关于基于网络的入侵检测系统的调查

李锭轩　　林静雯　　胡冬寅

1　导言

入侵检测系统 (IDS) 是维护系统网络安全的重要设施之一，它可以检测系统中疑似攻击的流量，及时向管理员提出警报或联合防火墙动态调整保护策略。依据数据来源，IDS 可以分成两大类: 使用主机数据的**基于主机的 IDS**(HIDS)，以及使用网络数据进行分析的**基于网络的 IDS**(NIDS)。本文将对近些年来，在 NIDS 领域取得最新进展进行梳理、总结与分类，同时提出了未来研究的可能方向。

目前，许多关于 IDS 的论述文章都是基于系统使用的方法进行分类的。这样的分类方法固然可以给读者了解到最新发展的技术是怎样的，但却缺少层次和整体感，给人一种琐碎的感觉。以本文将采取全新的组织方式: 按照**协议栈层次**对 IDS 进行分类。这样分类的好处是:

1. **体现 IDS 设计思想**

   由于协议栈各层有其给定的特点和功能，部署在不同层次上的 IDS 也会相应地考虑这些因素，以与该层特点适合的方式实现入侵检测。故 IDS 的设计总会其所在的层次相关联; 而采用层次分类的方法可以帮助读者更好的思考一个 IDS 为什么要这样的设计。

2. **分类更加简洁**

   即便按照 ISO 开放系统互连的七层模型，IDS 也至多会分为七大类; 但按照采用的方法来分类，可谓不胜枚举。因此，采用层次分类的方法可以以一种更加简洁、统一的方式去了解 IDS 的研究现状。

3. **以全新的角度思考**

   按照方法分类对 IDS 进行了解时，可能会陷入如何寻找新方法的困境中。但是，按照层次来看，可以帮助人们思考以往成熟的方法在其他层次应用时是否会有更好的效果。

本文按照如下顺序展开: 第2章按层次介绍目前 NIDS 的最新研究成果; 第3介绍未来可能的研究方向，希望可以抛砖引玉; 第4章包含其他与本文相关的内容。

2 NIDS 技术

2.1 应用层

在计算机网络体系中，应用层的特点是种类繁多，功能多样。正是由于这种多样性，针对不同应用服务的攻击有着很大的差异，因此对它们的检测方法也是各不相同。

总的来说，入侵检测的策略可以分为**基于协议的检测**和**基于内容的检测**。对于基于协议的方法，Zhang 等人根据 HTTP 错误生成模式对 HTTPS 协议中的恶意流量进行了检测;Manos 等人则提出了根据 NXDomian 流量来检测 DNS 攻击的方法。对于基于内容那个的检测，Sajjad Arshad 通过对 Web 网页元素进行细粒度的来源标识，达到检测基于浏览器扩展的广告注入的检测;Hugo 等人则提出了一种基于发送者的邮件结构和设置信息来检测异常邮件的方案。

2.1.1 基于协议的检测

利用 NXDomian 流量检测 DGA

Manos(2017)[1] 针对基于域名生成算法 (DGA) 的 DNS 攻击，提出了根据 NXDomian 流量来检测恶意域名的方法。

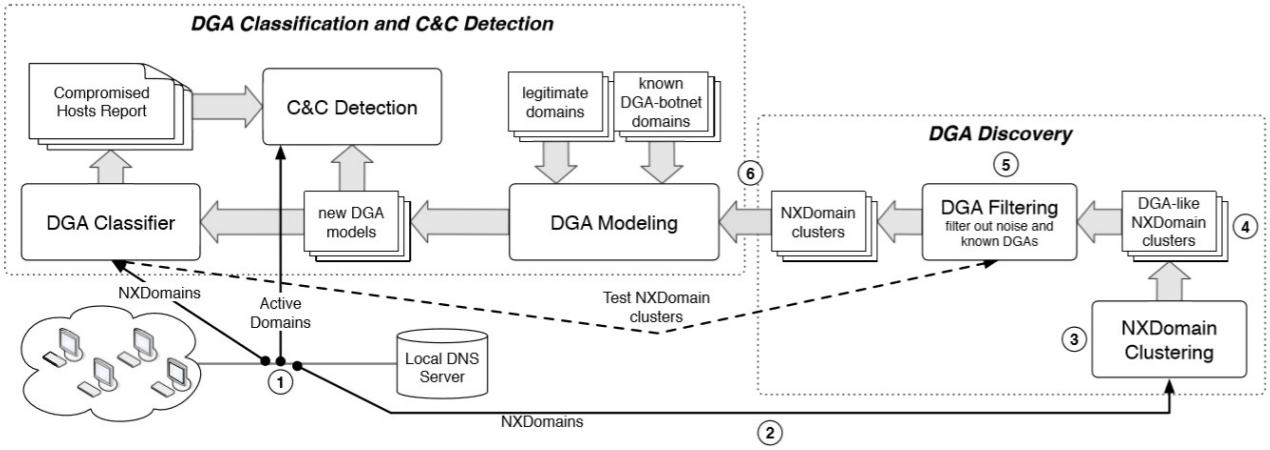
在传统 DNS 攻击中，攻击者会让被感染的僵尸主机解析一个只有恶意域名服务器才能解析的域名，使得二者之间建立 DNS 协议联系，为后续发送各种攻击代码或恶意命令提供了条件。但这种方法只使用一个 DNS 地址，因此很容易被加入到黑名单中过滤。而利用 DGA 可以生成大量域名，将其中部分域名注册到恶意 DNS 服务器上，然后僵尸主机使用同样的 DGA 和种子生成大量 DNS 域名，并依次进行解析。当僵尸主机成功获得了解析的域名，那么便与恶意域名服务器建立连接。

本文作者提出了根据 NXDomian 流量来检测恶意域名的方法。NXDomian 是在用户所查询的域名无法解析时服务器产生的响应。由于 DGA 产生的大量域名中只会有一小部分被注册在恶意服务器上，在对这些域名进行逐一查询时会产生大量 NXDomian 流量。故利用产生 NXDoamin 流量的域名特征就可以对 DGA 进行检测。利用这一点，作者提出了如图1所示的 Pleiades 系统。

此模型可以被部署在本地的 DNS 服务器。在这里，它首先会收集 NXDomian 流量，对这些域名送到模型中 NXDomain Clustering 模块进行聚类。聚类后在 DGA Modeling 模块中与已知 DGA 作比较，若吻合，则将它们归类到那一种 DGA 中; 若不吻合，则利用将这一类特征归为新的 DGA，并对其建模、存储在新 DGA 模型中，留待与之后比较。当完成这样的工作后，由某个 DGA 所产生的恶意域名会大概率被检测到与某一个 DGA 的特征相吻合，从而可以将它列入黑名单中。

值得一提的是，NXDomian 的产生并不一定是恶意的，它可能是由于用户 DNS 服务器配置错误，或者手误产生，如果检测提取了这种流量的特征显然是不合理的。为了避免这种情况，作者在模型中加入了一个过滤器，它可以检测出这类 NXDomain，并将其过滤，利用的方法

图 1: Pleiades 系统示意图



是检测域名与一些合法周知域名的相似度，若相似，则说明这类 NXDomain 大概率是人为失误造成，应该将其过滤。

基于 HTTP 的恶意软件检测

Zhang(2018)[13] 等观察到良性软件和恶意软件产生的 HTTP 错误以及所采取的恢复策略具有不同的模式，并以此作为依据设计了新的系统 *Error-Sensor*，该系统可以仅从 HTTP 错误以及其相邻时间的成功请求检测恶意软件流量。该文章主要关注的是 4xx 类状态码指示的 HTTP 错误，该类状态码用于指示由于客户端导致的错误。恶意软件的活动而生成的 HTTP 错误大多属于这些类别: 网络扫描攻击可能会因扫描不存在的目标产生 404 Not Found 错误，扫描受保护路径中的网页或策略违规的请求也可能导致 403 Forbidden 错误。

并非发生 HTTP 错误就意味着该 HTTP 发起者为恶意软件，但是基于大量的现实数据，Zhang 等观察到良性软件和恶意用户所生成的 HTTP 错误模式以及其恢复模式存在差异，例如：受恶意软件影响的客户端通常会产生比良性客户端更多的 HTTP 错误；而在产生 HTTP 错误后，恶意软件会可能会采取一系列恢复措施，比说连接到一个良性服务器来测试网络，而良好的客户端在 HTTP 错误面前经常缺乏恢复例程。

Error-Sensor 检测系统就是基于这些差异来进行检测恶意软件的。在进行流量过滤以及根据客户端，服务器，网页和错误代码四元组进行错误聚类后，将错误聚类的流量进行分析，生成了错误源模式、错误生成模式、错误恢复模式这三大类特征指标，作为随机森林分类器的输入。具体来说，错误源模式包括了客户端声誉、服务端声誉、意外错误率等特征，错误生成模式包括错误生成顺序、频率等特征，错误恢复模式包括时间相关性、URI 路径关联等。

结果表明，该系统在 0.005% 的 FPR 水平下达到了 99.79% 的检测率。要注意的是，该系统时 DNS 入侵系统的补充，因为 http 是在域名得到解析的前提下进行的，也就是该系统在 DNS 入侵系统检测失败后的进一步检测。

基于 HTTP 的负载聚合技术

Felix(2017)[8] 等提出了利用基于 HTTP 的载荷聚合技术 (HPA) 来识别完整的 HTTP 请求响应对，并据此设计了一个高性能，高效率的入侵检测系统。在评估中表明，该系统在仅分析 2.5%的有效负载时检测到 97.4%的事件。

HPA 载荷聚合技术沿用了前部载荷聚合 (FPA) 和基于对话的载荷聚合 (DPA) 的只转发每条 HTTP 消息的前 N 个字节的。经验证据表明，与入侵检测最相关的数据存储在应用层协议数据单元 (PDU) 的附近，而有效载荷的其余部分大多不相关，由此减少了转发到 IDS 中的数据量。

HPA 相比于 FPA 和 DPA 的优势在于其兼容了 HTTP2 和 HTTP1.1 中所包含的流水线操作，也就是一个 TCP 连接中会包含多个 HTTP 请求-响应对。HPA 的方法并不是单纯的保留每个 TCP 连接的每个方向的前 N 个字节，而是检测流水线的 GET 请求，以及其响应。并将该请求相应对作为单个消息导出，从而允许 IDS 分析所有的 HTTP 消息的前 N 个字节。

2.1.2 基于内容的检测

通过网页元素表示检测广告注入

Sajjad Arshad(2016)[2] 提出了一种通过细粒度的 Web 内容来源标识的方法来检测浏览器扩展程序插入的广告。

浏览器扩展程序很好地丰富了浏览器的功能，这为用户提供了很大的便利。但由于扩展程序一般拥有较高的权限，因此它也为攻击者提供了一个很好的平台，其中一种威胁就是基于扩展的广告注入。这种广告注入不仅可以通过播放广告来为注入者产生收益，它甚至还能使得网页发布者的合法收入转向第三方; 另一方面，它可能会让用户看到不想看的内容，而用户通常不能分清这是来自第三方注入还是来自网页发布者，这给网页发布者造成了口碑上的负面影响。

为此，作者提出了 *ORIGINTRACER*, 这是一种可以集成到浏览器中的插件，它可以在文档对象模型 (DOM) 的粒度上标识网页元素的来源。它实现的原理是利用**起源标签集**为网页中的 DOM 元素做注释，因为网页无论是在从发布者的服务器加载时，在按照脚本执行时，还是进行扩展时，浏览器总是可以知道添加或是修改 DOM 元素的主体是什么，这时候 *ORIGINTRACER* 就可以将这些主体添加到起源标签集中，以标识 DOM 元素的来源。

在分清 DOM 元素的来源后，*ORIGINTRACER* 的做法是将注入的 DOM 元素通过高亮或是其他一些易于发现的标识的形式展现给用户，而非直接将其删除。这样做的原因是，即便是广告注入，也很难判断用户是否需要它。因此过滤的权利被保留给用户，以帮助用户做出最有利于他们的选择。

基于非文本特征的鱼叉式钓鱼邮件检测

Gascon(2018)[11] 提出了一种基于发送者的邮件结构和设置信息，使用传统机器学习模型的鱼叉式钓鱼邮件检测方法。鱼叉式钓鱼邮件攻击是指针对特定目标构造的带有恶意链接或附件的邮件攻击，使被攻击目标无法判断该邮件是否真的来自其所声称的来源。由于不同的鱼叉式钓鱼邮件攻击相差甚远，故很难被统一抵御。在实际场景中，攻击者对其假冒的发送者的了解是有限的，而这种已知信息的有限性造成的模仿的差异性是该检测方法的关键。而 Gascon 提出的方法可以学习不同发送者与具体文本无关的特征，如果邮件的特征与其声称的发送者特征不相匹配，则被识别为欺骗性电子邮件。

该文章将发送者的信息提取为与具体文本无关的三类特征，分别是**行为特征**、**组成特征**以及

传输特征。行为特征是指发送者的习惯等特征，例如是否使用数字签名，附件类型和顺序等等。组成特征则是指客户端机器配置信息，例如公共头标的类型，顺序和语法。传输特征则与邮件传输路径有关，包括 Received 头标和顺序，以及 Received 头标中提供的传递协议和 TLS 功能。从结果可以看到，采用这些特征学习出来的不同发送者之间具有比较高的特异性 (根据曼哈顿距离)，说明了该特征提取是较为成功的。并且，传输特征在这三类特征中具有最大的辨别力，而且由于其涉及基础硬件设施，也最难被伪造，所以该特征在检测中是有效的。

该文章采取了 k 最近邻 (KNN) 和支持向量机 (SVM) 模型，所以其所需的样本量远远少于深度学习所需要的样本量。其中 KNN 最少只需要 2 个样本，SVM 为 5 个。这十分符合该论文所研究的特定攻击的场景——在这种特定攻击中，数据量是有限的。具体来说，来自于同一发送者的信件的量不大，故采取 KNN 和 SVM 模型适合的。特别的，SVM 分类器在少量可用的电子邮件中提供了更好的性能。

2.2 网络层及传输层

网络层与传输层信息是绝大多数 NIDS 会应用的信息。现行网络中最流行的网络层协议是互联网协议 (IP)，传输层协议是传输控制协议 (TCP) 和用户数据协议 (UDP)，绝大多数的工作在网络层与传输层的 NIDS 都依靠这几种协议的头标内容进行分析。

按照检测方法，网络层与传输层 NIDS 可以大致分为两类: **基于机器学习方法的**和**基于非机器学习方法的**。由于近些年来人工智能技术的急剧发展，越来越多的 NIDS 使用机器学习的方法，比如 Navaporn 等人利用循环神经网络 (RNN) 进行入侵检测。但非机器学习方法也有一定新的进展，比如 A. Gaurav 等人使用熵和评分手段进行检测，S. Jero 等人利用有限状态机发现新的攻击。本节按照 NIDS 采用的不同方法展开。

2.2.1 使用机器学习方法的

使用 TLS 证书进行检测

I. Torroledo 等人 (2018)[3] 首次实现了利用 TLS 协议的钓鱼攻击检测。由于 TLS 协议中数据载荷被加密，因此难以利用载荷进行检测。该方法利用 TLS 协议中的证书信息进行钓鱼攻击检测; 作者总结了 35 个与 TLS 证书相关的特征，使用单热 (One-hot) 编码方案对特征进行编码。编码后，将向量输入神经网络进行学习。

神经网络由三部分并行构成，两部分是长短时记忆模型 (LSTM)，用来处理证书中与被颁发主体和颁发机构相关的信息; 另一部分是全连接网络，用来处理从证书中提取出的 35 条特征。三部分的输出结果串接构成新的输入输进全连接网络，得到最终的评分。该评分标志着干网站属于恶意网站的概率。

该方法实现了 94.87% 的恶意软件检测率，88.64% 的钓鱼网站检测率。但该方法有一个不足之处——对于钓鱼网站的检测率较低，这是由于钓鱼网站的证书会伪装得与真实的证书十分相近。

使用数据报报头信息进行检测

Chockwanich(2019)[12] 实现了在不需要人工指定规则的前提下，使用深度学习算法检测网络攻击的方法。该方法只使用了报头的相关信息，并不涉及到载荷，以此保护数据隐私。在数据预处理中，先将通过 tcpdump 命令获得的 pcap 格式流量，并删除了数据包的有效负载。然后将包数据转换为流数据。流数据不同于包数据，它是其覆盖的某个通信下的数据包的统计量和基础信息的总结，包括 IP 地址，持续状态，历史状态以及包速度，字节速度供 16 个特征等等。

在该文章的方法中，使用网络流的 16 个元数据作为特征供神经网络训练。为了检测和对攻击方式进行分类，该文章使用了 RNN, Stacked RNN, CNN 进行训练，检测了 DoS 攻击，端口扫描，使用 UDP 的网络扫描，使用 TCP 的网络扫描，以及使用 ICMP 的网络扫描这五种攻击，并实现了对这些攻击的分类。同时，作为比较，将同样的数据包放到 Snort 系统中检测。Snort 是一个基于规则的入侵检测系统，但是其使用载荷信息进行分析，并且计算的时间很长。测试结果表明，该方案在准确率和效率上均优于 Snort 入侵检测系统。

2.2.2 使用非机器学习方法的

使用熵和评分方法检测

Akshat Gaurav 等人 (2017)[6] 提出了一种基于熵和评分机制的层次性 DDoS 攻击检测方法。该方法基于这样一个观察: 在 DDoS 攻击期间，服务器接收到的数据包头标中的源 IP 地址的熵会比平常要高。该方法同时解决了如何将 DDoS 攻击与特定时段用户的大量访问进行区分的问题。

该方法将接收到的源 IP 地址按照它们到预先设置好的中心值的距离进行分组。每个组的概率 p_i 由该组内的包个数 N_i 除以数据包总数 N 计算，即 $p_i = \frac{N_i}{N}$，那么源 IP 地址的熵表示为 $H = -\sum_i p_i log_2 p_i$; 分数由当前该组的概率与历史记录中的概率相除得到。如果最近一段时间内源 IP 的熵如果小于阈值，当前数据会被记录下来更新分数表; 如果大于阈值，那说明系统目前可能遭遇 DDoS 攻击，但也有可能是大量用户的正常访问。为了区分这两种情况，接着计算各组的分数，如果分数高于阈值，说明发生了 DDoS 攻击，不然则是正常的用户访问。

该方法综合了利用熵和计分方法的优势，做到了较低的虚警率。遗憾的是作者没有提及检测的精确率，也没有提及对于 DDoS 攻击和大量用户访问之间区分性有多好。但这种分析的方法在一众机器学习方法之中令人耳目一新。

检测新型 DDoS 攻击——目标链路泛洪攻击

K. Sakuma 等人 (2017)[5] 介绍了一种应对新型 DDoS 攻击——目标链路泛洪攻击 (Target link flooding attack) 的方法。该方法利用 traceroute 进行检测。目标链路泛洪攻击不直接针对目标区域发起 DDoS 攻击，而是对互联网中特定的链路进行攻击，从而将目标区域与其他区域隔离。

为实现这种攻击，攻击者需要对目标区域附近的网络拓扑结构进行探测。这就需要攻击者发送大量 ICMP 数据包，且这些包会集中在距离目标服务器若干跳的范围内。因此通过分析距离目标服务器的不同跳数 ICMP 数据包可以对该类型 DDoS 攻击进行探测。

首先，距离目标服务器一定范围内会部署一定量的监控器，对 traceroute 进行记录，并将包按照距离目标服务器的跳数进行分类; 总分析服务器收集各个分散的监控器的信息，计算分数。如果两个相邻时刻的分数相差超过阈值，那么就会发出攻击警报。

这种按跳数进行恩类的思想是这篇文章的亮点所在。它克服了只利用 traceroute 数量变化的方法的缺点，更加健壮。但是需要部署额外的探测器是一个问题，作者没有在文章中介绍探测器数目以及部署的策略对该方法的影响。

使用有限状态机检测 TCP 拥塞攻击

Samuel Jero 等人 (2018)[7] 提出了一种基于有限状态机的 TCP 拥塞攻击发现方法。该方法受基于模型的测试 (MBT) 和模糊测试启发，利用 TCP 拥塞控制中的有限状态机作为模型指导模糊测试，从而发现新的攻击方式。

首先，该方法在有限状态机中寻找所有的可能攻击路线; 然后将该路线转换成实际的攻击动作与数据包。这两部分别被称作抽象攻击策略与具体攻击策略。接着，作者在虚拟网络中进行了攻击测试，以判断该具体攻击策略是否确实显著地影响了网络流量。通过对可行的攻击策略进行分类，作者总结了 11 种 TCP 拥塞攻击，其中 8 种是目前文献中未被报道过的。

该方法的缺点在于，目前对于攻击的分析仍需通过日志文件由人工进行，即便可以在一定程度上自动化分析一部分，约 11% 的攻击方案仍需要人工分析。

2.3 数据链路层及物理层

本节介绍了位于数据链路层和物理层的检测方法。物理层与数据链路层位于 TCP/IP 栈的底层，不涉及高层协议，故具有通用性，但相对缺少高层协议中的语义信息。对于位于物理层的检测，通常需要与网络层或数据链路层的包或帧的信息结合，这是由于物理层缺乏必要的语义信息，比如对于帧的划分。另外，物理层的检测通常需要运用大量的通信方面的知识，例如对信道进行建模，计算信噪比等等。

位于数据链路层的检测通常适用于无线网络、工业网络等场合。对于无线网络，由于其共享介质的原因，容易受到干扰，这也是无线安全中的一大问题。针对 IEEE 802.11 协议，E. Lichtman 等人利用信噪比对具体协议干扰攻击进行检测。对于工业网络，由于其协议众多，难以针对不同协议进行统一的检测。Peter Schneider 等使用字节流进行通用检测，是解决该问题的方法之一。

直接利用字节流进行入侵检测

Peter Schneider 等人 (2018)[4] 提出了一个可对多种现场总线协议进行异常检测的实时、统一平台。该方法直接利用线路中的字节流数据进行异常检测，因此对多种上层协议有通用性。作者提出的检测方法是，将线路中的数据包按照定长进行划分，将每一个划分的字节流作为向量输入到层叠的去噪自动编码机中。其中，层叠指隐含层是多层的，去噪是指在训练时会对训练集的数据人工添加噪声，以避免过拟合。

自动编码机的架构包含编码机和解码机，训练时二者成对的被训练。损失函数定义为由解码机恢复的向量与原向量之间均方误差。对于正常的字节流，该编码机输出的向量与原始数据

之间会有很小的均方误差; 而对于异常数据，编码机的输出会有明显的误差。在给定阈值的情况下，只要是误差高于阈值的数据流都被判断为存在异常。

由于该方面免除了对包中协议内容的解析，其速度明显地优于需要解析信息的方法。同时，该方法可以解构成三个模块 (数据获取、数据预处理以及数据分析) 并行实现，因此具备了实时性的优势。

该方法实现了对较长攻击数据流的 99% 检测率。不足之处是，对于较短的攻击数据流会漏检，这是来自于该方法对数据进行批处理。

基于信噪比的干扰攻击入侵检测

E. Lichtman(2015)[9] 提出了用于检测协议感知的干扰攻击的异常入侵检测框架，可被用于 802.11 网络。协议感知的干扰攻击是指针对于 MAC 或者网络层的控制信息发起的占用网络节点通信信道的拒绝服务攻击。

文章提出的方法基于一个前提——无线接收端对于包的类型是已知的，也就是说接收端可以识别关键包和非关键包。该方法在物理层检测的基础上结合了数据链路层的协议信息，对关键包与非关键包的信噪比的统计信息进行追踪。首先，统计关键包和非关键包在一定时间段内的平均信噪比，并计算两者的比值。假设协议感知干扰存在，关键包的平均信噪比就会比较低，比值也会较低，当时间趋于无穷大时，比值趋于零。故将比值低于一定阈值的场景判定为具有协议感知的干扰攻击的存在。该框架中所设计的信噪比计算以及阈值推导，运用了大量的通信知识，这是物理层检测的一大特点。

基于 802.11 的欺骗与干扰检测

Garcia-Villegas(2015)[10] 提出了一种针对检测作弊行为和干扰攻击的方案，该方案应用在使用 IEEE 802.11 协议的局域网中。在 IEEE 802.11 中，**干扰攻击**可以防止节点正确执行 MAC 协议，或者可以导致强制重复退避的帧的冲突，因此，在 IEEE 802.11 其他客户端的干扰信号期间总是监测到到介质忙，从而无法使用信道。而**作弊行为或者欺骗性干扰**试图修改 MAC 协议的约束以获得带宽增益，使得作弊节点可以快速访问媒体介质。

该方案是利用信标访问时间 (BAT) 来检测作弊行为和干扰攻击。BAT 被定义为从信标在信标预定传送时间 (TBTT) 时间点产生并被放到传输序列头部开始，一直到该信标真正被传输的时间。可以证明，当不存在作弊行为或者干扰攻击时，BAT 是可以预测的。具体来说，将取决于与活动站相关的物理传输因子: 站的数量，传输帧的大小，物理传输速率和提供的负载。

当作弊行为和干扰攻击存在时，BAT 明显大于接入点 (AP) 的预测值，故可以通过 BAT 的偏离程度来判定是否存在作弊行为和干扰攻击。其依据为，具有作弊行为的设备可能会采取降低分布式帧间间隙 (DIFS) 值以更快地获得共享信道的访问，或者减小退避时间选择的竞争窗口的值来更快地获得增大访问信道的概率。而这两个行为都会导致 BAT 的增加。对于干扰器，BAT 值随着占用时间的增加而增加，随着静默时间的增加而减少。当占用时间比较大时，占用时间主导了对 BAT 的影响。这是无线网络中利用链路层协议检测攻击的方法。

3 未来工作

本章讨论 NIDS 未来可以发展的方向。虽然目前的 NIDS 已经能够达到十分高的准确率，但是仍存在一些固有问题:

- **数据来源**: 机器学习算法的训练需要大量数据，那么数据从何而来?
- **面向连接型网络**: 随着中间节点能力的迅猛发展，未来网络有着向连接型发展的趋势; 面对连接型网络，NIDS 该何去何从?
- **协议栈各层**: 当前 NIDS 在各层上的实现仍存在一些问题; 同时层与层之间的协作应该如何实现?
- **NIDS 检测的攻击类型**:NIDS 应该如何处理未知的攻击以及蓄意构造的针对性攻击?
- **使用非统计方法**: 机器学习或者统计的方法有其天生的弊端，结合非统计方法是否回台给我们新的启发?

此外，目前设计的 NIDS 还存在**实用性**的问题——随着网络流量的不断增大，NIDS 采用的检测算法是否能有效应对现实中的网络流量? 以下将针对这些问题进行详细讨论。

3.1 关于数据来源

对使用机器学习或者统计的方法的入侵检测系统，需要大量的数据，比如第2.1.1小节中提到的 DGA 分析算法，它需要大量的 DNS 数据集作为支撑。而在论文中，作者们总是理所当然地获得了这些数据，从中提取了所需要的特征。但这些数据并不是这么容易获得的，有些甚至是受法律限制的。因此，我们认为关于何让获取入侵数据的研究是一个值得在未来研究的方向。虽然这并非直接涉及到入侵检测，但却是做好入侵检测所必不可少的铺垫性工作。

我们设想了两大研究领域，包括:

1. **数据收集**

   我们如何在法律允许的前提下，有效地收集各种数据，并形成一个开放的数据库供安全工作者们使用。这不仅为在研究新型方法的人们提供了便利，让他们不需要花费太多无谓的精力在数据收集上，还使得已提出的各种受限于数据来源的方法有可能投入到实际使用中。

2. **数据生成**

   另一方面，受到了 DGA 的启发，我们想到数据集除了收集之外，是否还可以通过一个算法来生成攻击数据呢? 我们将之命名为攻击生成算法 (AGA)。如果我们可以设计一个有良好性能的 AGA，可以根据需要伪随机地生成特定攻击类型的数据，那么相关研究的数据问题就可以得到根本性的解决了。当然，生成的数据是否符合实际，是否真的随机等问题肯定是存在的，这需要我们日后的研究来逐一解决。

3.2 关于面向连接型网络

目前网络传输中占据主导地位的是 TCP/IP 协议族，因此在讨论入侵检测时，一般来说也是基于 TCP/IP 协议族来进行讨论的; 部署检测系统或者方案时，也常常将目光放在计算能力较强的端点处。

但随着网络节点计算、存储能力的提升，倾向连接型网络的协议，比如 ATM 与 MPLS，可能会得到进一步的发展。因此在未来，我们应该将一部分精力分配到适应连接型网络的入侵检测系统。比如，我们可以很自然的将 IDS 部署在各个**中间节点**当中，并使得节点可以将它们检测到的信息进行相互交换、整合，从而获得比只在端点检测的系统更多，更全面的对网络状况的掌握。这样一来，攻击者产生的恶意流量会更加难以躲开安全系统的检测，甚至在到达客户端之前，就会在传输过程中被某个节点过滤。从安全工作者的角度来说，他们能使用的方案、策略会更加灵活多样。

另一方面，由于这种网络是面向连接的，通信双方都维护了连接的状态，这样一来 IP 网络中一个非常严重的问题: **源地址的不确定性**就得到了一定程度的解决。直接地，一些典型的与源地址修改相关的攻击的威胁会大大降低，例如 IP 地址欺骗、DRDoS 攻击等等; 间接地，在有了源地址可靠这么一个安全前提之后，我们就不需要将精力太多投入到源地址的确认上，这样自然可以提出一些新的更可靠的检测方案。

3.3 关于协议栈各层

1. 应用层 IDS 缺少统一性方案

在本文涉及的文章中，面对 DNS 服务、HTTPS 协议、浏览器插件广告插入以及电子邮件等，解决的方案都是特定的、针对性的。这是受应用层特性决定的: 由于应用层希望为用户提供多种多样的服务，因此应用的数据格式高度分化，很难找到统一的监测方案。是否能找到一种通用性的解决方案为应用层提供更好的安全保障呢?

受到 HTTPS 下层的 TLS 协议启发，由于上层注定无法统一，因此从其下层进行处理是更好的策略。目前有许多流行的中间件，比如 Google 的 QUIC 协议，利用这些中间件为上层应用层服务提供更好的安全性是很有意义的研究方向。

是否可以设计一种能满足多种不同应用层需求的协议? 它可以灵活的组装各种模块，就像 IPSec 中的 IKE 一样，可以为差异很大的不同应用提供服务; 而它又可以有效地被 IDS 检测和分析。这种**中间件安全**引人思考。

2. 针对数据链路层以及物理层的监测方案较少

在寻找合适的文章过程中，一个现象是很少有针对数据链路层或物理层进行安全监测的方案。这样的检测是有益的，考虑到:

1. **应对信息受限情况**

   正如 Peter Schneider 等人这篇文章所述，在安全检测时，有时对上层协议不能获得完全的了解和知识，一个通用性的解决方案可以处理这样的问题; 同时通用性也避免了重复开发的花费。

2. **处理上层加密情况**

   TLS 被越来越多的应用所使用;IPSec 也可以对数据包进行保护; 但是他们有共同的特点: 数据内容被加密。这虽然为用户信息安全提供了一定的保障; 但是也为攻击者提供了方便——加密的数据意味着许多方法不能被利用。而直接在数据链路层或物理层展开监测可以处理加密信息的安全监测。

3. **对下层攻击的预防**

   在安全领域中，很常见的一种攻击方法是下层攻击。虽然本层的防御体系已经十分完善，但是对于来自下层的攻击还是力有不逮。如果，攻击者直接对电磁信号进行调整或伪造，或者在链路层中对帧进行修改，都可能会使上层的防御手段失效。

但由于数据链路层以及物理层所包含的语义信息较少，各种检测方法有时难以实现较高的准确率。但这不代表这一层上不能够或不应该进行入侵检测。

3. 各层之间的协作

受到深度包检测 (DPI) 启发，充分利用各个层之间的信息也对实现更好的入侵检测是很有帮助的。Peter Schneider 等人 [4] 文章中所述，一些针对工业联网设备的攻击是控制逻辑上的攻击，也即其内容本身可能符合安全标准，但是多个包连续的组合会产生特定的攻击。这样的攻击，如果能通过最初对字节流的检测，然后更进一步对报文内容进行分析处理，想必可以更进一步提高准确性，并降低虚警率。

如果凭借本层信息不能完全确定，那么可以将本层的分析结果传送到上层的检测系统中，用于辅助进一步的分析。这样的协作模式可以对检测系统的提升有很大帮助。但是这同样会面里一个问题: 开销的增长。面对越来越庞大的数据流，多层协作的 IDS 能否实时地对系统进行保护? 这需要平衡协作的程度与开销，达到一个令人满意的程度。

3.4 关于 NIDS 面向的攻击类型

在入侵检测中我们所研究的各个类型的攻击中往往可以分为三类，它们是**已知的攻击**，**未知的攻击**，**针对已有入侵检测系统的漏洞的针对性攻击**。目前在研究中最成熟的是第一种攻击，我们可以针对性的提出各种各样的方法来检测，这样对于这一种特定攻击来说，检测效果一般会非常好。但良好的针对性往往牺牲的是通用性，在面对后面两种攻击的时候，这些检测方法的效果就会大打折扣。特别是最后一种，魔高一丈的入侵者在找准漏洞后所作出的改进后的攻击，将会在最大程度上削弱检测系统的效果。从这个角度上来说，后面这两种攻击会带来更大的安全威胁。

因此如何提出一种检测方法或者检测模型，使得它可以对可能出现的未知攻击具有很好的普适性是我们值得研究的方向。比如说，当它遇到未知的数据时，能否对它进行分析建模，主动学习，化未知为已知。一方面可以及时与安全社区分享新的攻击手段; 另一方面，若日后发生了类似的安全时间，则可以据此快速响应，避免危害的进一步扩大。

此外，由于大多数入侵检测系统使用机器学习的方法，对抗攻击成为了一个必须注意的攻击手段。当攻击者针对性地对发起此类攻击，检测系统将很大可能失去检测能力，直到产生较

大危害时才会被其他手段注意到。综上所述，在未知攻击和对抗攻击的方面仍有工作等待解决。

3.5 关于使用非统计方法

目前的入侵检测系统的检测方法主要有**基于规则的**和**基于统计的**。随着人工智能的兴起和算力的大幅度提高，以传统机器学习和深度学习为代表的统计方法是目前研究的主流方向。基于规则的入侵检测系统的弱点在于其性能依赖于其系统中的规则数目，而规则越多，系统的处理时间越长，效率越低。同时，由于其无法识别未知攻击，所以**漏检率**比较高。

而与基于规则的入侵检测系统不同，基于统计的方法可以检测未知攻击，但是会带来比较高的**误报率**。从理论上来看，漏检率和误报率是无法同时降低的。故如何在这两者之间进行折中处理，以符合实际的使用，是需要被关注的。

受此启发，非统计学的方法也许可以做到同时降低漏检率和误报率。由于网络中的数据主要是信息流，各种行为模式皆以信息的方式呈现，引入**信息论**的知识，也是未来的方向之一。信息论已被广泛地运用于检测 DoS 攻击的研究中，并表现出了比其他方法更低的误报率。如何将信息论更广泛的运用于其他的攻击检测，也是值得探讨的方向。

随着网络节点计算能力的增加，网络拓扑图的信息能够更好的被传递和储存，这使得**图论**可能成为未来入侵检测的重要方法。在计算能力大幅提升的背景下，将图论和统计结合也是一大未来的方向。

此外，传统的机器学习需要大量的人工特征提取，而这有赖于人们的先验知识，虽然对比于深度学习，前期的准备工作量更大，但更有利于实现更精准的检测，也更适用于场合特定且数据量不大的攻击检测。在统计方法中，通常需要事先将大量的数据进行预处理，这些处理过程通常是基于规则的，故如何结合基于规则和基于统计的检测系统，并发挥各自的优势，也是值得探讨的。

4 其他

本文准备全过程可以在这个Github repo中更详细的了解，包括选题、选定论文、阅读分享以及写作润色等。

本文的分工如下:

第2.1.1节中导言，利用 NXDomian 流量检测 DGA，通过网页元素表示检测广告注入、第3章中关于数据来源、关于面向连接型网络、关于 NIDS 面向的攻击类型及文章润色由李锭轩同学负责。

第2.1.1节中基于 HTTP 的恶意软件检测，负载聚合技术，鱼叉式钓鱼邮件检测、第2.2节使用数据报报头信息进行检测、第2.3节中导言，基于信噪比，802.11 的检测及文章整体润色由林静雯同学负责。

第1章、第2.2节中导言，使用 TLS 证书，熵和评分方法，目标链路泛洪攻击，有限状态机检测进行检测、第2.3节直接利用字节流进行入侵检测、第3章关于协议栈各层内容、文章整体润色及协调统筹由胡冬寅同学负责。

本文的完成与组内每一位成员的努力密不可分，在此向每位成员表示衷心的感谢。同时，感谢您的阅读。

参考文献

[1] Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., ... & Kumar, D. (2017). Understanding the mirai botnet. In 26th USENIX Security Symposium (USENIX Security 17) (pp. 1093-1110)

[2] Arshad, S., Kharraz, A., & Robertson, W. (2016, September). Identifying extension-based ad injection via fine-grained web content provenance. In International Symposium on Research in Attacks, Intrusions, and Defenses (pp. 415-436). Springer, Cham.

[3] Torroledo, I., Camacho, L. D., & Bahnsen, A. C. (2018, October). Hunting Malicious TLS Certificates with Deep Neural Networks. In Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (pp. 64-73). ACM.

[4] Schneider, P., & Böttinger, K. (2018, October). High-Performance Unsupervised Anomaly Detection for Cyber-Physical System Networks. In Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy (pp. 1-12). ACM.

[5] Sakuma, K., Asahina, H., Haruta, S., & Sasase, I. (2017, December). Traceroute-based target link flooding attack detection scheme by analyzing hop count to the destination. In 2017 23rd Asia-Pacific Conference on Communications (APCC) (pp. 1-6). IEEE.

[6] Gaurav, A., & Singh, A. K. (2017, May). Entropy-score: A method to detect DDoS attack and flash crowd. In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (pp. 1427-1431). IEEE.

[7] Jero, S., Hoque, E., Choffnes, D., Mislove, A., & Nita-Rotaru, C. (2018, July). Automated attack discovery in TCP congestion control using a model-guided approach. In Proceedings of NDSS.

[8] Felix Erlacher, Falko Dressler, "High Performance Intrusion Detection Using HTTP-Based Payload Aggregation" Local Computer Networks (LCN) 2017 IEEE 42nd Conference on, pp. 418-425, 2017.

[9] Lichtman, M, Reed, JH, "Anomaly-based intrusion detection of protocol-aware jamming." IEEE Military Communications Conference (MILCOM'15); October 26-28, 2015;Tampa, FL, USA.

[10] E. Garcia-Villegas, M.S. Afaqui, E. Lopez-Aguilera, "A novel cheater and jammer detection scheme for IEEE 802.11-based wireless LANs." Elsevier J Comput Netw, 86 (2015), pp. 40-56.

[11] Gascon H., Ullrich S., Stritter B., Rieck K, "Reading Between the Lines: Content-Agnostic Detection of Spear-Phishing Emails." In: Bailey M., Holz T., Stamatogiannakis M., Ioannidis S. (eds) Research in Attacks, Intrusions, and Defenses. RAID 2018. Lecture Notes in Computer Science, vol 11050. Springer, Cham

[12] N. Chockwanich and V. Visoottiviseth, "Intrusion Detection by Deep Learning with TensorFlow," 2019 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang Kwangwoon_Do, Korea (South), 2019, pp. 654-659.

[13] Zhang J., Jang J., Gu G., Stoecklin M.P., Hu X. (2018) Error-Sensor: Mining Information from HTTP Error Traffic for Malware Intelligence. In: Bailey M., Holz T., Stamatogiannakis M., Ioannidis S. (eds) Research in Attacks, Intrusions, and Defenses. RAID 2018. Lecture Notes in Computer Science, vol 11050. Springer, Cham