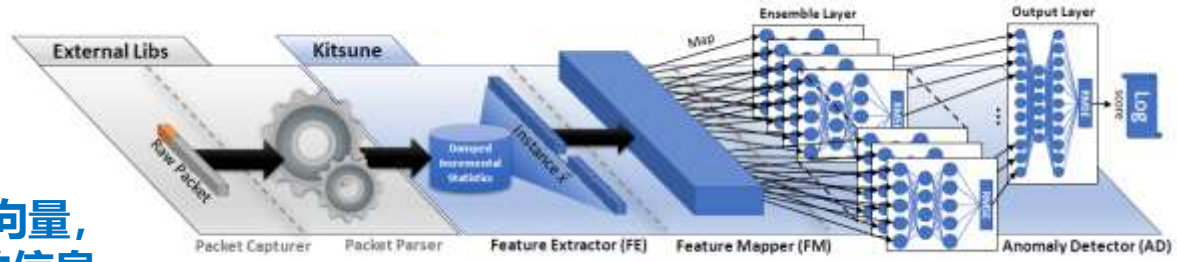


Fig. 3: An illustration of **Kitsune**'s Architecture.

Kitsune方法为每一个包构建一个特征向量，其中包含了数据流、地址等不同粒度的信息

开放讨论：这类方法可能存在哪些问题

Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection, NDSS 2018.

基于包粒度特征的检测方法：总结

对包粒度特征的检测方案的评价

1. 包粒度的特征检测相比于与传统负载特征分析的优点在于：

- 不依赖数据包负载信息：可以识别加密流量
- 采用的统计信息可以对多种流量提取：实现通用检测，识别多种攻击流量
- 不用维护流状态：一定程度降低了检测开销

2. 包粒度特征检测仍然存在的不足：

将数据包视为孤立个体，难以分析数据包之间的高层次的依赖关系

此外，逐一分析每一个数据包带来显著的性能开销，无法适应快速增长的流量规模

为解决这一问题，基于流粒度的检测方法被大量提出于应用

第2节 统计特征驱动的流量识别方案

- 基于数据包粒度特征的检测方法
- **基于流粒度特征的检测方法**
- 编程网络设备上的攻击流量检测
- 加密攻击流量检测

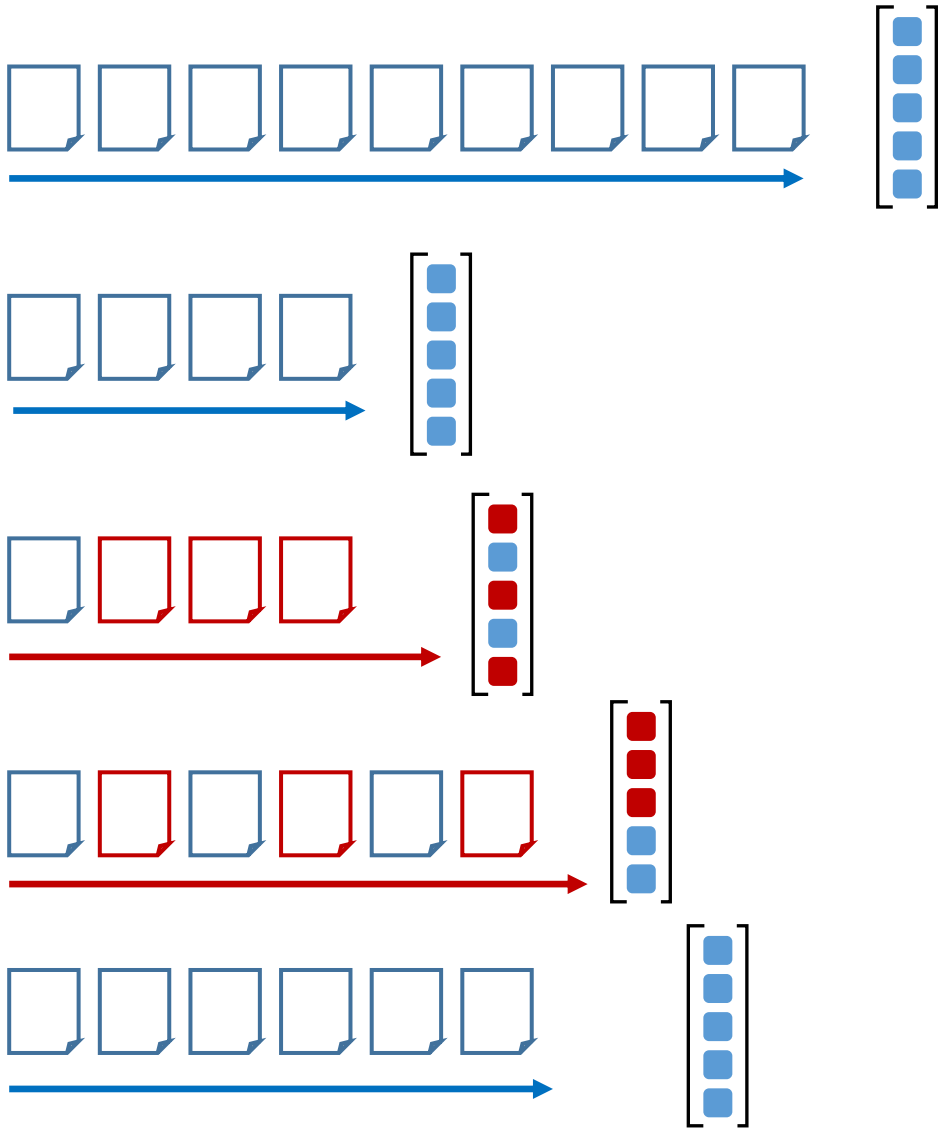
基于流粒度特征的检测方法

技术方法

基于流粒度特征的检测方法

为每一条流抽取一个特征向量，作为机器学习的输入，分类每一条流是否为正常的数据包

流一般被定义为具备相同五元组的数据包序列，即具备相同源目的地址+源目的端口+传输层协议类型



技术评价

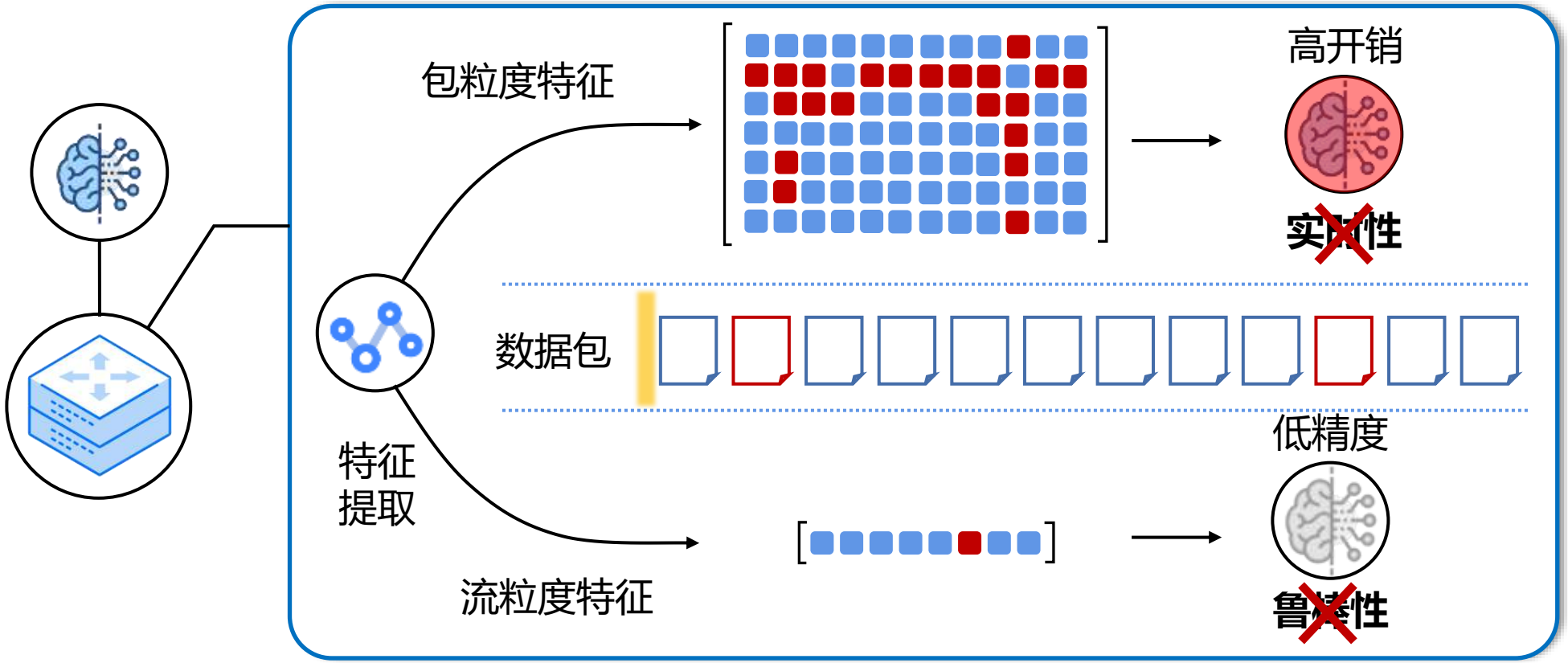
优势：1. 可以在流的基础上构造复杂特征，检测复杂攻击；2. 在没有专用硬件的条件下达到良好的效率

劣势：1. 分析粒度比较粗糙；2. 将数据包重组为流，引入额外的开销

比较：基于 流粒度 / 包粒度 特征的检测方法

包粒度特征和流粒度特征的比较

1. **包级粒度征的方案** 为每个数据包抽取特征：特征规模过大造成检测开销上升
2. **流级粒度征的方案** 为每条流抽取特征：粒度过于粗糙，攻击者很容易逃逸检测



两大类流量特征比较

基于流粒度特征的检测方法

代表方案: Whisper

分析目标：传统方法的流量特征，通常采用统计特征，例如单位时间内数据包的数量

虽然能保证实时检测，但是这些特征过于粗糙，容易被攻击者逃逸

而包粒度的检测无法保证实时性，是否存在流特征包特征之间的折衷方案

Table 1: Comparing the Existing Malicious Traffic Detection Methods

Category of Detection Systems		Feature Extraction Methods	Zero-Day Detection	High Accuracy	Robust Detection	Realtime Detection	High Throughput	Task Agnostic
Rule based		Preconfigured fix rules [6, 29, 35]	×	✓	×	✓	✓	×
ML based	Packet-level	Packet header fields [53]	✓	✓	×	✓	×	✓
		Context statistics [42]	✓	✓	×	✓	×	✓
		Payload statistics [68]	✓	✓	×	×	×	✓
	Flow-level	Flow-level statistics [5, 37, 77]	✓	×	×	×	✓	×
		Application usage statistics [4, 28, 49]	✓	✓	×[1]	×	×	×
		Frequency domain features, Whisper	✓	✓	✓	✓	✓	✓

[1] Bartos *et al.* [4] only considered evasion strategies for malicious Web traffic.

现有方案设计目标比较

Realtime Robust Malicious Traffic Detection via Frequency Domain Analysis, ACM CCS, 2021.

第2节 统计特征驱动的流量识别方案

- 基于数据包粒度特征的检测方法
- 基于流粒度特征的检测方法
- **可编程网络设备上的攻击流量检测**
- 加密攻击流量检测

可编程网络设备上的流量检测：可编程交换机

技术目标

在线速TB/s的高带宽网络环境下，实现对每一个数据包的可编程处理逻辑，相比于固定功能的传统转发芯片，能否实现灵活的交换功能

技术思路

PISA 架构：将芯片电路分为如下几个部分

Parser单元：实现自定义的数据包解析逻辑

Match单元：实现对包内特定信息进行内存查表

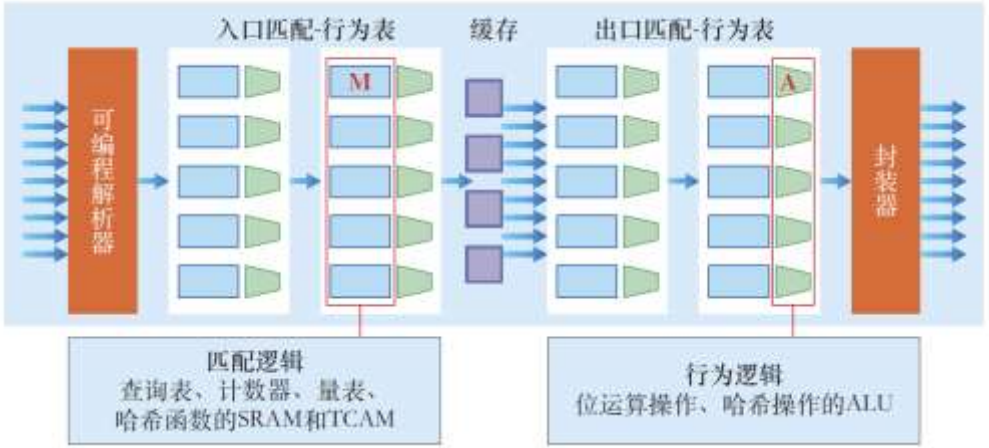
Action单元：实现对包头的操作或简单计算

Deparser单元：实现对数据包的重组

注意其中每一个Match-Action的组合被称为一个Stage

其中的Match模块和Action模块是高度可重定义的，辅助实现动态网络任务，例如负载均衡等

可编程交换机的编程逻辑模型PISA



APS Network 的可编程交换机产品
支持6.4Tb/s的交换容量

可编程网络设备上的流量检测：可编程交换机

技术应用于流量分析

将流级别特征抽取或者机器学习算法全部或者部分的实现在可编程交换机上，例如Intel Tofino芯片，他们采用P4语言进行编程，可以达到Tb级别的线速检测

这些方法分为纯粹数据面方法，和数据面控制面混合方法

可编程网络设备资源有限

Parameter	Intel Tofino 1 (up to 6.4 Tb)	Intel Tofino 2 (up to 12.8 Tb)
Process	16 nm	7 nm
Num of MAU stages/pipe	12	20
Total SRAM/pipe	120 Mb	200 Mb
Total TCAM/pipe	6.2 Mb	10.3 Mb
Scheduler	1-level	2-level
Number of queues/port	32/100 Gb port	Up to 128/400 Gb port
CPU port queues	32	Up to 128
Maximum SerDes speed	25 Gbps	56 Gbps
Port speeds supported	100 Gb/50 Gb/40 Gb/ 25 Gb/10 Gb	400 Gb/200 Gb/100 Gb/ 50 Gb/40 Gb/25 Gb/10 Gb
Maximum port contexts	256	256

技术评价

优势：1. 高吞吐低延迟，向着真实世界可用迈了一大步；2. 在高吞吐网络，例如广域网上完成检测，因此可以保护大量的合法用户

劣势：1. 硬件资源限制导致可以实现的模型和特征十分有限，检测精度也存在折损

Intel® Tofino™ 2

The next generation of programmable Ethernet switch, Intel® Tofino™ 2 is the best choice for meeting the needs of hyperscale data centers.

Built with the same architecture as Intel® Tofino™, it's capable of delivering twice the bandwidth of its predecessor—up to 12.8 Tb/s.

可编程网络设备上的流量检测：可编程交换机

代表方案: NetBeacon

分析目标：在高带宽网络环境下，完成通用的网络的攻击流量识别任务，同时保证极低的检测延迟

流量特征：借助可编程交换机，提取每一条流的粗粒度统计信息

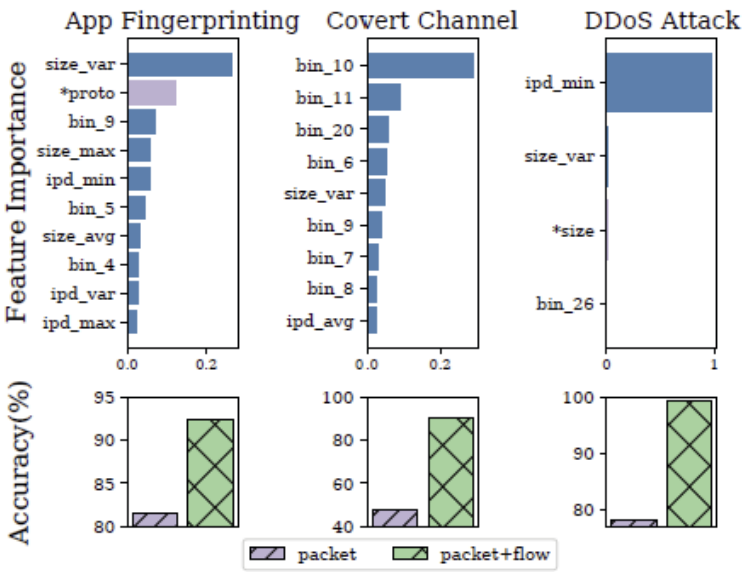
分析算法：同时在数据平面上，完成对每一个特征向量的决策树推理

技术优势：最早的可以部署到可编程交换机的智能流量分析系统，同时支持在可编程交换机上实现复杂的特征抽取操作，例如，方差信息等等

An Efficient Design of Intelligent Network Data Plane, USENIX Security, 2023.

方案细节：如何采用Match-Action结构操作高效计算均值方差等**复杂特征**：

将除法运算、平方运算转化成为查表问题，即在Match阶段对计算结果进行直接查表，而无需在Action阶段进行复杂操作和不支持的操作



复杂特征对提升检测性能效果显著

第2节 统计特征驱动的流量识别方案

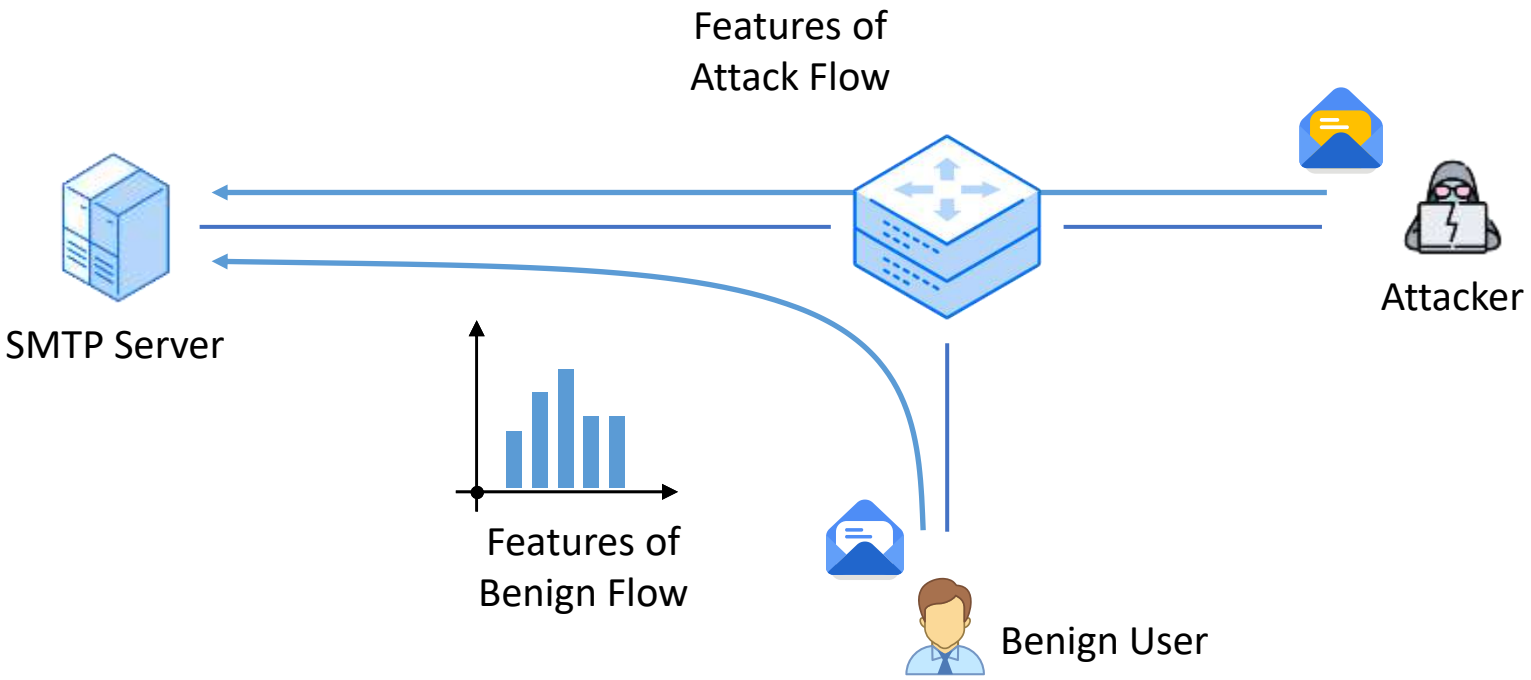
- 基于数据包粒度特征的检测方法
- 基于流粒度特征的检测方法
- 可编程网络设备上的攻击流量检测
- **加密攻击流量检测**

应对加密攻击流量

现有的方法无法实现通用加密流量检测

1. **通用检测方法**：无法检测加密攻击，因为**从传统流量特征角度无法区分出攻击**
2. 流量加密技术显著混淆流量特征，包力度和流粒度特征都不很明显

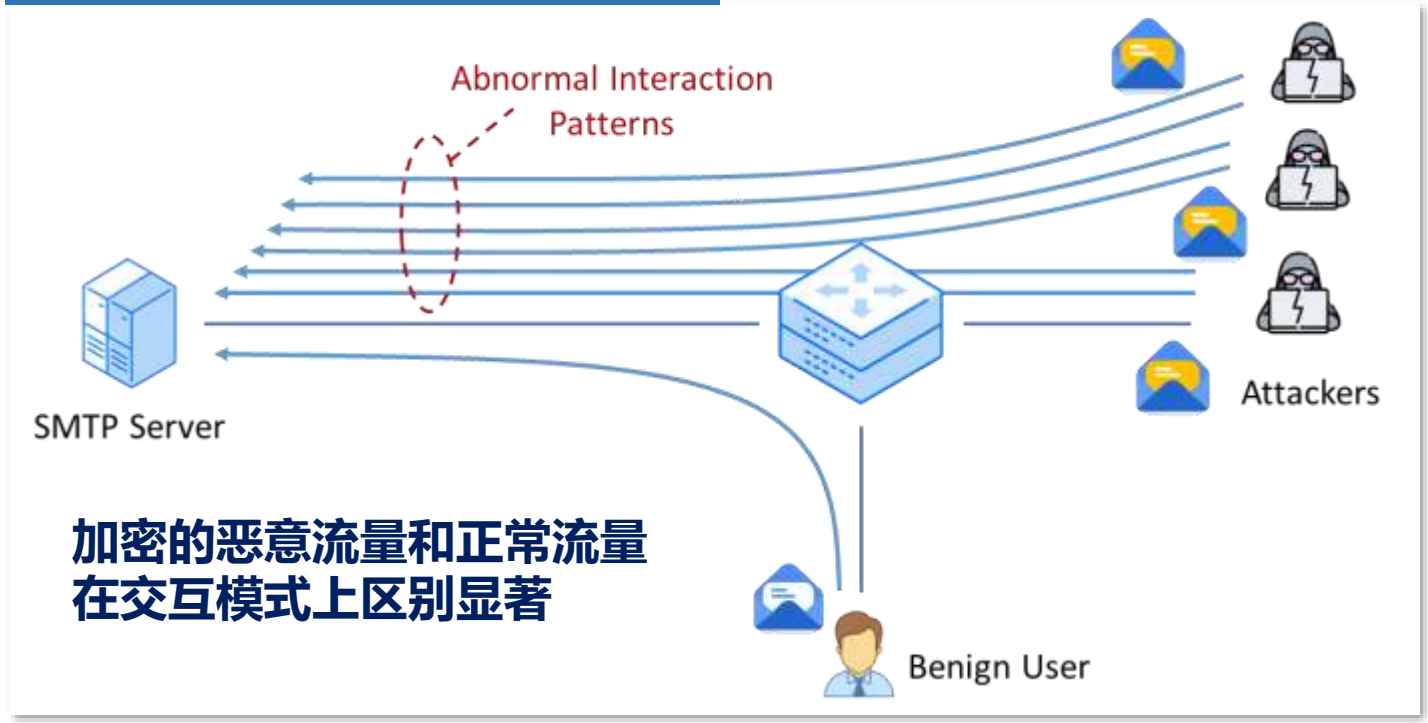
例如，加密的正常邮件流量 和 加密的垃圾邮件流量，流长度等传统特征无法区分他们



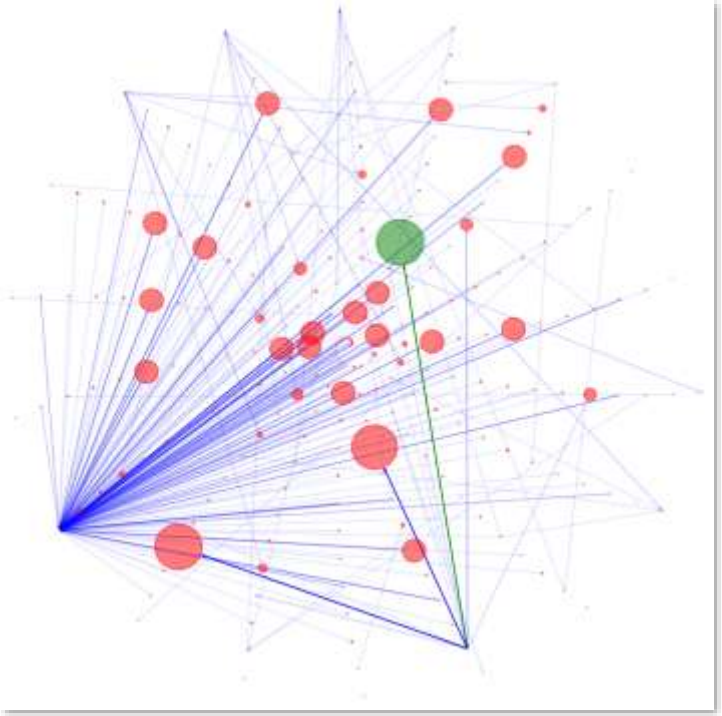
基于流量交互图的加密恶意流量识别

- **分析目标：识别未知地加密恶意流量**
  - 虽无法从单个流的特征检测，但可从流量**长期交互模式**中区分加密的攻击流量
  - **构建图结构表示节点之间的流量交互关系**

Motivation



Flow Interaction Graph



基于流量交互图的加密恶意流量识别

■ **流量特征：构建流量交互图**

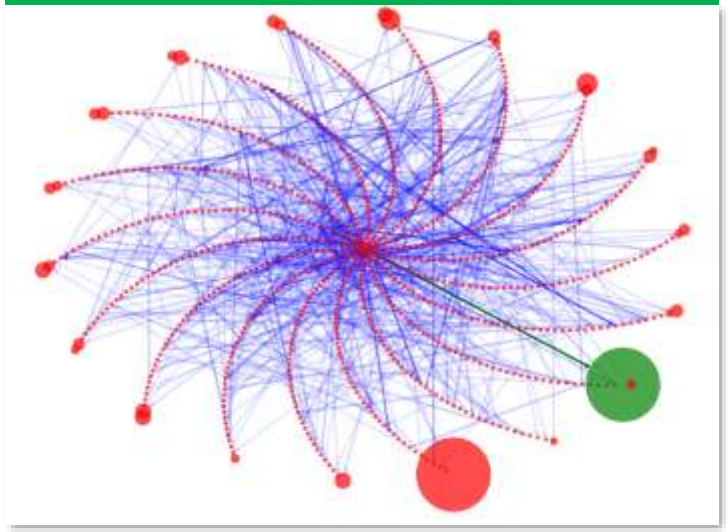
图上的节点表示用户，图上的边表示流，带来的问题是图的稠密宽度过高：

在AS2500当中存在超过**五万个活跃用户**和每小时**三百万条流**

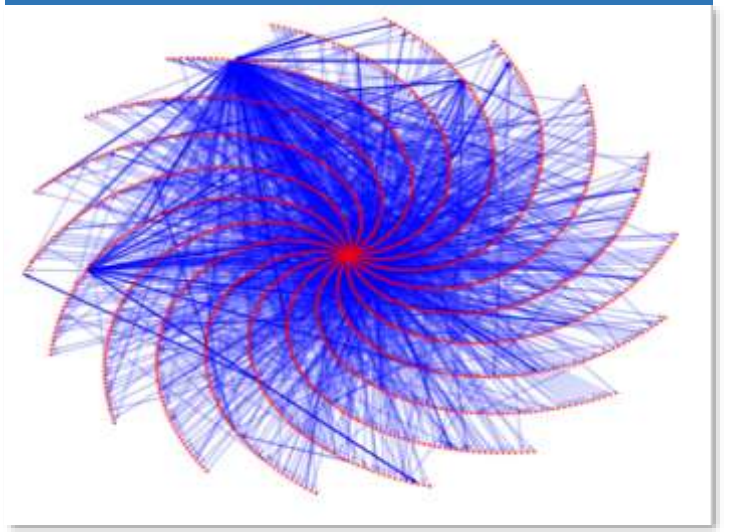
因此我们不可用一条边表示一条流，一个节点表示一个地址

研究方案中提出一种图压缩算法，将相似的流合并为一个边，相似的节点合并为一个地址

密度压缩后



密度压缩前



第3节 检测后的防御方案

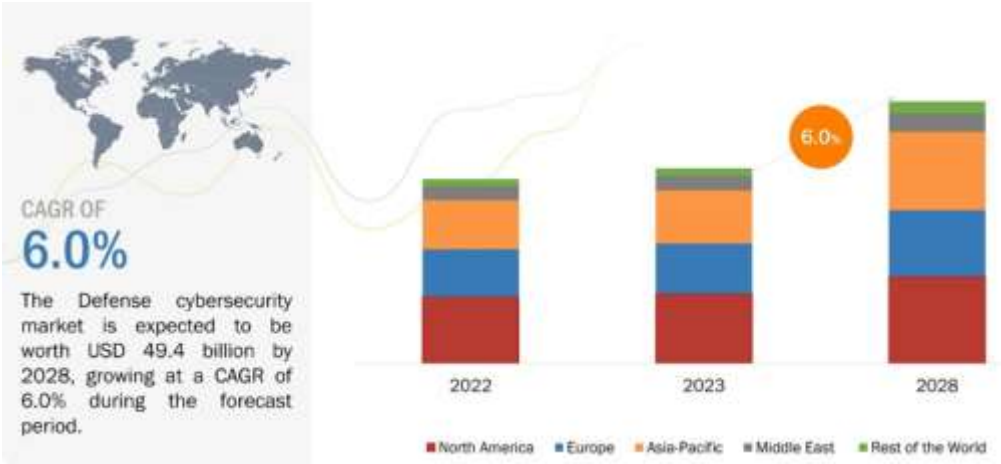
- 网络攻击防御的目标
- 传统防御方案及其局限性
- 可编程交换机上的防御方案

检测后的防御手段

技术目的

在识别出恶意流量后，得知恶意流量相关的信息，例如，攻击者的地址和一系列流量统计信息之后，如何有效拦截攻击者的攻击流量

网络防御产品市场规模逐步上升



技术难点

1. **防御泛化性**：攻击者的IP地址和恶意流量特征变化，无法简单通过固定规则实现防御
2. **防御流量规模**：对海量因特网流量进行匹配会存在较高的性能开销，尤其是对统计特征进行匹配，需要维护大量状态
3. **防御的时间效力**：部署规则快，并且在攻击消失后尽快回复原状
4. **防御的空间效力**：尽可能在接近攻击者的位置拦截攻击流量
5. **防御的误伤问题**：防御策略不应该损害正常流量

检测后的防御手段：可编程交换机方案

技术方法

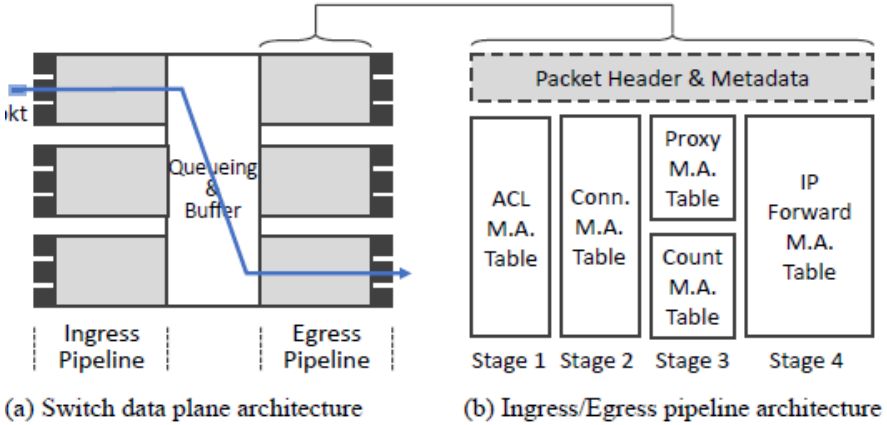
在交换机上高效部署过滤规则

在检测到攻击流量后，需要对恶意流量加以限制

为了提升传统固定规则防火墙性能，现有方法大量采用可编程数据面来实现更灵活的防御

在应对大规模恶意流量例如DDoS有良好效果

可编程交换机高层次电路结构



(a) Switch data plane architecture (b) Ingress/Egress pipeline architecture

技术评价

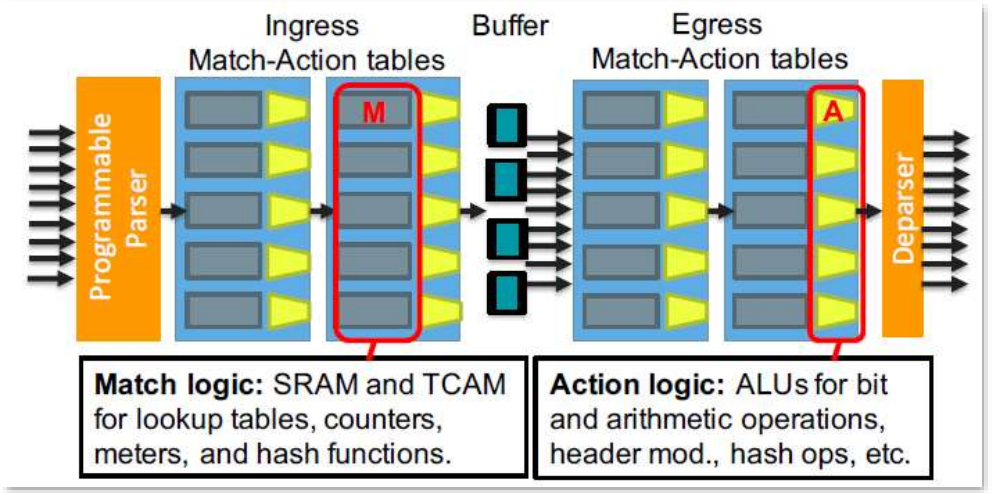
优势:

1. 防御方法灵活;
2. 可以应对大规模恶意流量

劣势:

1. 可编程数据面的资源限制
2. 细粒度防御一直难以实现

PISA的高层硬件架构



第4节 流量分析攻击

- **网站指纹攻击**
- 其他流量分析攻击
- 流量分析攻击的防御手段

网站指纹攻击：威胁模型

网站指纹攻击

网站指纹攻击通过分析用户生成的Tor流量推测被访问的网站

从而完全破坏Tor网络隐藏被访问的网站的功能

注意：网站指纹攻击采集用户加密流量的位置是用户到Tor第一跳的位置，即距离用户最近的节点

网站指纹攻击的两大类

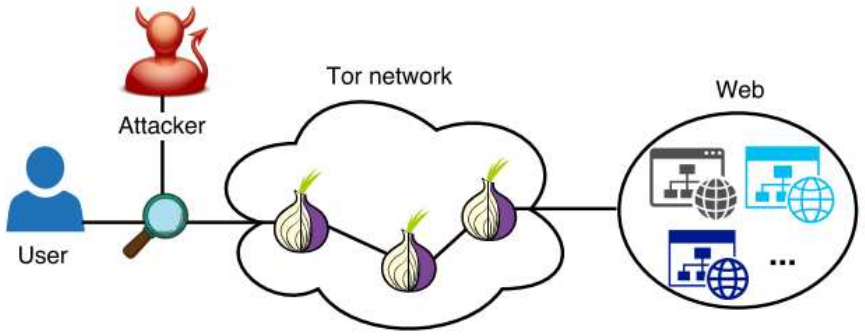
闭合世界假设：仅考虑用户是否访问一个集合中的站点

开放世界假设：考虑用户可能通过Tor访问任意站

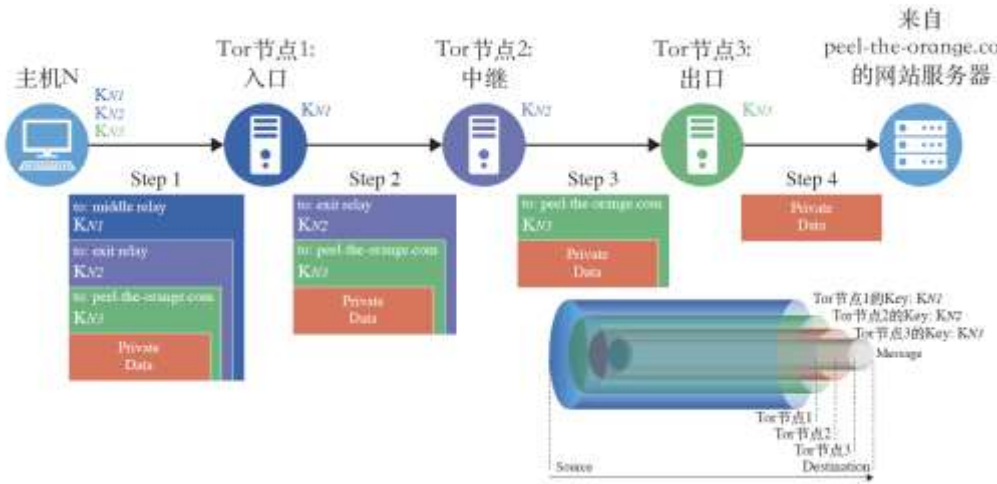
在网站指纹攻击中通常采用有监督学习算法分析流级别的特征

网站指纹攻击 威胁模型

注意攻击者的位置：在第一跳之前窃听流量



洋葱网络转发



第4节 流量分析攻击

- 网站指纹攻击
- **其他流量分析攻击**
- 流量分析攻击的防御手段

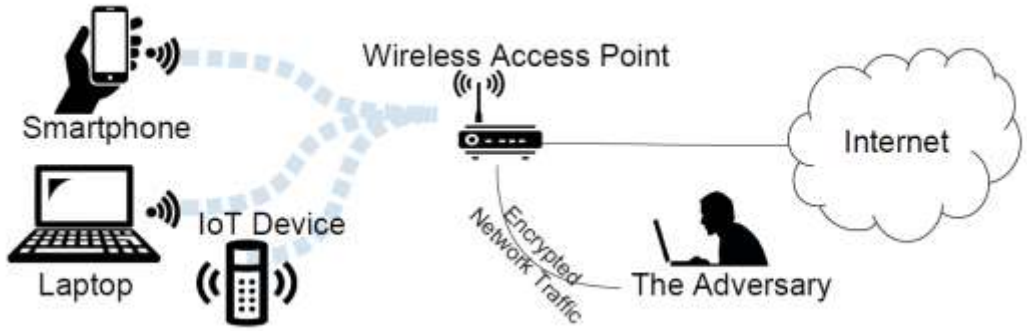
其他流量分析攻击

针对APP应用的方案：

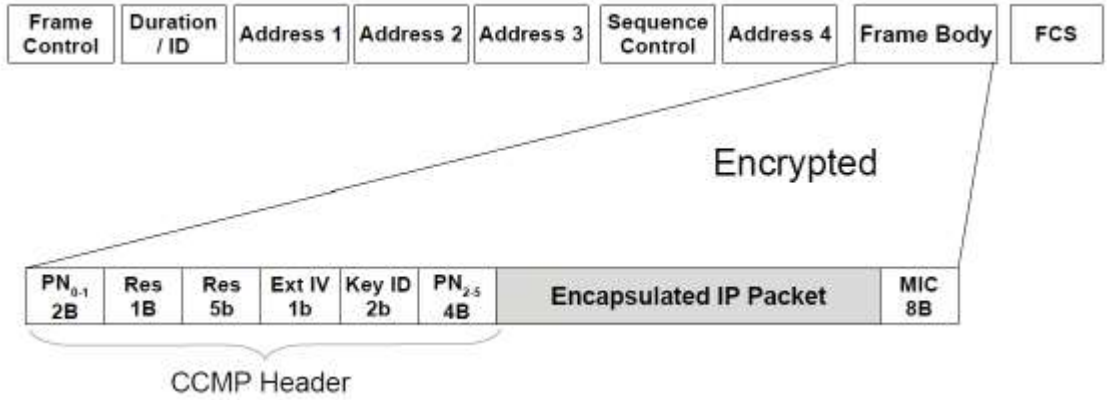
分析目标：根据智能手机生成的流量，判定用户的手机有什么APP在运行，从而破坏WiFi对通讯内容的保密性

流量特征：为包粒度的特征，因为在WiFi场景下数据包被封装在加密的帧当中，无法重组为流

分析算法：通常为有监督学习算法，同时将一组数据包的特征作为输入



攻击者通过WiFi帧嗅探发起流量分析攻击



WiFi帧当中的数据被完全加密

Packet-Level Open-World App Fingerprinting on Wireless Traffic, ISOC NDSS, 2022.
FOAP: Fine-Grained Open-World Android App Fingerprinting USENIX Security 2022.

第4节 流量分析攻击

- 网站指纹攻击
- 其他流量分析攻击
- **流量分析攻击的防御手段**

流量分析攻击的防御手段

基于特征扰动的方案：

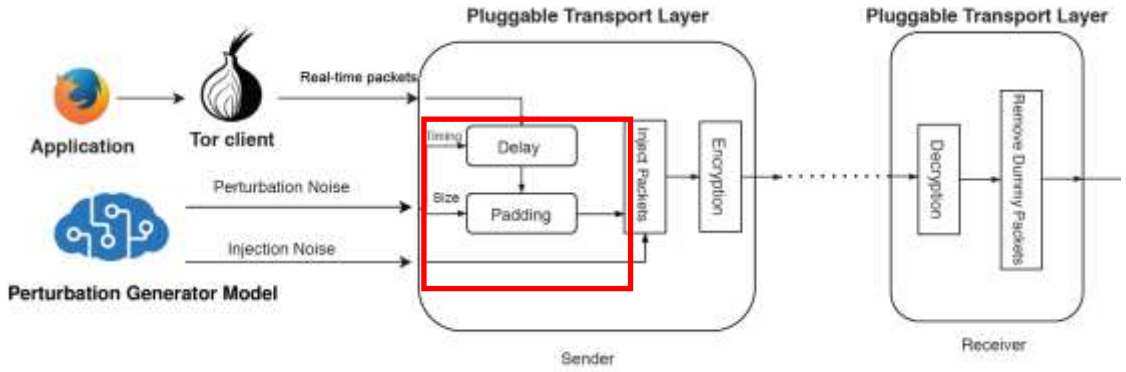
在发送端：传输的过程当中随机加入扰动，例如：

插入无意义的包、延迟发送数据包，为数据包添加无意义的负载等等

在接收端：消除扰动产生的影响，例如丢弃无意义的数据包和无效负载等等

本质上，这类方法为流量分析策略构建了**对抗样本**

即引起流量分析系统错误的样本



三种关键的扰动机制：
插入无意义的包、延迟发送数据包，为数据包添加无意义的负载

Defeating DNN-Based Traffic Analysis Systems in Real-Time With Blind Adversarial Perturbations, USENIX Security, 2022.

第5节 总结与展望

- **对于流量分析系统的批判**
- 流量识别的假阳性问题
- 流量分析的可解释性问题
- 流量分析技术知识体系
- 流量分析发展的总结与展望

对于防御方法：恶意流量识别研究的不足

广域网场景下的恶意流量识别是长期以来难以解决问题

- 入侵检测系统之父Vern Paxon教授曾批判智能检测方案：
- *It is crucial to realize that activity found in a small laboratory network differs fundamentally from the aggregate traffic seen upstream where NIDSs are commonly deployed.*
- 呼吁在实验室外的大型网络评估方法


Outside the Closed World:
On Using Machine Learning For Network Intrusion Detection

IEEE S&P'10 & Test-of-Time Award'20

- 十余年后，智能检测方案依旧只应用在小规模实验网络
- *It is essential to move away from a laboratory setting and approximate a real-world setting as accurately as possible.*
- 并认为这是全部机器学习安全应用的**共性问题**


Dos and Don'ts of Machine Learning in Computer Security

USENIX Security'22 & Best Paper Award

该问题
十年来未得到良好解决

逃逸、加密、隐私 三大问题

对于攻击方法：真实世界可行性被质疑

目标网站数的规模远大于现有方案可以实现的数量

- 近期研究表明在真实世界环境下，分析Tor路由节点的流量在5分类时可以达到90%以上的准确度，但是当目标网站类别超过20的时候准确度快速下降到80%以下
- 呼吁在数据集当中加入真实Tor节点采集的流量

Evaluating Website Fingerprinting Attacks on Tor in the RealWorld, USENIX Security 2022.

网络环境多样性
攻击目标多样性

真实世界WF难以实现

- 发现了在不同网络环境下需要不同的WF模型，这代表着WF模型针对Tor节点的迁移性差
- 呼吁采用多样化数据集或者生成的人工流量解决这一问题

Realistic Website Fingerprinting By Augmenting Network Traces, ACM CCS, 2023.

第5节 总结与展望

- 对于流量分析系统的批判
- **流量识别的假阳性问题**
- 流量分析的可解释性问题
- 流量分析技术知识体系
- 流量分析发展的总结与展望

流量识别系统的假阳性问题

在真实世界部署当中的人工处理压力：对假阳性问题的批判从来没有停止过

- 近期研究发现在Security Operating Centers (SOCs) 中存在大量的假阳性警报，即正常行为触发检测系统的警报
- 这种假阳性警报至多占据99%的比率


Outside the Closed World:
On Using Machine Learning For Network Intrusion Detection

99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms USENIX Security, 2022

- 近期研究进一步发现假阳性对真实世界部署识别系统至关重要，直接影响流量识别系统在真实世界的评价、测试、和运行


Dos and Don'ts of Machine Learning in Computer Security

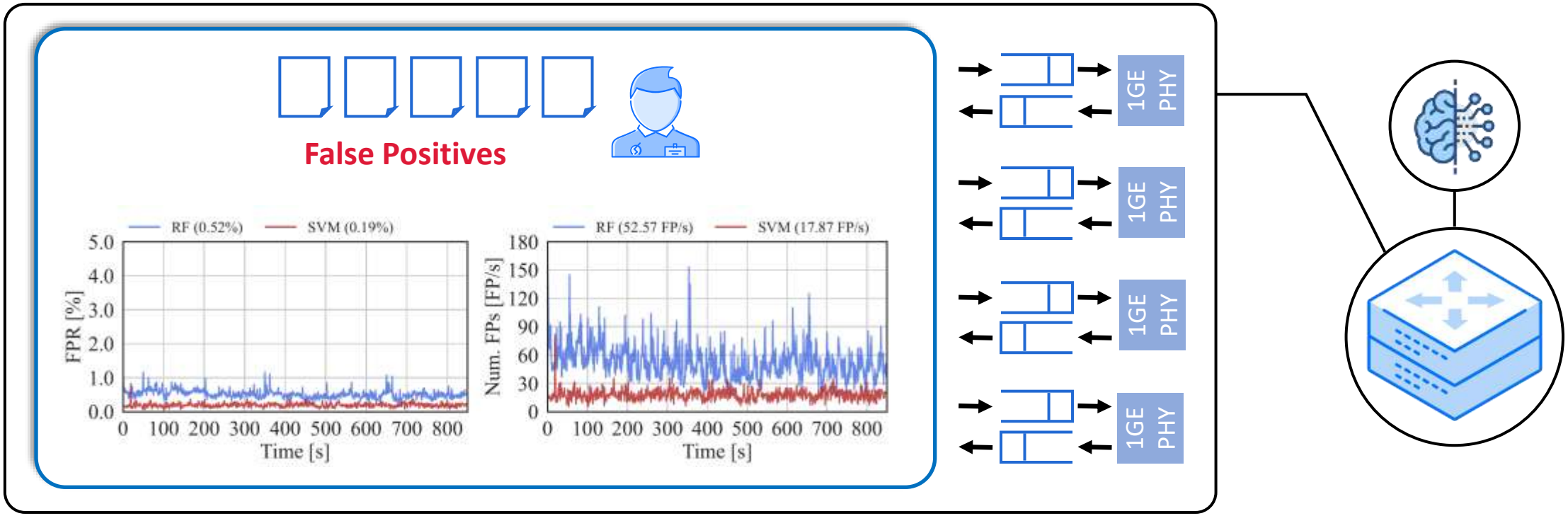
Alert Alchemy: SOC Workflows and Decisions in the Management of NIDS Rules, ACM CCS, 2023.

流量识别系统的假阳性问题

- **假阳性问题十分严重：**

大量产生假阳性的根本原因是流量规模本身就会特别大

导致人类专家必须手动从全体警报当中分理出假阳性警报



1M Flows * 0.01% FPR = 10K False Alarms

流量识别系统的假阳性问题

假阳性问题十分严重：

传统的降低假阳性率的两种方法：均是利用已知假阳性警报和其他任务相关信息

1. 重训练方法：

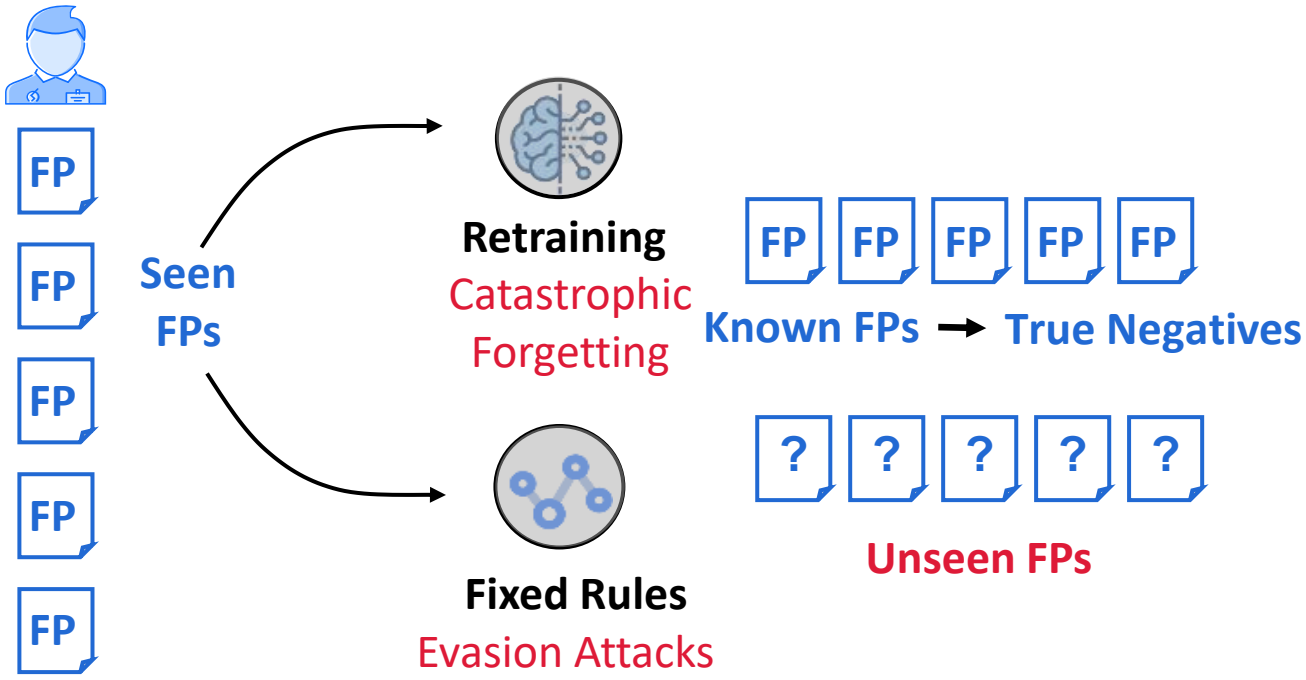
将被人工识别出来的假阳性警报加入数据集，对模型进行重训练，即可降低相似的假阳性警报

方法的不足：需要额外的计算开销，并且对黑盒模型不可用，例如成熟的商业流量识别软件

2. 固定规则方法：

对全体阳性样本进行规则匹配，过滤出部分的假阳性警报，例如采用IP地址白名单和AS信用度量

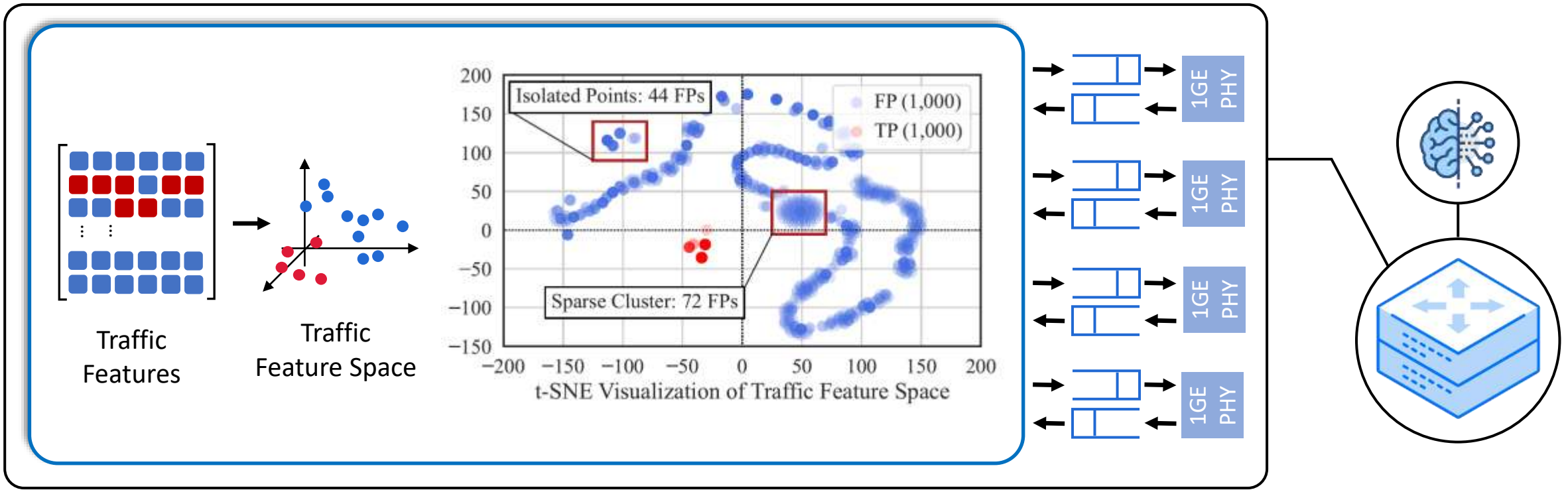
方法的不足：需要额外的领域知识，并且泛化性差，只能适用于已知的网络环境



流量识别系统的假阳性问题

- **解决流量分析的假阳性问题：pVoxel方案 (ACM CCS 2023)**

利用流量特征在特征空间当中的分布特性

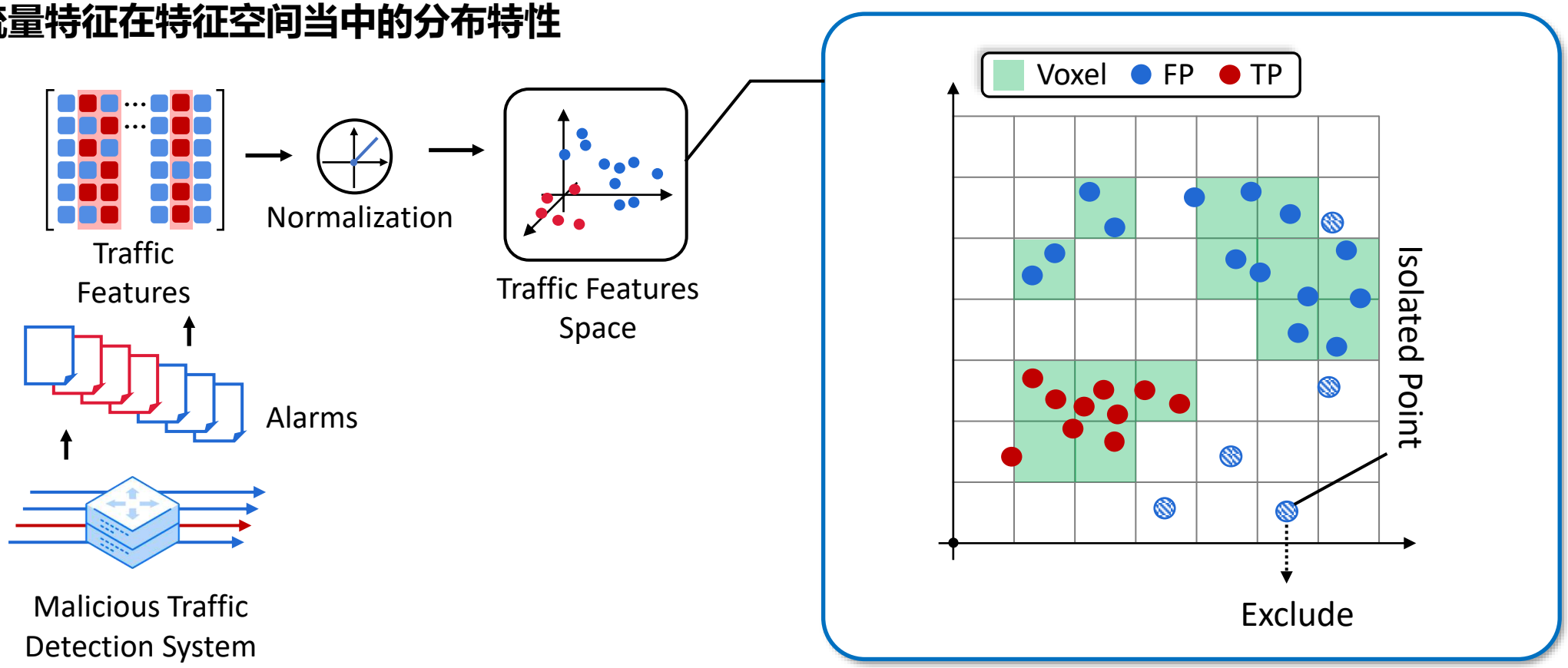


正常流量的特征在特征空间稀疏分布，异常流量的特征在特征空间当中稠密分布

流量识别系统的假阳性问题

- **解决流量分析的假阳性问题： pVoxel方案 (ACM CCS 2023)**

利用流量特征在特征空间当中的分布特性

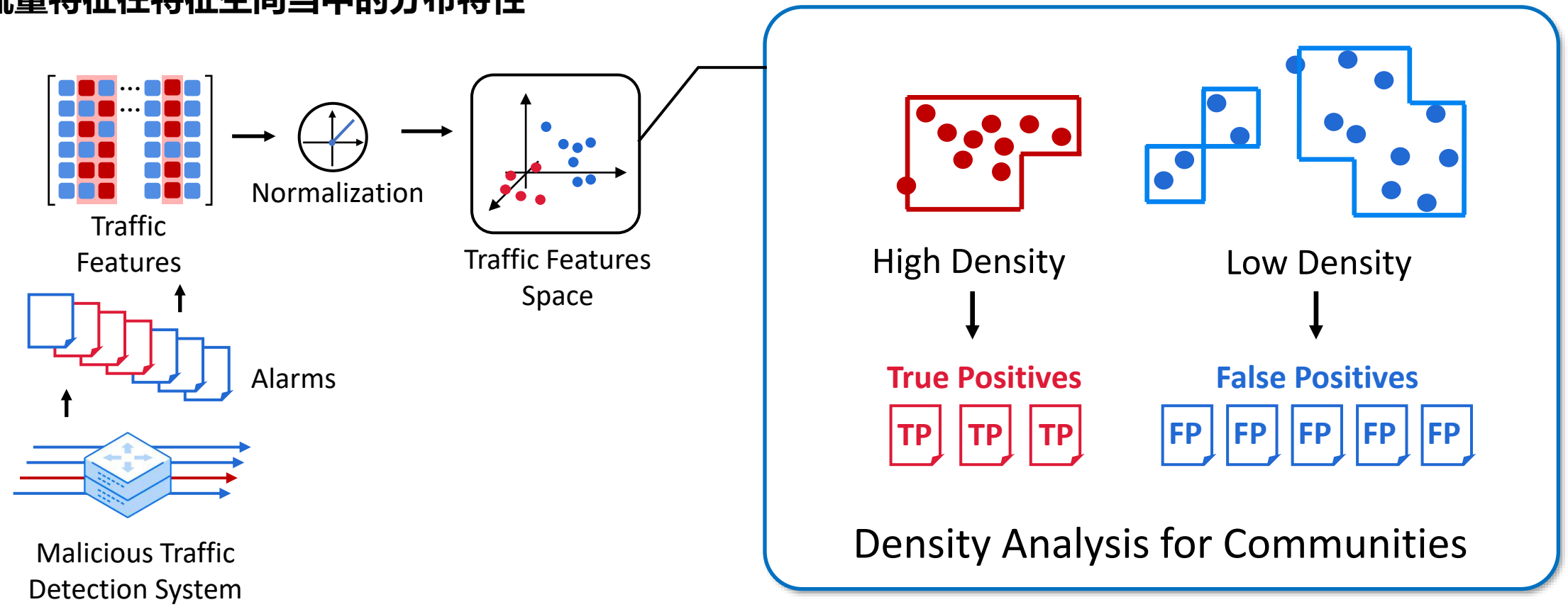


收集全体阳性样本，分析在特征空间的分布特性：
将稠密分布的分类为真阳性，将稀疏分布的分类为假阳性

流量识别系统的假阳性问题

- **解决流量分析的假阳性问题： pVoxel方案 (ACM CCS 2023)**

利用流量特征在特征空间当中的分布特性



该方案可以为目前最先进的11种流量识别系统降低超过95%的假阳性警报

第5节 总结与展望

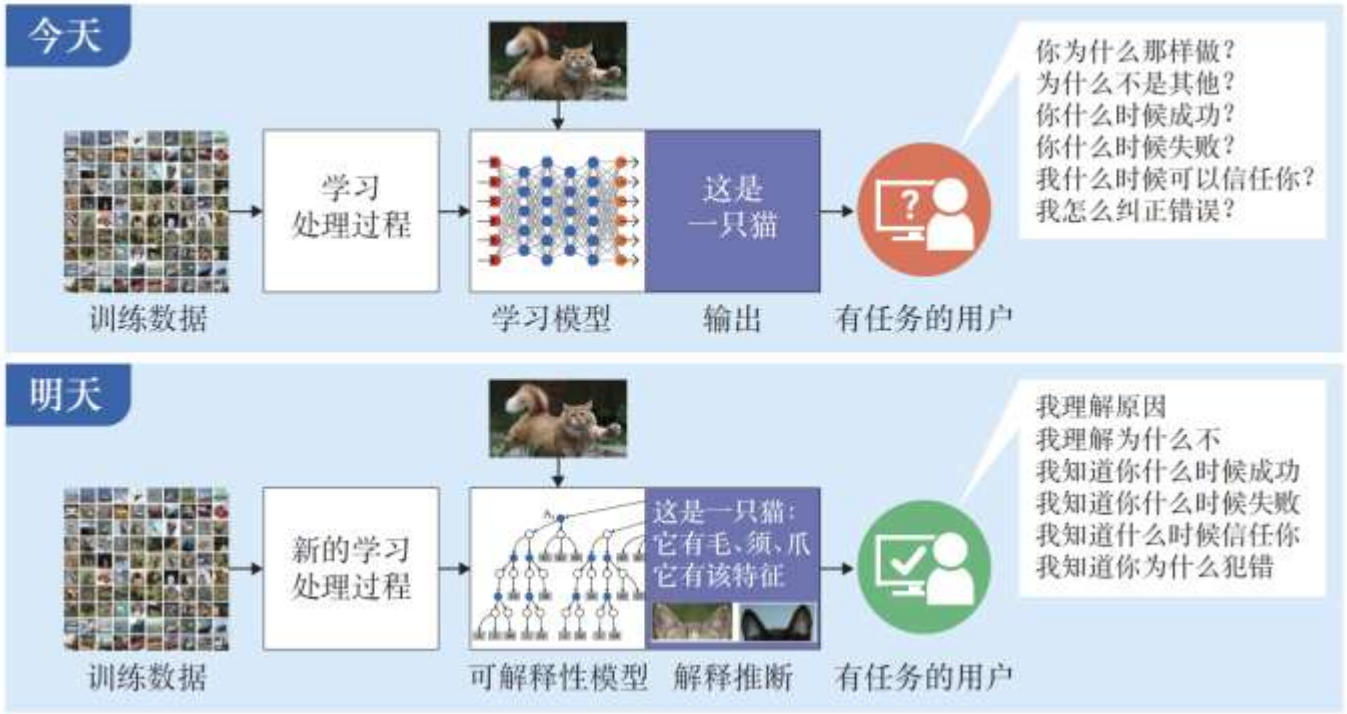
- 对于流量分析系统的批判
- 流量识别的假阳性问题
- **流量分析的可解释性问题**
- 流量分析技术知识体系
- 流量分析发展的总结与展望

机器学习的可解释性问题

可解释机器学习 Explainable AI

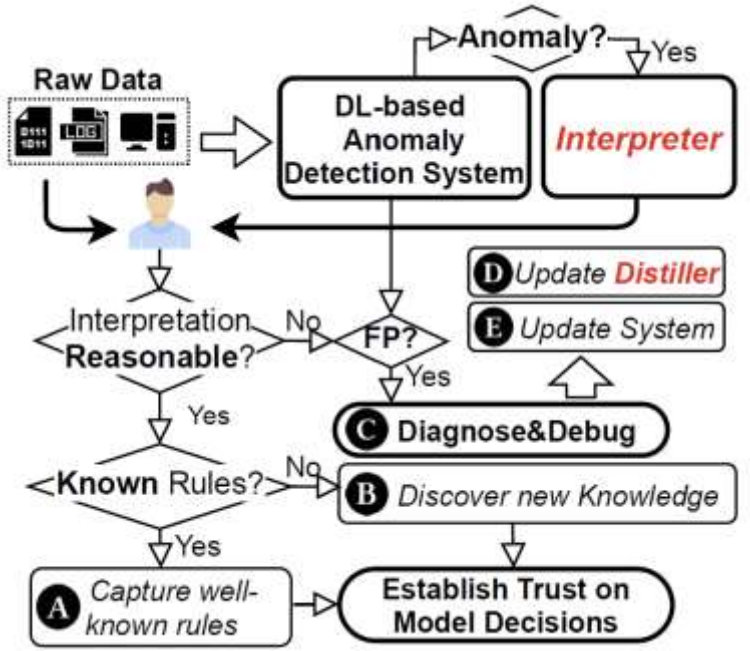
目的在于解释为什么机器学习做出这类决策

现有方法普遍标记重要特征，解释基于哪一特征做出了决策



可解释性对安全任务极为重要

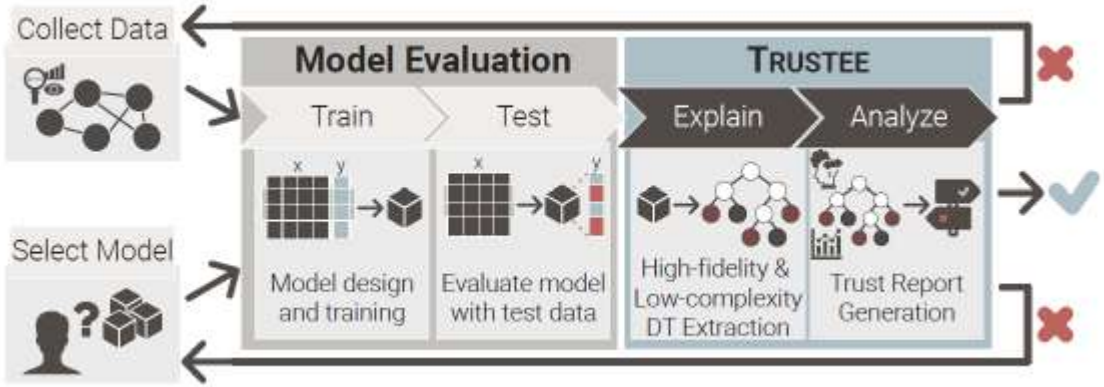
必须知道模型为何做出决策，才能在真实世界部署，及解决异常为何是异常的难题



机器学习的可解释性问题

代表性模型解释方法：Truesteeml

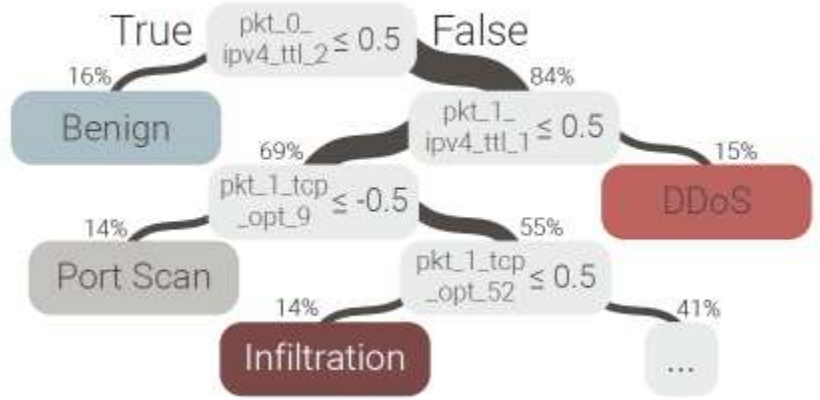
将DL/ML模型转换成为决策树，通过树状结构发现用于模型决策的关键特征和决策路径



通过解释模型，可以得到生成运维人员对模型的信任，进一步地约束假阳性和进行模型调优

AI/ML for Network Security: The Emperor has no Clothes, ACM CCS, 2022.

识别出数据包中对分类贡献度大的字段



将决策过程转换成为一个决策树

第5节 总结与展望

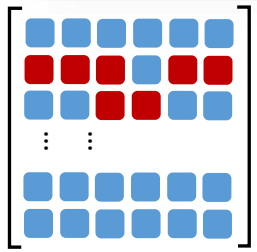
- 对于流量分析系统的批判
- 流量识别的假阳性问题
- 流量分析的可解释性问题
- **流量分析技术知识体系**
- 流量分析发展的总结与展望

流量分析的应用盘点

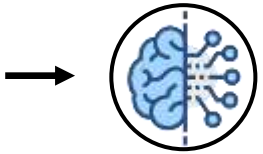
总结流量分析系统的目的

流量分析系统根据目推测的目标被分成两大类，但其实质性技术均为流量分析：

根据流量特征推测流量信息



流量特征



分析算法

目标信息

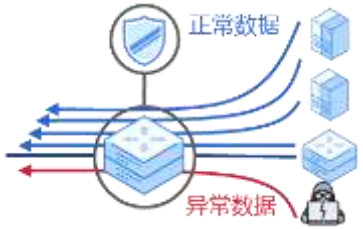
好的目的

入侵检测和防御

推测流量异常与否

目标未知

攻击流量检测系统



目标已知

攻击流量防御系统

坏的目的

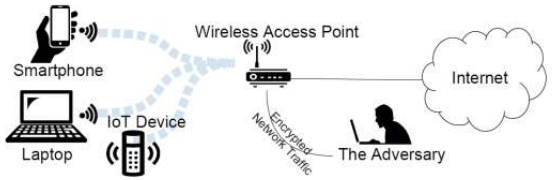
流量分析攻击

推测流量中的隐私

针对Tor

网站指纹生成

其他应用



APP指纹生成与其他加密流量指纹生成

如何系统化地分析流量分析系统：
抓住流量分析系统的三要素

流量分析的系统的架构的三要素和分类法

- **流量分析系统的三要素：流量特征、分析算法、分析的目的**

› **流量特征：**系统采用了何种方法表示流量

  › 负载特征：直接将数据包当中地内容作为特征

  › 统计特征：例如对于数据包头部内容设计统计量

其中，流量特征是影响系统性的关键

› **检测算法：**固定规则和机器学习算法

  › 固定规则：由人类专家设计

  › 机器学习：算法根据已知数据训练

› **分析的目的：**

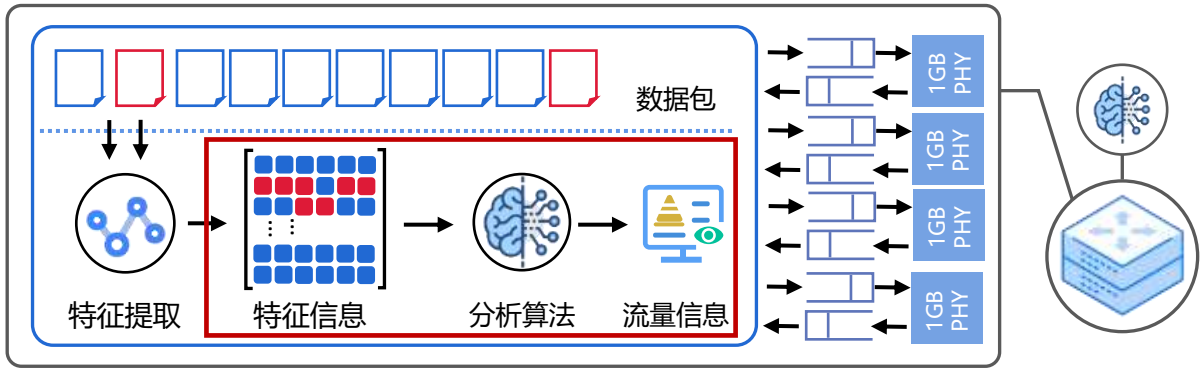
  › 窃取隐私：流量分析攻击

  › 保护用户：流量的识别和防御

流量分析系统的分类法

分析目的	检测算法	特征	实际应用
保护用户	固定规则	统计特征	攻击流量防御系统 (3.1-3.3)
		负载特征	传统固定规则入侵检测 (1.1)
	机器学习	包头特征	智能入侵检测系统 (2.1 - 2.5)
		负载特征	网站防火墙应用 (1.3)
隐私窃取	固定规则	包头特征	网络侧信道攻击 (第八章)
		负载特征	流量窃听 (第八章)
	机器学习	包头特征	流量分析攻击 (4.1 - 4.3)
		负载特征	------

本课程通过三要素模型：
将全部的流量分析系统放在同一框架下研究



本章内容关联性和教学目标

攻击流量的检测和防御

从海量的数据包当中抽取流量特征，推断攻击流量是否为正常或者异常，从而实现检测和拦截攻击流量，保护大量互联网用户的安全性

基于流量特征的隐私窃取

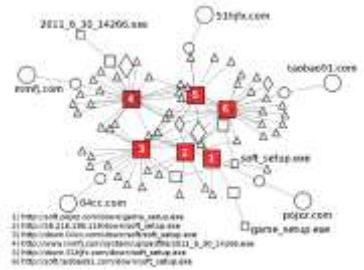
根据加密流量的特征，预测流量通讯内容相关的隐私信息

负载特征流量识别

研究内容:

传统NIDS

网站防火墙

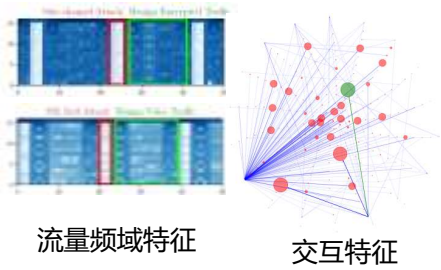


统计特征流量识别

研究内容:

流粒度、包粒度特征

加密流量识别



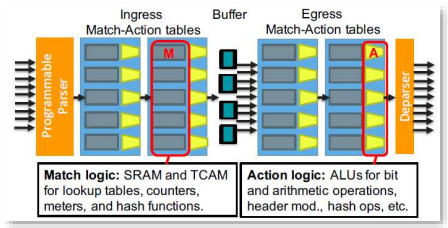
流量频域特征　　交互特征

攻击流量防御系统

研究内容:

传统匹配过滤机制

可编程数据面防御机制

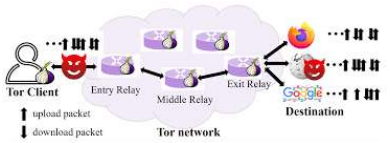


流量分析攻击

研究内容:

网站指纹生成

其他流量分析攻击



WF

未知攻击

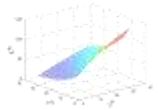
高效信息使用

未知攻击

未知流量模式

流量规模

海量流量匹配



窃取信息

高效信息使用

我们以流量分析系统三要素为切入点，系统化分析各种流量识别系统

第5节 总结与展望

- 对于流量分析系统的批判
- 流量识别的假阳性问题
- 流量分析的可解释性问题
- 流量分析技术知识体系
- **流量分析发展的总结与展望**

技术价值观：网络技术的双刃剑属性

- **流量分析技术可以用于攻击，也可用于防御**

› **流量检测系统：**根据流量特征推测流量是否为恶意

› **流量分析攻击：**根据流量特征推断用户隐私信息

› 二者在技术上的相似性，必然导致了流量分析在真实世界的困难性：

  › 例如，如何防止流量识别系统抓取流量进行流量分析攻击

  › 目前抓包分析被大量网络服务提供商禁止

要解决流量分析技术的隐私难题，利用流量分析技术保护网络安全需要行政法规保障，同时也需要研究人员面有正确的技术价值观

将流量分析技术应用于真实世界

- **真实世界应用流量分析技术，释放人工智能的强大能量保护网络安全，需要解决一系列难题：**

- **隐私保护难题：**
  - 流量数据携带用户隐私，部署技术受到严格显著
- **真实世界泛化性：**
  - 真实世界的网络环境复杂多变，为分析鲁棒性提出挑战
- **人类用户可用性：**
  - 解决假阳性人工分析开销问题，和可解释性问题
- **硬件设备部署难题：**
  - 真实流量规模大，如何部署到高性能网络设备上

目前，真实世界大规模部署流量分析，保护网络基础设施安全仍然面临巨大挑战