

基于时空特征融合深度学习的NXNSAttack流量识别方法

刘宇翔¹, 莫秀良¹, 宫良一^{2,3}, 吕昆娴^{2,3}, 左鹏⁴

(1. 天津理工大学计算机科学与工程学院, 天津 300384; 2. 中国科学院计算机网络信息中心, 北京 100083;
3. 中国科学院大学, 北京 101408; 4. 中国互联网信息中心, 北京 100070)

摘要: 域名系统 (domain name system, DNS) 是互联网的重要基础设施, 也是网络攻击的重灾区。作为 DNS 的重要组成部分, 权威服务器对域名解析起到关键作用, 也因此成为攻击者的重点攻击对象。研究人员发现, 攻击者可以利用 DNS 递归解析器的漏洞发起针对权威服务器的基于 NXNSAttack 的分布式拒绝服务 (distributed denial of service, DDoS) 攻击。此外, 该研究发现一种新的基于 NXNSAttack 的变种 DDoS 攻击, 攻击者可以利用分布式自建权威服务器发起 NXNSAttack。该攻击利用 DNS 递归解析器在 NS (name server) 记录查询关联分析方面的缺陷, 以空间协同方式发起分布式 NXNSAttack (D.NXNSAttack)。针对上述两种 NXNSAttack 变种, 提出了一种基于时空特征融合深度学习检测方法。所提方法包括流量采集、数据预处理、时空特征融合学习及攻击分类等模块。在流量信息图像化基础上, 利用空间学习模型 ShuffleNet 和时序学习模型 Mamba 的学习优势来捕获正常流量与攻击流量的图像特征的差异, 实现对两种不同类型的 NXNSAttack 流量的检测。基于大量网络靶场攻击流量的实验结果显示, 所提方法的检测精确率超过 98%、F1 分数达到 98.8%。

关键词: NXNSAttack; 域名系统; DDoS 攻击; 权威 DNS

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025050

NXNSAttack traffic detection method based on spatio-temporal feature fusion deep learning

LIU Yuxiang¹, MO Xiuliang¹, GONG Liangyi^{2,3}, LYU Kunxian^{2,3}, ZUO Peng⁴

1. School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China

2. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100083, China

3. University of Chinese Academy of Sciences, Beijing 101408, China

4. China Internet Network Information Center, Beijing 100070, China

Abstract: The domain name system (DNS) is recognized as a critical infrastructure of the Internet and is frequently

收稿日期: 2025-01-14; 修回日期: 2025-05-15

通信作者: 宫良一, lygong@cnic.cn

基金项目: 国家重点研发计划青年科学家项目 (2023YFB3105700); 国家自然科学基金 (62272440)

Foundation Items: The National Key R&D Program for Young Scientists of China (2023YFB3105700), The Natural Science Foundation of China (62272440)

引用格式: 刘宇翔, 莫秀良, 宫良一, 等. 基于时空特征融合深度学习的NXNSAttack流量识别方法[J]. 网络与信息安全学报, 2025, 11(4): 173-188.

Citation Format: LIU Y X, MO X L, GONG L Y, et al. NXNSAttack traffic detection method based on spatio-temporal feature fusion deep learning[J]. Chinese Journal of Network and Information Security, 2025, 11(4): 173-188.

targeted by cyber-attacks. As an essential component of the DNS system, authoritative servers played a key role in domain name resolution and were therefore subject to frequent attacks. Recently, researchers found that vulnerabilities in DNS recursive resolvers could be exploited to launch distributed denial-of-service (DDoS) attacks based on NXNSAttack against authoritative servers. Furthermore, a new variant of DDoS attack based on NXNSAttack was identified in this study. Attackers were found to utilize distributed self-built authoritative DNS servers to launch the NXNSAttack. This attack leveraged the vulnerability of DNS recursive resolvers lacking correlation analysis for NS queries, and distributed NXNSAttack attacks (D.NXNSAttack) were conducted through spatial collaboration. To address these two types of NXNSAttack-based DDoS attacks, a detection method based on the fusion of spatial and temporal features using deep learning was proposed. The method comprised modules for traffic collection, data preprocessing, spatial-temporal feature fusion learning, and attack classification. By converting traffic information into images, the advantages of the spatial learning model ShuffleNet and the temporal learning model Mamba were utilized to capture differences between normal and attack traffic, enabling the detection of two distinct types of NXNSAttack attacks. Experimental results based on large volumes of attack traffic from multiple network ranges show that the detection accuracy of the proposed method exceeds 98%, and the $F1$ -score reaches 98.8%.

Keywords: NXNSAttack, DNS, DDoS attack, authoritative DNS

0 引言

随着互联网的快速发展, 域名系统 (domain name system, DNS) 已成为网络基础架构中不可或缺的组成部分。作为互联网的核心服务之一, DNS 的主要功能是将用户输入的域名解析为对应的资源地址^[1], 从而实现资源的精确定位和高效访问。然而, 正是由于其基础性和重要性, DNS 已成为网络攻击的主要目标^[2-3]。其中, 分布式拒绝服务 (distributed denial of service, DDoS) 攻击尤为典型且危害严重。

DNS 递归解析域名的过程如图 1 所示, DNS 采用分层查询机制, 递归解析器会按照域名层级结构依次向各级域名服务器发送查询请求, 最终从二级域名服务器获取域名对应的资源记录^[4]。由于二级域名服务器存储着最具体的域名解析信息, 其在 DNS 层级结构中处于关键位置, 因此更容易成为网络攻击的目标。

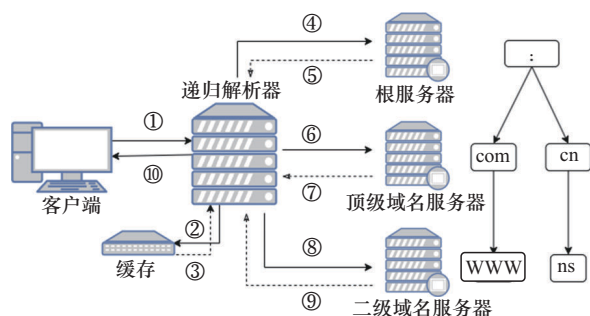


图 1 DNS 递归解析域名的过程

Figure 1 Process of DNS recursive resolution of domain names

NXNSAttack 是一种新型的 DNS 攻击技术, 它通过利用 DNS 递归解析器在设计上的安全缺陷, 对权威服务器发起 DDoS 攻击^[5]。由于当前 DNS 协议针对名称服务器 (name server, NS) 记录的查询缺乏有效的验证机制, 攻击者可以向递归解析器发送大量的域名解析请求, 并在每个查询请求中引入伪造的子域名记录。如图 2 所示, 递归解析器在收到这些伪造的域名解析请求后, 会按照标准 DNS 解析流程尝试解析这些不存在的子域名。由于这些子域名在 DNS 层级结构中并不存在, 递归解析器无法获得有效的结果便会持续向权威服务器发送 NS 记录查询请求^[6], 这种无效查询的持续累积会快速消耗系统资源。攻击者正是利用大量递归解析器对无效 NS 记录的无限制递归查询这一特性, 导致权威服务器性能显著下降甚至宕机, 使其成为一种有效的 DNS DDoS 攻击方式。

为抵御基于 NXNSAttack 的 DDoS 攻击, 各大 DNS 软件服务商推出了防御策略, 其中最具有代表性的是 MaxFetch(k)策略^[5]。该策略引入了查询限制机制, 规定每个原始客户端请求最多能触发 k 次额外的 NS 记录查询, 从而显著减少处理大量无效 NS 记录而导致的资源消耗。然而, 由于互联网中大量递归 DNS 服务器未及时更新软件版本, 上述防御策略未能得到广泛应用, 因此基于 NXNSAttack 的 DDoS 攻击仍存在较大潜在风险, 相关威胁尚未完全消除。

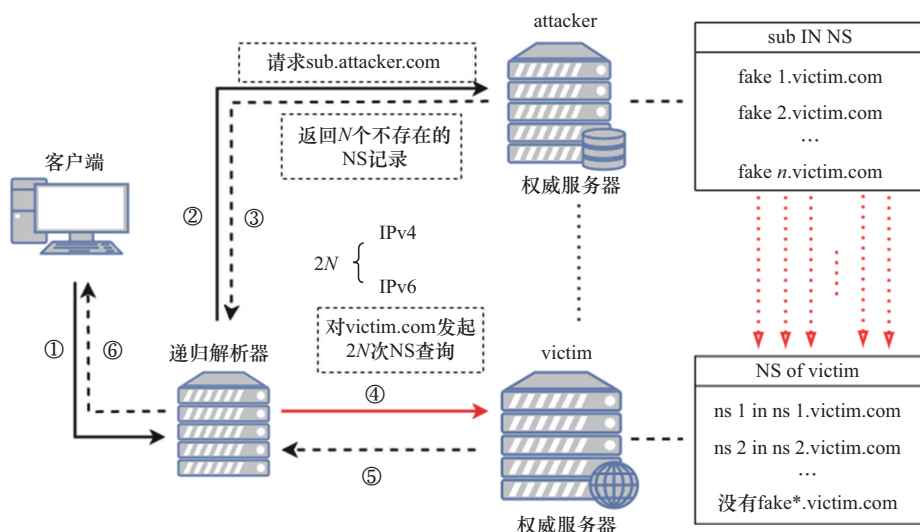


图2 NXNSAttack示例

Figure 2 An example of NXNSAttack

尽管现有的修复策略已对基于NXNSAttack的DDoS攻击有所防御，但递归DNS在NS记录关联性处理机制上仍存在根本性缺陷。当面对无法解析的域名时，递归解析器仍会机械式地查询权威服务器指向的所有NS记录，且难以关联识别出“多个虚假域名查询指向同一权威服务器”这一特征。基于这一发现，本文提出一种新型的NXNSAttack。攻击者通过部署大量伪造的权威服务器，并为合法域名构建大量虚假子域名的NS记录，再向递归解析器持续发起不同攻击域名的查询请求。这一过程会诱发大量虚假NS记录查询，迫使目标权威服务器陷入频繁响应的状态，最终实现与NXNSAttack类似的资源耗尽效应。

与传统的集中式NXNSAttack（centralized NXNSAttack, C.NXNSAttack）相比，新型的分布式NXNSAttack（distributed NXNSAttack, D.NXNSAttack）能够产生更加显著的群体攻击效应。本文在修复与未修复NXNSAttack漏洞的BIND服务器上开展了攻击对比实验。实验结果显示，D.NXNSAttack在两种环境下均能显著消耗服务器资源，充分体现了其潜在威胁。

尽管目前已有诸多研究^[7-11]提出了基于机器学习和深度学习的DNS流量检测方法，但这些方法大多未考虑攻击流量在空间结构和时间行为上的协同特征，导致其在处理NXNSAttack及其

变种时存在准确率下降和识别能力不足的问题。针对这一局限性，本文创新性地提出了一种基于DNS流量时空特征融合的深度学习检测方法。首先，该方法对DNS递归解析器查询流量进行捕获和提取特征，并采用流量图像化技术提取时间窗口内的流量用来构建流量特征灰度图。其次，考虑到C.NXNSAttack与D.NXNSAttack流量与正常DNS查询流量具有高度相似性，该方法进一步引入了时空特征融合学习机制，借助ShuffleNet的空间学习能力与Mamba的时序学习能力，增强对攻击流量和正常流量的识别能力。最后，通过融合分类标记获得3类流量的概率分布，实现高效的3类别分类。

为验证所提方法的有效性，本文在半开放网络靶场环境下进行了流量采集和攻击模拟实验。实验过程中，基于搭建的DNS共采集了65万个样本，包括正常流量、C.NXNSAttack流量和D.NXNSAttack流量。实验结果表明，该方法的检测精确率超过98%、F1分数达到98.8%，展现出优异的检测性能。

1 相关工作

DNS作为互联网的核心基础设施之一，正面临着日益严峻的安全挑战。Liu^[12]等通过对NX-Domain响应进行深入分析，揭示了无效响应在DNS中普遍存在的现象及其潜在的安全风险。研究发现，频繁接收NXDomain响应的域名存在较

高的安全风险, 这些域名可能被恶意利用于僵尸网络控制、恶意软件分发等攻击活动, 这凸显了对这类域名进行实时监控和主动防御的必要性。值得注意的是, 当前出现的 NXNSAttack 正是利用海量 NXDomain 响应引发流量放大效应, 对目标系统发起 DDoS 攻击。Zheng^[13]等重点探讨了胶水记录对 DNS 安全的影响, 描述了递归解析器处理胶水记录的过程, 并指出胶水记录可能导致产生缓存污染漏洞。Moura^[14]等对 TsuNAME 漏洞进行了分析。该漏洞是一种利用 DNS 记录中的循环依赖关系实施的放大攻击, 可能导致部分递归解析器大幅放大查询流量, 进而引发 DDoS 攻击。Afek 等^[15]发现的 NRDelegationAttack 利用 DNS 解析器处理委托响应的漏洞, 通过恶意构造的请求让解析器 CPU 负载急剧增加, 造成系统资源耗尽。

为有效防御上述 DNS 攻击, 基于模式识别的攻击流量识别技术成为当前的研究热点。现有的 DNS 流量识别方法主要分为两类: 一类基于机器学习方法^[7-8], 这类方法侧重于从数据包或流级别提取统计特征; 另一类基于深度学习方法, 这类方法通过解析原始流量数据的结构与内容, 过滤冗余信息来构建定制化训练数据集^[9]。

Niktabe 等^[10]提出一种基于统计模式识别的 DoH (DNS over HTTPS) 恶意流量检测方案, 采用线性回归和逻辑回归模型, 通过数据预处理和模式识别两个阶段识别 DoH 恶意流量的特征模式。Tu^[11]等提出一种基于 DNS 查询与响应对 (Q&R) 的 DNS 隧道检测方法, 通过结合编码特征与行为特征实现 DNS 隧道流量识别。史国振等^[16]提出一种基于深度残差胶囊网络 (deep residual capsule network, DRCN) 与注意力机制的加密流量识别框架, 结合 3D 卷积动态路由、通道注意力机制和残差网络, 显著提升了检测准确率和效率。Li 等^[17]设计了一种融合改进 Transformer 和卷积神经网络的恶意 DNS 流量检测模型, 通过优化 Transformer 的注意力头和编码结构, 并采用基于块的集成分类器实现恶意 DNS 流量检测。然而, 基于 Transformer 的预训练流量基础模型因其自注意力机制具有平方级复杂度^[18], 在实际应用中面临着实时流量分类的计算效率和内存占用等挑战。

2 分布式 NXNSAttack 原理

本节将详细阐述递归解析器的脆弱性、分布式 NXNS 威胁模型, 以及在存在关联性的多个 DNS 权威服务器环境下, 基于分布式 NXNSAttack 的 DDoS 攻击原理。

2.1 递归解析器的脆弱性

当递归解析器查询某个域名时, 若目标服务器返回的响应中仅包含用于指引下一步查询的引荐信息 (即 NS 记录), 而非最终的解析结果, 则递归解析器会依据这些 NS 记录中的信息发送新的查询请求, 这种行为被称为 “NS 引荐响应”^[5]。当递归解析器收到一个未关联 IP 地址的引荐响应时, 它会启动额外的解析流程, 以获取这些 NS 记录所指向服务器的 IP 地址。

旧版 DNS 服务器软件 (如 BIND 9.11.18) 在处理 DNS 引荐响应时存在一个安全漏洞^[5], 即未对单次引荐响应中的 NS 记录数量进行限制。攻击者可以利用该漏洞构造特定的引荐响应, 诱导递归服务器发送大量查询以处理这些伪造的引荐请求, 进而引发 NXNSAttack。为消除 NXNSAttack 的潜在风险, 新版 DNS 服务器软件 (如 BIND 9.11.19) 引入了 NS_RR_LIMIT 机制。当递归解析器接收到包含多个 NS 记录的引荐响应时, 针对每个客户端查询最多解析 k 条 NS 记录。如果连续 k 次响应均返回 NXDomain, 则终止递归解析流程。

为探究递归解析器处理引荐响应的行为, 本文构建了一个网络靶场, 包括 4 台权威服务器、1 台递归解析器和 1 台客户端服务器, 注册 5 个新域名并与权威服务器进行绑定。所有服务器均使用 BIND 9.18.28 版本, 该版本修复了 NXNSAttack 漏洞, 并在安全性和性能方面进行了优化。实验模拟了不同的引荐响应场景, 测试递归解析器在新旧机制下的行为表现。

为分析递归解析器处理多级 NS 记录的行为, 实验配置了 3 层权威服务器结构, 如图 3 所示。以第 1 层权威服务器的子域 “www.auth1.com” 为例, 其 NS 记录指向第 2 层的 “fake.auth2.com” 和 “fake.auth3.com”, 而这两个子域的 NS 记录 (www.auth2.com, www.auth3.com) 又进一步指向第 3 层权威服务器。需要说明的是, 所有 NS 记

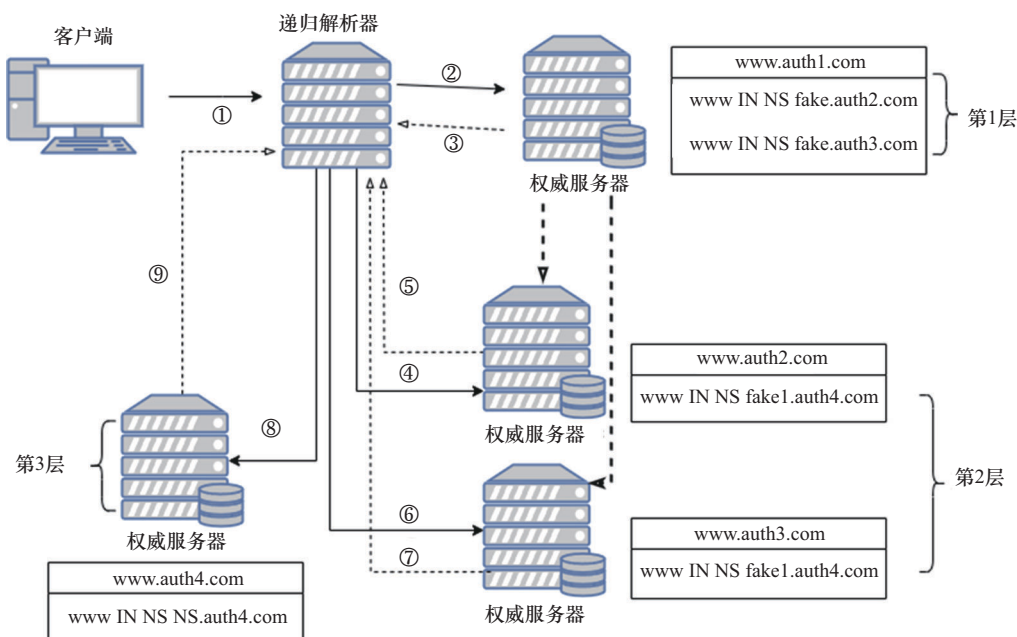


图3 3层权威服务器结构

Figure 3 Three-tier authoritative server architecture

录均未配置对应的IP地址。

实验中，客户端首先查询“www.auth1.com”的A记录（步骤①），递归解析器则根据引荐响应依次与各层权威服务器进行交互（步骤③、④、⑥、⑧）。使用LiBPCAP工具捕获并分析相关数据包及日志，发现递归解析器存在以下行为特征。

1) 接收到第1层的引荐响应后，递归解析器会继续向第2层权威服务器发起查询。

2) 在未获取A记录但接收到不含IP地址的NS记录时，递归解析器并未按照预期返回NX-Domain响应，而是继续向下一层发起查询，进而形成引荐响应的递归查询循环。

3) 当递归解析器收到攻击者精心设计的NS引荐响应时，每查询一条NS记录都会进行 k 次解析，最终触发的查询总数超过了NS_RR_LIMIT机制的限制。

上述行为表明，递归解析器虽能成功阻断大量NS引荐响应查询，但在NS递归查询的有效关联性分析方面仍存在短板。

2.2 D.NXNSAttack威胁模型

由于递归DNS难以对NS记录的虚假递归查询进行有效关联分析，攻击者可以利用大量伪造的授权服务器发起NXNSAttack，实现D.NXNSAttack。

为实现基于递归解析器和权威服务器的D.NXNSAttack，攻击者需要控制一台或多台具备互联网访问能力并能够发起DNS查询的客户端服务器。这些服务器将作为攻击的入口点，负责向递归解析器发送大量的DNS查询请求。此外，攻击者需要控制大量的二级域名权威服务器。由于权威服务器的操控者可以自由构造NS引荐响应，并将其作为DNS查询的回复，因此攻击者能够精确地控制引荐响应中涉及的服务器数量和目标地址。即使递归DNS已配置NS限制策略，攻击者依然能够构建一个高度关联的分布式分层授权查询网络，进而发起大规模的NXDomain攻击。

2.3 基于D.NXNSAttack的DDoS攻击

递归解析器缺乏对NS递归查询的关联性分析能力，攻击者可以利用递归DNS的这一脆弱性构建基于分布式NXNSAttack的DDoS攻击。该攻击的形成，源于递归解析器中多种DNS机制的相互作用，具体如下。

(1) 搭建分布式分层授权解析网络

攻击者需要构建一个分布式分层授权解析网络。首先选定一台权威服务器（如图4中的nxns-boss服务器），将其配置为负责接收DNS查询的域名权威服务器。该权威服务器的职责是将NS

引荐响应分发到下一层的权威服务器，而下层服务器可以进一步将查询分发到更深的层级，或者直接将NS记录指向受害者的权威服务器。

(2) 发起恶意 DNS 查询

如图4所示，攻击者通过客户端向递归解析器发起针对恶意子域名“sub.nxnssboss.com”的查询（步骤①）。对于该子域名下的任何DNS查询，都会被委派到第2层的恶意权威服务器。当递归解析器试图解析该子域名时，首先将检查缓存中是否存在这些NS记录的有效解析结果。如果缓存没有命中，那么递归解析器必须先解析这些NS记录。

(3) 递归解析器的异常查询行为

由于NS_RR_LIMIT机制的存在，每次委派查询触发NS引荐响应时，递归解析器最多会查询前 k 台NS服务器。作为对比，在BIND 9.11.18版本（未修复NXNSAttack漏洞），递归解析器会逐一查询配置的所有NS服务器（步骤④）。每次查询都会被递归转发到下一层权威服务器，进而触发新一轮的NS查询。这些查询最终会指向

受害者（victim）的解析器，导致资源消耗。虽然DNS存在一定的查询数量限制，可避免攻击放大系数呈指数级增长，但攻击规模仍足以对目标系统造成显著影响。当递归解析器的查询最终到达受害者的权威服务器时（步骤⑥），受害者的解析器会被迫处理大量无效或伪造的查询请求。这些查询来自攻击者精心设计的分布式分层NS解析网络，并通过递归解析器逐步放大。受害者服务器会不断收到NXDomain响应，直到超时或达到查询最大限制，最终返回FAIL（步骤⑧）。

2.4 D.NXNSAttack评估

本节将计算D.NXNSAttack的成本及其放大效益。首先，定义受害者服务器在每轮攻击中接收和发送的数据包总量为 C_v^{pkt} ，由攻击者发送的数据包总量为 C_a^{pkt} 。此外，定义 P_{max} 为在一次客户端查询中，攻击者引发的数据包请求的最大数量。定义 b_{max} 为递归解析器在解析一个客户端请求时可发送的最大请求数，默认设置为75。当开启诸如DNSSEC等安全机制时，DNS响应可能会超过UDP数据包的常规大小限制（一般为512 byte），

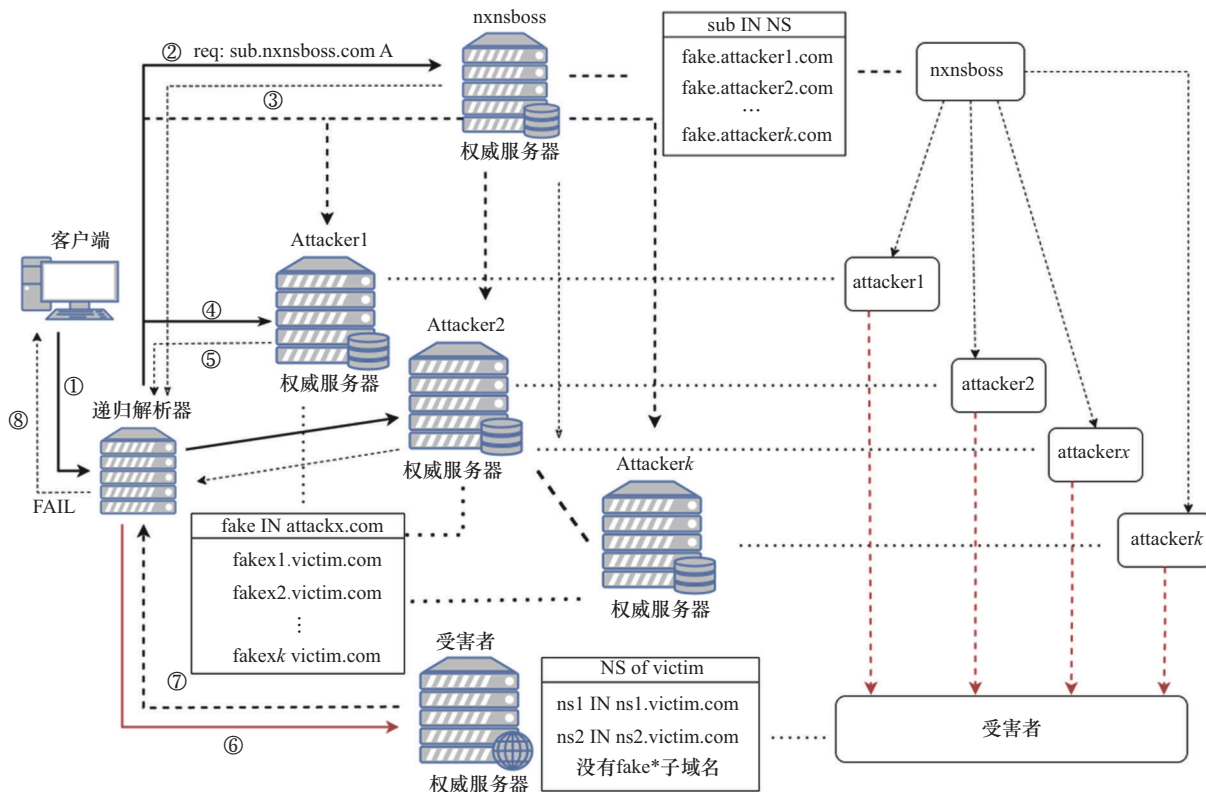


图4 3层权威服务器的D.NXNSAttack示例

Figure 4 Example of D.NXNSAttack on three-tier authoritative servers

此时需要通过TCP传输。由于目标服务器无法建立TCP连接，不仅增加了额外的TCP控制数据包，还可能引发TCP重传机制。本文使用 r 表示TCP重传的次数，并假设 r 的值为5。

为了衡量D.NXNSAttack的攻击效果，本节使用包放大系数（packet amplification factor, PAF）计算受害者处理的数据包 C_v^{pkt} 与攻击者发送的数据包 C_a^{pkt} 之间的比例。PAF值越大，意味着攻击者能够用较少的数据包诱使受害者处理更多的数据包，攻击的放大效果越显著。其定义如下：

$$PAF = \frac{C_v^{pkt}}{C_a^{pkt}} \quad (1)$$

以一个3层NS网络且未触发TCP机制的D.NXNSAttack为例，攻击者控制客户端向递归解析器发起一个恶意请求后，递归解析器会按规则查询前 k 台NS服务器，每个查询均需请求恶意权威服务器的IPv4和IPv6地址，此时共计产生 $(1 + 2k)$ 个数据包。对于第2层权威服务器产生的 $2k$ 个查询，每个查询同样会请求下一层权威服务器的IPv4和IPv6地址，因此第2层共产生 $4k^2$ 个数据包。 k 值越大，攻击者通过单次请求触发的数据包数量越接近递归解析器的最大请求阈值 b_{max} 。因此，在该3层NS网络中，攻击者服务器一次查询触发的数据包请求数量为：

$$P_{max} = \begin{cases} 1 + 2k + 4k^2, & P_{max} < b_{max} \\ b_{max}, & P_{max} \geq b_{max} \end{cases} \quad (2)$$

本文对实验环境中部署的BIND服务器进行了测试，其 k 值默认为5。攻击者通过3层NS网络发出的最大数据包数量为 $1 + 2k + 4k^2 = 111$ ，该数值大于递归解析器的最大请求阈值 b_{max} 。此时， $P_{max} = b_{max} = 75$ 。

(1) 针对二级权威服务器的攻击

由式(2)可知，攻击者针对权威服务器发起攻击时共引发75个请求数据包。其中有6个请求数据包被发送至攻击者控制的权威服务器，69个请求数据包被发送给受害者服务器。此时受害者服务器需要处理的数据包总数为：

$$C_v^{pkt} = \begin{cases} 2(P_{max} - k - 1)(1 + r), & TCP \\ 2(P_{max} - k - 1), & UDP \end{cases} \quad (3)$$

由于本实验设置的受害者权威服务器未启用DNSSEC安全机制，且攻击场景中未涉及TCP重

传，此时 $C_v^{pkt} = 2(P_{max} - k - 1) = 138$ 。在上述攻击中，攻击者服务器需要发送两种类型的数据包，一种是由攻击者控制的客户端发起的一个初始查询请求，另一种是权威服务器发起的 k 次NS引荐响应。因此，攻击者发送的数据包总量为 $C_a^{pkt} = 1 + k = 6$ ，由式(1)可知， $PAF = 23x$ 。

尽管此时攻击的PAF值少于C.NXNSAttack(74x)，但攻击者接收到的数据包将根据权威服务器的分布情况被均匀分配到 k 台权威服务器上。同时，D.NXNSAttack引发的数据包查询总量与C.NXNSAttack均达到了 b_{max} 。此外，若权威服务器启用了DNSSEC等安全设置，则会触发TCP重传机制，此时 C_v^{pkt} 可能会被放大至6倍。

(2) 针对递归解析器的攻击

攻击者发起针对“nxnsboss.com”的恶意子域名“sub.nxnsboss.com”的域名解析请求，导致递归解析器向顶级域名服务器发起递归查询，进而对递归解析器形成攻击。此时，递归解析器接收的数据请求包括攻击者控制的客户端向递归解析器发起的一个查询请求和攻击者权威服务器向递归解析器发起的 P_{max} 个查询请求。由式(2)可知， $P_{max} = 75$ 。此时，递归解析器需要处理的数据包总数为

$$C_v^{pkt} = \begin{cases} 2(P_{max} + 1)(1 + r), & TCP \\ 2(P_{max} + 1), & UDP \end{cases} \quad (4)$$

当.com域名服务器查询该子域名发现其不存在时，会返回NXDomain响应。此外，该响应包中还包含SOA记录、DNSSEC签名等额外信息^[5]，其长度远远超过UDP的最大响应大小(512 byte)。因此，后续数据传输将自动切换至TCP，进而触发TCP重传机制。此时， $C_v^{pkt} = 2(P_{max} + 1)(1 + r) = 912$ ， $C_a^{pkt} = 1 + k = 6$ 。由式(1)可知， $PAF = 152x$ 。文献[5]研究表明，C.NXNSAttack在攻击递归解析器时， $C_v^{pkt} = 3\,243$ ， $PAF = 1\,620x$ 。尽管此时D.NXNSAttack的攻击效益低于C.NXNSAttack，但仍能实现可观的放大效果。

实验结果表明，在修复了NXNSAttack漏洞的BIND软件中，D.NXNSAttack仍能对服务器造成显著的资源消耗。具体而言，随着每轮查询涉及的NS服务器数量 k 的增加，以及恶意NS服

务器层级的增加, D.NXNSAttack 产生的数据包总量也会相应增长。这一现象表明, D.NXNSAttack 具备更高的隐蔽性和灵活性, 同时也进一步暴露了当前 DNS 防护机制在 NS 递归查询关联性分析方面的不足。

3 NXNSAttack 流量识别模型

针对 NXNSAttack, 目前业界仍缺乏有效的流量识别与检测机制, 其技术挑战主要体现在以下 3 个方面。

1) 流量隐蔽性强: 攻击流量在协议格式、报文结构和查询模式等方面与正常 DNS 查询流量高度相似, 导致传统基于特征匹配的检测方法难以实现有效区分。

2) 行为复杂性高: NXNSAttack 涉及的递归查询和引荐响应过程具有较高的复杂度, 使得常规的检测方法难以准确识别攻击流量与合法流量之间的差异。

3) 攻击架构的对抗性: 攻击者采用分布式部署的权威服务器集群, 不仅增强了攻击路径的隐蔽性, 还通过构建多跳查询链显著提升了溯源分析的复杂度, 导致传统的 IP 黑名单等防御措施失效。

为有效应对 NXNSAttack 及其变种对 DNS 基础设施的威胁, 本文提出了一种基于时空特征融合的 DNS 攻击流量检测模型。该模型的创新性在于将流量的空间分布特征与时间序列特征进行深度融合, 通过以下 4 个核心模块实现攻击流量的精准检测。

1) 流量捕获模块: 负责实时采集 DNS 网络中的各类报文, 为后续检测提供原始数据支持。

2) 数据预处理模块: 对采集的流量数据进行清洗和特征提取等处理。

3) 时空特征学习模块: 采用深度学习算法挖掘流量数据的时空特征和攻击模式。

4) 攻击识别模块: 基于学习到的特征实现 C.NXNSAttack 流量、D.NXNSAttack 流量和正常流量的精准识别。

3.1 流量捕获

本文采用多源数据采集策略构建攻击学习数据集, 具体包含以下 3 个组成部分。

(1) 基于网络靶场的攻击流量捕获

本文搭建了基于 BIND 9.11.18 (未修复漏洞版本) 和 BIND 9.18.28 (修复漏洞版本) 的 DNS 服务器靶场环境, 用来模拟真实网络环境下的 C.NXNSAttack 和 D.NXNSAttack。在攻击实施过程中, 使用 LiBPCAP 工具捕获流量, 获取具有真实流量特征的攻击数据。

(2) 基于 Scapy 的模糊测试生成流量^[19]

为提高数据集的多样性和鲁棒性, 本文设计了基于 Scapy 的模糊测试方案, 具体实现过程如下。

1) 协议层变异: 利用 Scapy 强大的 DNS 报文构造能力对 RFC 标准定义的可扩展字段进行定向变异, 通过随机化查询类型 (如 A、NS、CNAME、AAAA)、动态设置非常规响应状态码 (NOERROR、NXDomain、SERVFAIL), 以及注入 Base64 编码或随机字符填充的异常载荷, 实现协议层变异。

2) 模拟复杂攻击场景: 运用随机算法构造 3~8 层嵌套 NS 记录引用链, 采用 IP 地址伪装技术将权威服务器地址替换为随机生成的 IPv4 地址组合, 并在 DNS 响应流中插入 1~500 ms 的随机时延扰动, 以模拟复杂网络环境。

3) 采用正交实验设计方法对查询类型、响应状态、权威服务器关联的恶意 NS 数量 (20~50 条) 及载荷混淆方式等维度特征进行全组合覆盖, 结合多线程速率控制机制 (50~1 000 packet/s 可调) 生成涵盖不同攻击频率 (突发型/持续型)、攻击深度 (浅层探测/多层递归) 和载荷复杂度 (标准报文/混淆数据) 的混合流量库, 最终生成包含 50 种 NXNSAttack 及其变种的数据包, 显著提高了流量的多样性和攻击场景的覆盖能力。

(3) 正常流量对比数据集

本文部署了专用递归解析器, 对 Cloudflare 提供的 10 万个二级域名 (second-level domain, SLD) 进行系统化解析。这些域名涵盖 A 记录、CNAME 记录、MX 记录等常见资源记录类型, 能够全面反映正常 DNS 查询的行为特征。使用 LiBPCAP 工具捕获完整的流量数据, 为模型训练提供了可靠的正常流量基准, 确保检测模型能够准确区分正常流量与异常流量。

3.2 数据预处理

为了将流量数据转换为适合深度学习模型训练的格式，本文采用图5所示的流程对流量数据进行预处理，具体步骤如下。

(1) 流量文件解析与流量提取

本文采用基于五元组（源IP地址、目的IP地址、源端口号、目的端口号、传输层协议）的流重组技术，对捕获的网络流量进行会话还原。考虑到NXNSAttack主要依赖UDP和TCP实施DNS查询放大，本文设计了一套严格的协议过滤机制。首先通过深度包检测（deep packet inspection, DPI）技术识别传输层协议类型，保留UDP/53和TCP/53端口的DNS流量，过滤ICMP、SCTP等非相关协议的数据包。该机制不仅可以确保数据集的纯净性，还能够有效提升攻击检测的针对性。

(2) 流量数据特征提取与表示

提取流量数据后，将其转化为适合深度学习模型处理的形式。首先，抽取TCP/IP头部的关键字段，包括源IP地址、目的IP地址、端口号、协议类型、请求与响应结果等信息。同时，将负载数据解析为字节序列，为模型提供更加丰富的上下文信息，从而提升对复杂攻击和异常行为的识别能力。

为防止神经网络直接依据源IP和目的IP地址进行分类而影响模型的泛化能力，本文模型对IP地址进行了模糊处理，具体如下。①将源IP和目的IP地址替换为固定值（如0.0.0.0），以屏蔽具体的地址信息。②保留IP地址头部的其他特征（如协议类型、TTL等），确保模型更专注

于流量的整体特征而非具体的地址标识，增强模型在不同场景下的泛化能力。③抽取数据包的时间戳、包间隔等时序特征，为后续的图像化处理提供更全面的基础数据。

(3) 二维数组的构建及灰度图像的生成

将每条数据流的字节序列填充或截断为1600 byte，然后重构为40×40的二维数组，并进一步生成40×40像素的灰度图像。实验表明^[20]，不同类型的流量通常在图像中呈现出独特的纹理和形态特征，从而为流量分析和模式识别提供重要依据。

3.3 时空特征融合学习

时空特征融合学习模型由特征嵌入、基于ShuffleNet^[21]的空间特征学习模块，以及基于Mamba的时间特征学习模块与基于PSFM的特征融合模块3个部分组成。特征嵌入模块部分负责将原始图像数据转化为高维度、更具表达力的特征向量。基于ShuffleNet^[21]的空间特征学习模块部分负责提取和优化模型的局部特征。基于Mamba的时间特征学习与基于PSFM的特征融合模块部分负责实现时间特征的高效学习与跨模态融合。为解决数据集样本数量失衡问题，模型采用带权重的交叉熵损失函数进行训练。

本文模型的框架如图6所示，特征嵌入部分采用PatchStride嵌入^[20]方法，通过对输入图像进行分块和步幅嵌入操作，将二维图像转化为一维特征序列。假设输入图像的尺寸为 $H \times W$ ，步幅为stride_size，则划分后的块数 N 可由式(5)计算得出。

$$N = \frac{H}{\text{stride_size}} \times \frac{W}{\text{stride_size}} \quad (5)$$

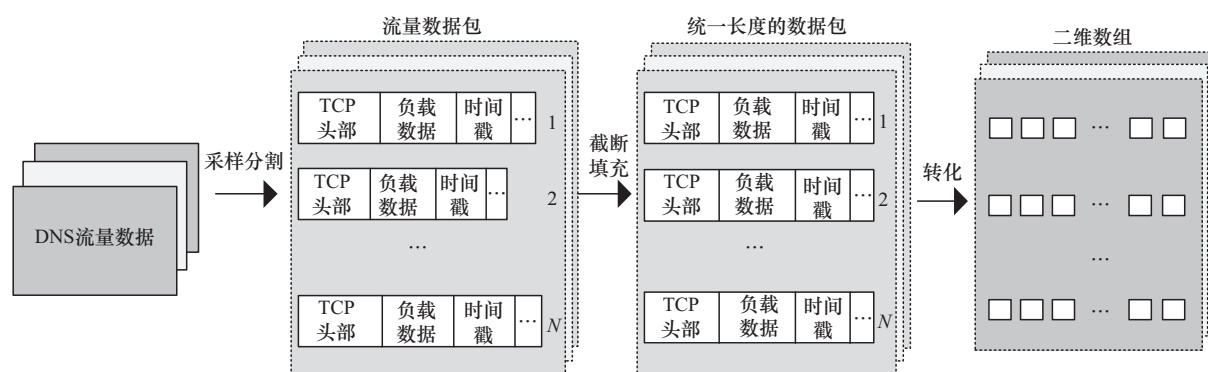


图5 流量处理流程

Figure 5 Traffic processing flow

为使模型能够捕获输入图像的空间位置信息, 本文在嵌入后的特征上添加了一个位置嵌入 pos_embed 。假设输入的嵌入特征的维度为 $(B, N, \text{embed_dim})$, 位置嵌入 pos_embed 的维度为 $(1, N+1, \text{embed_dim})$, 则位置嵌入可由式(6)表示。

$$x_{\text{embed}} = x_{\text{patch}} + \text{pos_embed}[:, N:, :] \quad (6)$$

经过上述处理, 原始图像数据被转化为高维度、更具表达能力的特征向量, 为后续模块的处理提供了有效输入。

基于 ShuffleNet 的空间特征学习模块部分接收从特征嵌入模块传输过来的特征向量, 并进一步提取复杂的局部特征。其采用轻量级 ShuffleNet 架构, 核心组件如下。

1) 动态分组的 1×1 卷积: 采用组间通道重排策略, 可减少 80% 的通道变换计算量, 并实现跨组特征的高效交互。

2) 改进的深度可分离卷积: 将 3×3 空间卷积与 1×1 通道变换解耦, 显著降低了计算复杂度, 使得在 40×40 分辨率下计算量仅为标准卷积的 $1/8$ 。

3) 具备跨层特征融合功能的残差连接^[21]: 确保梯度有效传播, 并减少内存占用空间。

此外, 该部分将 Relu 激活函数与 Layer Normalization 相结合, 确保了信息传递的稳定性。这种架构设计使该模块部分在中低分辨率场景下既能保持特征完整性, 又能优化计算效率。相较于 MobileNet^[22], ShuffleNet 具有两大优势: 一是通道重排机制提升了跨通道信息的交互效率; 二是深度卷积的受限使用有效避免了特征碎片化问题。

在基于 Mamba 的时间特征学习模块与基于 PSFM 的特征融合模块部分, 其中基于 Mamba 的

时间特征学习模块集成了单向 Mamba、RMS 归一化和残差连接, 既保证了特征建模的能力, 又在数值稳定性和计算效率方面进行了优化。单向 Mamba 在处理序列依赖任务时表现优异, 能够逐步提取输入特征并捕获序列中前后的依赖关系。RMS 归一化通过对输入特征进行均方根归一化, 避免数据分布不均导致的训练不稳定问题。基于 Mamba 的时间特征学习模块的每一层均采用 RMS 归一化, 以抑制梯度爆炸或梯度消失问题。残差连接将输入特征直接叠加在输出特征上, 确保每一层的关键信息得以保留。

基于 Mamba 的时间特征学习模块的核心计算公式如下:

$$X_t = \text{Linear}_x(\text{SSM}(A, B, C)(x') \odot \text{SiLU}(z)) + X_{t-1} \quad (7)$$

其中, $\text{SSM}(A, B, C)$ 表示单向 Mamba 的核心操作, 即采用硬件优化方式高效地处理序列数据。它结合了输入依赖的参数 (A, B, C) , 动态捕获序列中前后时间步的依赖关系, 增强对序列特征的建模能力。 x' 为输入特征 x 经过非线性激活函数 SiLU 和卷积操作处理后的中间特征表示。 z 为自门控机制中的特征表示, 用于动态控制 x' 的信息流动。 X_{t-1} 是残差连接的输入特征, 表示前一层的输出直接叠加在当前层的结果上。

在 DNS 流量分类任务中, 选择 Mamba 作为模型架构主要基于其在时序数据处理方面的独特优势。虽然 DNS 流量数据通常呈现较短的序列特征 (大多数样本长度不超过 2 000 byte), 但其蕴含的时序特性 (如数据包间隔、响应时延等) 对流量分类起着至关重要的作用。Mamba 基于选择性状态空间模型 (SSM) 构建, 其核心优势主要体现在以下 3 个方面。

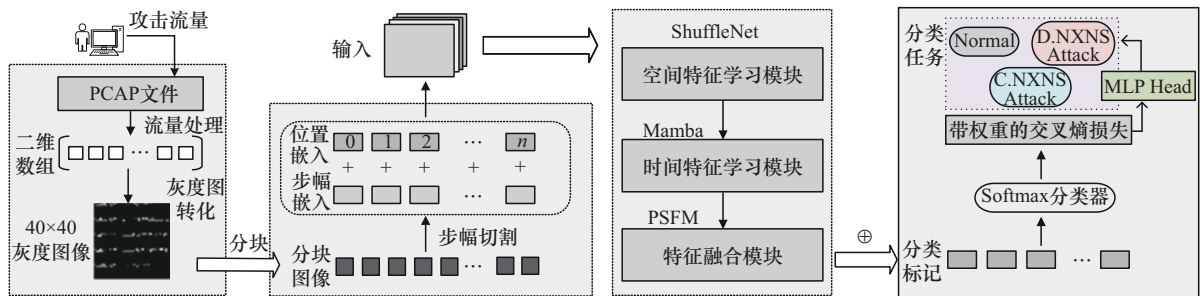


图 6 基于时空特征融合的 NXNSAttack 流量识别模型框架

Figure 6 The framework of the NXNSAttack traffic identification model based on spatio-temporal feature fusion

1) 长距离依赖建模能力：借助并行计算机制，Mamba能够在一次前向计算中完成全局信息的提取，进而高效捕捉数据中的长距离依赖关系。

2) 短序列优化：针对DNS数据的短序列特点，Mamba在建模过程中能够在保证特征捕捉精度的同时，有效提升处理效率，适配性极强。

3) 计算效率优势：Mamba拥有 $O(n)$ 的线性时间复杂度，能够充分利用现代硬件的并行计算能力，提高模型的训练和推理效率。

与传统的时序神经网络（如LSTM或GRU）相比，Mamba的状态空间建模方法从根源上避免了序列数据处理中的梯度消失或爆炸问题。实验结果也进一步验证，Mamba在处理DNS流量数据时，不仅能有效捕捉细微的时序特征，还能在计算效率和扩展性方面展现出显著优势，最终实现流量分类模型整体性能的提升。

基于PSFM的特征融合模块将经过空间特征学习模块处理得到的特征与单向Mamba处理后的特征进行融合，这种设计增强了模型对关键信息的提取能力，提高了其在不同任务上的泛化能力，进而提升模型的整体学习能力和对序列信息的捕捉能力。

在实际网络环境中，正常流量与异常流量的分布存在显著不平衡，正常流量占比远高于异常流量。为准确反映这一现实场景，本文构建的数据集中，正常流量的样本数量明显多于C.NXNSAttack和D.NXNSAttack这两类异常流量。针对这种类别不平衡问题，本文采用了带权重的交叉熵损失函数，通过为少数类别（即异常流量）分配更高的权重，使模型在训练过程中更加关注这种类别流量的学习，从而缓解其对多数类别（正常流量）的过度偏向。该损失函数的定义如下：

$$\zeta(x, y) = - \sum_{c=1}^C \omega_c \cdot y_c \cdot \log(\hat{y}_c) \quad (8)$$

为进一步缓解类别不平衡问题，本文采用数据欠采样策略在训练过程中动态调整数据分布，确保每个训练批次中各类别样本的比例相对均衡。在模型处理流程中，输入的流量特征经过深度特征提取与转换后，通过分类标记汇总获取全

局特征表示。随后，线性分类头将高维特征映射至类别空间，输出对应于3个类别的概率分布。最终，模型选择概率最高的类别作为预测结果。

4 实验结果及分析

4.1 实验环境

本文采用PyTorch 2.1.0深度学习框架，基于Python 3.10与CUDA 12.1环境提出的NXNS流量识别模型。所有实验均在Ubuntu 20.04操作系统环境下进行，硬件配置为Intel Xeon Silver 4214R（12 v CPU @ 2.40 GHz）、NVIDIA RTX 3080 Ti GPU（12 GB GDDR6X显存）。

首先模拟C.NXNSAttack环境。实验部署了4台服务器（1台作为客户端，1台作为递归解析器，1台作为攻击者的权威服务器，1台作为受害者的权威服务器）模拟C.NXNSAttack环境。在该网络结构下，捕获了100种不同NS记录数量的攻击样本，总计约8万个数据包。此外，利用Scapy进行模糊测试，获得了7万个更具多样性的攻击数据包。最终获得约15万个C.NXNSAttack数据包。

然后模拟D.NXNSAttack环境实验部署了9台服务器（其中，1台作为第1层攻击者的权威服务器，4台作为第2层攻击者的权威服务器，其他为辅助节点）模拟D.NXNSAttack。在该网络结构下，捕获约10万个攻击数据包。同时，通过Scapy模糊测试获得约10万个更深层次、拓扑结构更加复杂的D.NXNSAttack数据包。总计约20万个攻击数据包。此外，实验通过配置递归解析器对Cloudflare提供的10万个二级域名进行解析，最终获得约30万个正常网络流量数据包。

本文采用五折交叉验证方法对模型性能进行评估。将数据集划分为5个子集，在评估过程中轮流使用其中4个子集进行训练，剩余1个子集则用于验证，从而全面评估模型在不同数据子集上的性能表现。此外，本文还采用了多种策略提升模型的性能和稳定性。首先，利用带权重的交叉熵损失函数应对正常流量与攻击流量之间存在的样本不平衡问题。其次，类别权重根据训练集中样本数量动态计算，为少数类别分配更高的权重，从而提升模型对少数类别样本的识别能力。

为防止模型出现过拟合现象, 本实验实施了早停策略, 当验证集准确率连续 10 个 epoch 未出现提升时, 便提前终止训练。此外, 采用 ReduceLROnPlateau 调度器监控验证集的准确率, 当验证集性能陷入停滞状态时, 通过降低学习率来避免模型陷入局部最优, 进一步保障模型的学习效果。

4.2 评价指标

本文选用准确率 (Accuracy, ACC)、精确率 (Precision, PR)、召回率 (Recall, RC) 和 F1 分数 (F1-score) 4 个典型指标来评估模型的性能。

(1) 准确率

准确率表示模型预测正确的样本数占总样本数的比例, 其定义如下:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

(2) 精确率

精确率衡量的是模型预测的正类中有多少是实际的正类, 其定义如下:

$$PR = \frac{TP}{TP + FP} \quad (10)$$

(3) 召回率

召回率衡量的是全部实际为正类的样本中, 被模型正确预测为正类的比例, 其定义如下:

$$RC = \frac{TP}{TP + FN} \quad (11)$$

(4) F1 分数

F1 分数是精确率和召回率的调和平均, 能够综合反映模型的性能, 其定义如下:

$$F1\text{-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

其中, TP (true positive) 表示预测为正样本, 实际也为正样本的特征数, TN (true negative) 表示预测为负样本, 实际也为负样本的特征数, FP (false positive) 表示预测为正样本, 实际为负样本的特征数, FN (false negative) 表示预测为负样本, 实际为正样本的特征数。

4.3 实验结果

基于训练数据集完成模型训练后, 本文在独立测试集上评估了模型的分类性能。如表 1 所示, 本文模型在 3 类流量 (Normal、C.NXNSAttack、D.NXNSAttack) 的分类任务中展现出优异

的性能, 各类指标均达到较高水平, 尤其在处理 D.NXNSAttack 流量时, 表现接近完美。实验结果显示, 所有 (10 007 个) D.NXNSAttack 样本均被正确分类且无误判, 这充分证明模型对 D.NXNSAttack 样本具有极高的辨识能力。然而, 模型在识别 Normal 和 C.NXNSAttack 流量样本时存在一定程度的混淆。其中, 514 个 Normal 样本被误判为 C.NXNSAttack 样本, 12 个 C.NXNSAttack 样本被误判为 Normal 样本。

表 1 本文模型的流量识别性能
Table 1 Traffic identification performance
of our model

流量类型	ACC	PR	RC	F1-score
Total	98.95%	98.36%	99.39%	98.85%
Normal	98.29%	99.96%	98.29%	99.12%
C.NXNSAttack	99.88%	95.11%	99.88%	97.44%
D.NXNSAttack	100%	100%	100%	100%

进一步分析 Normal 和 C.NXNSAttack 流量的误报样本, 发现导致分类错误的主要原因有以下两点。

1) C.NXNSAttack 流量与正常流量在某些特征上存在一定程度的重叠 (某些特征值的分布较为接近), 使得模型在对这两类流量进行区分时容易产生混淆。

2) 收集的 C.NXNSAttack 数据集中包含高版本 BIND 的 C.NXNSAttack 流量。由于受到 NS_RR_LIMIT 的限制, 这些攻击流量的影响被显著减弱, 因此这部分 C.NXNSAttack 流量在流量包数量和请求频率等特征上, 与正常流量中的 NXDomain 响应流量包高度相似, 进而导致模型难以准确分辨。

如图 7 所示, 本文模型在 3 类 (D.NXNSAttack、Normal 和 C.NXNSAttack) 流量识别任务中展现出卓越的判别能力, 3 种流量的 ROC 曲线均紧贴坐标图左上角, 表明模型在不同分类阈值下均能保持极高的真阳性率和极低的假阳性率。此外, 3 条 ROC 曲线与随机分类的基准线 (对角线) 的距离均较远, 证实模型的分类效果远优于随机猜测。

如图 8 所示, 3 类流量的特征在二维空间中

呈现出较明显的聚类现象，这一现象说明模型能够有效地提取和区分不同类型流量的特征。具体而言，Normal流量（橙色）形成了一个较大的中心簇；D.NXNSAttack流量（蓝色）主要分布在图的上半部分，与其他类型的流量有着显著的分离性；C.NXNSAttack流量（绿色）主要集中在图的下半部分，虽然与Normal流量存在少量重叠，但整体上保持了良好的区分度。该分布结果与数据集本身的分布特征和各类流量的性质相符，这进一步验证了模型对特征提取的有效性。

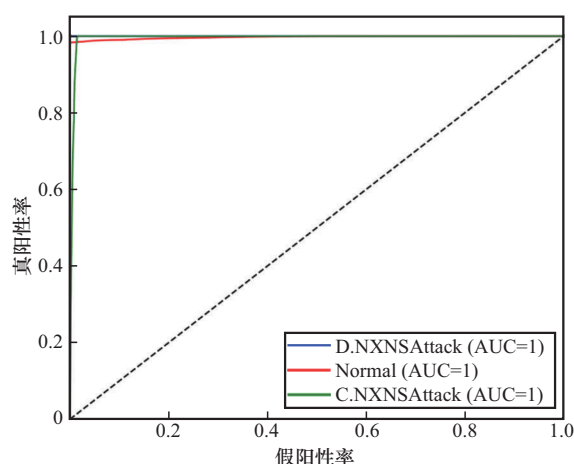


图7 模型分类ROC曲线

Figure 7 ROC curve of the model classification

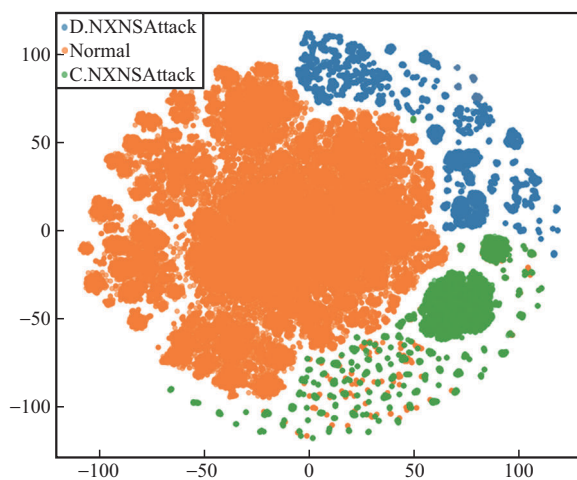
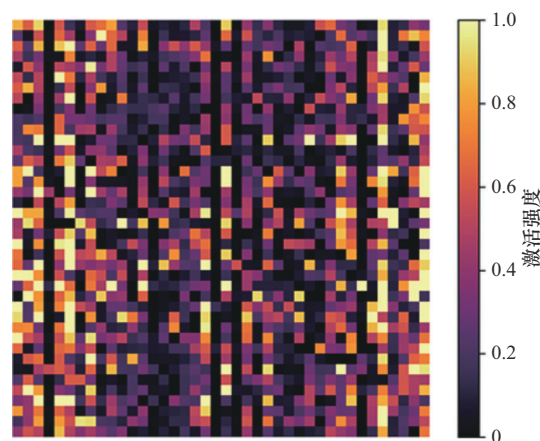


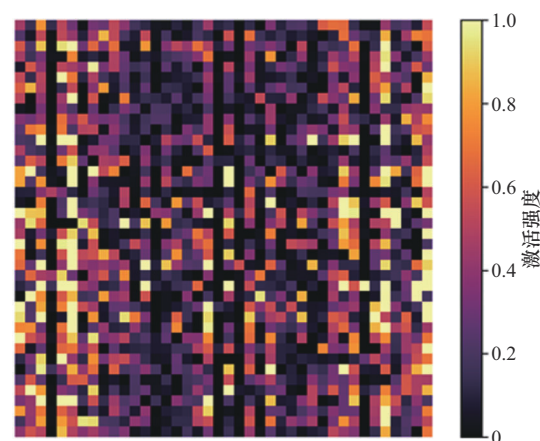
图8 特征向量分布

Figure 8 Distribution of eigenvectors

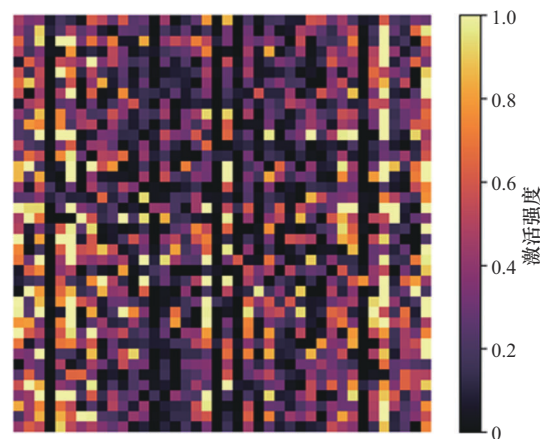
为了深入理解模型的决策机制，本文采用Grad-CAM技术对3类DNS流量的训练情况进行了可视化分析，结果如图9所示。



(a) 类别: D.NXNSAttack, 置信度: 99.97%



(b) 类别: C.NXNSAttack, 置信度: 92.35%



(c) 类别: Normal, 置信度: 99.95%

图9 3种DNS流量的Grad-CAM热力图

Figure 9 Grad-CAM heatmap of three types of DNS traffic

1) 对于D.NXNSAttack流量，模型主要关注具有明显垂直条带状分布的特征区域，这种模式反映了攻击流量中存在的周期性和规律性特征。模型对此类攻击的识别置信度高达99.97%，表明这些特征具有很强的判别性。

2) 对于 C.NXNSAttack 流量, 热力图呈现出相对分散但仍具有一定规律性的激活模式, 置信度为 92.35%。这种激活模式的差异性反映了集中式攻击与分布式攻击在时间分布特征上的区别。

3) 对于 Normal 流量, 热力图呈现出更加均匀分散的激活模式, 图中没有明显的规律性条带, 这与正常 DNS 查询的随机性特征相符。模型以 99.95% 的高置信度成功识别了这种特征模式。

为了系统评估模型架构中各模块对整体性能的贡献程度, 本文采用控制变量法设计了消融实验。所有实验在相同环境下重复 5 次, 并对结果取平均值和标准差。表 2 展示了各模块消融后的性能变化情况, 并提供了置信区间 (confidence interval, CI) 和配对 t 检验的统计学分析结果。具体而言, 均值差异表示模块消融后模型的准确率与完整模型的准确率之间的差值。95% 置信区间给出了均值差异的区间估计, 表明在 95% 的置信水平下, 真实均值差异有较大概率落入该区间。 t 统计量用于量化样本均值差异的显著程度, 其数值由样本均值差异和样本标准差共同决定。 t 统计量的值越大, 表明样本均值差异的显著性越强。 p 值用于评估假设检验的显著性, 表示在零假设成立的前提下观察到当前结果的概率。根据统计学惯例, 当 p 值小于显著性水平 (通常设定为 0.05) 时, 实验结果具有统计学意义, 即模块消融导致的性能差异是显著的。

如表 2 所示, 本文提出的完整模型在所有评估指标上均显著优于其消融变体。移除 ShuffleNet 模块导致模型的准确率下降 2.93% (95% 置信区间: [2.91%, 2.95%]), 移除 PSFM 模块导致模型的准确率下降 3.03% (95% 置信区间: [2.98%, 3.08%]), 移除 Mamba 模块同样引起了显著的性能衰退, 准确率下降 3.18% (95% 置信区

间: [3.09%, 3.27%])。所有消融实验均通过配对 t 检验, 结果显示出极高的统计显著性 ($p < 0.001$), 且 95% 置信区间均不包含零值, 进一步验证了各模块对模型性能的正向贡献。此外, 完整模型的准确率标准差仅为 0.01%, 表明该模型具有卓越的稳定性和可靠性。综上所述, 实验结果充分证明所提出的架构中每个模块均对模型的整体性能起到不可或缺的作用, 任何模块的缺失都会导致显著的性能下降。

为全面评估模型性能, 本文将所提方法与 MobileViT^[23]、ConvNeXt^[24] 及 NetMamba^[20] 3 种主流深度学习分类方法开展了对比实验。实验结果如表 3 所示, 本文方法在 4 项指标上均表现优异, 整体性能最佳。MobileViT 在多数指标上与本文方法表现接近, 但其精确率稍逊一筹。此外, MobileViT 基于 Transformer 架构的训练耗时显著多于本文采用的 Mamba 模型的训练耗时。ConvNeXt 的各项性能指标均落后于本文方法, 这种性能差距主要源于传统卷积神经网络在捕获 DNS 流量时空特征方面的局限性。NetMamba 作为 Mamba 架构的改进方案, 其性能显著优于传统卷积网络 ConvNeXt, 印证了状态空间模型在时空特征建模中的优势。然而, 其性能较本文方法仍存在小幅差距。

表 3 对比实验结果

Table 3 Comparison of experimental results

方法	F1-score	RC	PR	ACC
本文方法	98.85%	99.39%	98.36%	98.95%
MobileViT	97.76%	99.23%	96.46%	98.50%
ConvNeXt	96.80%	96.81%	96.97%	96.81%
NetMamba	98.15%	98.10%	98.19%	98.10%

上述实验结果表明, 本文模型在 3 类流量 (Normal、C.NXNSAttack、D.NXNSAttack) 分类任务中

表 2 消融实验结果及统计显著性分析

Table 2 Ablation experiment results and statistical significance analysis

模块	ACC	PR	RC	F1-score	均值差异	95% 置信区间	t 统计量	p 值
本文模型	98.96% $\pm$ 0.01%	98.4% $\pm$ 0.03%	99.4% $\pm$ 0.02%	98.9% $\pm$ 0.01%	—	—	—	—
Mamba+PSFM	96.05% $\pm$ 0.02%	94.5% $\pm$ 0.03%	97.7% $\pm$ 0.04%	95.9% $\pm$ 0.02%	2.93%	[2.91%, 2.95%]	147.32	<0.001
ShuffleNet+Mamba	95.77% $\pm$ 0.09%	94.1% $\pm$ 0.17%	97.5% $\pm$ 0.08%	95.5% $\pm$ 0.12%	3.03%	[2.98%, 3.08%]	134.76	<0.001
ShuffleNet+PSFM	95.96% $\pm$ 0.05%	94.4% $\pm$ 0.08%	97.7% $\pm$ 0.06%	95.7% $\pm$ 0.06%	3.18%	[3.09%, 3.27%]	112.45	<0.001

的表现都很出色，尤其是在识别D.NXNSAttack流量时，展现出极高的辨识能力。此外，ROC曲线、特征向量分布和Grad-CAM热力图的分析结果，从多个维度验证了该模型在分类性能上的优越性及特征提取环节的准确性。通过消融实验结合统计学分析可知，各个模块之间的协同作用对模型性能的稳步提升起到了至关重要的作用。而对比实验的结果表明，在要求高精度与高召回率的任务场景中，本文模型的分类性能和训练效率均更具优势。总体而言，该模型通过模块间的深度融合实现了性能的均衡提升，为复杂流量分类问题提供了可靠的解决方案。

5 结束语

本文深入探究了基于NXNSAttack的DDoS攻击原理，以及DNS协议中递归解析器处理NS记录时存在的安全性问题，提出了一种新型的分布式NXNSAttack DDoS攻击方式，并依托自建网络靶场对其攻击效益展开评估。结果显示，即便在新版本BIND针对NXNSAttack的补丁限制下，D.NXNSAttack仍能对受害者服务器发起高效的放大攻击，展现出极强的威胁性。

为实现对NXNSAttack流量的精准识别，本文设计并实现了基于时空特征融合的深度学习的流量检测模型。该模型有机整合了ShuffleNet的高效特征提取能力、Mamba框架的序列建模优势，以及PSFM模块的特征融合优势。实验结果证实，该模型对C.NXNSAttack和D.NXNSAttack两类恶意流量具有较高的分类性能，对这两类NXNSAttack恶意流量识别的F1分数均超过97%。

本文提出的基于时空融合的深度学习的NXNSAttack流量识别模型，主要以DNS流量为切入点，实现对恶意流量的精准识别。在未来的研究工作中，计划引入域名知识图谱，结合攻击模式的特征构建更加全面的预警体系；通过深度挖掘攻击行为与域名特征之间的关联，进一步实现更高效、更精准的网络安全防护。

参考文献：

- [1] MOCKAPETRIS P, DUNLAP K J. Development of the domain name system[C]//Symposium Proceedings on Communications Architectures and Protocols. New York: ACM, 1988: 123-133.
- [2] MIRKOVIC J, REIHER P. A taxonomy of DDoS attack and DDoS defense mechanisms[J]. ACM SIGCOMM Computer Communication Review, 2004, 34(2): 39-53.
- [3] ANTONAKAKIS M, APRIL T, BAILEY M, et al. Understanding the mirai botnet[C]//Proceedings of the 26th USENIX Security Symposium (USENIX Security 17). Berkeley: USENIX Association, 2017: 1093-1110.
- [4] GAO H Y, YEGNESWARAN V, JIANG J, et al. Reexamining DNS from a global recursive resolver perspective[J]. IEEE/ACM Transactions on Networking, 2016, 24(1): 43-57.
- [5] AFEK Y, BREMLER-BARR A, SHAFIR L. NXNSAttack: recursive DNS inefficiencies and vulnerabilities [C]//Proceedings of the 29th USENIX Security Symposium (USENIX Security 2020). Berkeley: USENIX Association, 2020: 631-648.
- [6] DAGON D, PROVOS N, LEE C, et al. Corrupted DNS resolution paths: the rise of a malicious resolution authority[C]//Proceedings of the Network and Distributed System Security Symposium. Virginia: the Internet Society, 2008.
- [7] TAYLOR V F, SPOLAOR R, CONTI M, et al. Robust smartphone app identification via encrypted network traffic analysis[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(1): 63-78.
- [8] HAYES J, DANEZIS G. K-fingerprinting: a robust scalable Website fingerprinting technique[C]//Proceedings of the 25th USENIX Security Symposium (USENIX Security 2016). Berkeley: USENIX Association, 2016: 1187-1203.
- [9] BIERBRAUER D A, DE LUCIA M J, REDDY K, et al. Transfer learning for raw network traffic detection[J]. Expert Systems with Applications, 2023, 211: 118641.
- [10] NIKTABE S, LASHKARI A H, SHARMA D P. Detection, characterization, and profiling DoH Malicious traffic using statistical pattern recognition[J]. International Journal of Information Security, 2024, 23(2): 1293-1316.
- [11] TU Y, LIU S, SUN Q. DNS tunnelling detection by fusing encoding feature and behavioral feature[J]. Computers & Security, 2023, 132: 103357.
- [12] LIU G N, JIN L, HAO S, et al. Dial "N" for NXDomain: the scale, origin, and security implications of DNS queries to non-existent domains[C]//Proceedings of the 2023 ACM on Internet Measurement Conference. New York: ACM, 2023: 198-212.
- [13] ZHENG W. The availability and security implications of glue in the domain name system[J/OL]. CoRR, abs/1605.01394, 2016.
- [14] MOURA G C M, CASTRO S, HEIDEMANN J, et al. TsuNAME: exploiting misconfiguration and vulnerability to DDoS DNS[C]//

Proceedings of the 21st ACM Internet Measurement Conference. New York: ACM, 2021: 398-418.

- [15] AFEK Y, BREMLER-BARR A, STAJNROD S. NRDelegationAttack: Complexity DDoS attack on DNS recursive resolvers[C]//Proceedings of the 32nd USENIX Security Symposium (USENIX Security 2023). Berkeley: USENIX Association, 2023: 3187-3204.
- [16] 史国振, 李昆阳, 刘瑶, 等. 基于深度残差胶囊网络与注意力机制的加密流量识别方法[J]. 网络与信息安全学报, 2023, 9(1): 32-41.
- SHI G Z, LI K Y, LIU Y, et al. Encrypted traffic identification method based on deep residual capsule network with attention mechanism[J]. Chinese Journal of Network and Information Security, 2023, 9(1): 32-41.
- [17] LI H Y, LI Z, ZHANG S Y, et al. Malicious DNS detection by combining improved transformer and CNN[J]. Scientific Reports, 2024, 14(1): 30248.
- [18] GU A, DAO T. Mamba: linear-time sequence modeling with selective state spaces[EB]. 2023.
- [19] S R R, R R, MOHARIR M, et al. SCAPY- A powerful interactive packet manipulation program[C]//Proceedings of the 2018 International Conference on Networking, Embedded and Wireless Systems (ICNEWS). Piscataway: IEEE Press, 2018: 1-5.
- [20] WANG T Z, XIE X H, WANG W D, et al. NetMamba: efficient network traffic classification via pre-training unidirectional mamba[J]. arXiv preprint arXiv:2405.11449, 2024.
- [21] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6848-6856.
- [22] SINHA D, EL-SHARKAWY M. Thin MobileNet: an enhanced MobileNet architecture[C]//Proceedings of the 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). Piscataway: IEEE Press, 2019: 0280-0285.
- [23] MEHTA S, RASTEGARI M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer[J]. arXiv preprint arXiv:2110.02178, 2021.
- [24] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 11966-11976.

[作者简介]



刘宇翔 (2001-), 男, 安徽亳州人, 天津理工大学硕士生, 主要研究方向为域名安全。



莫秀良 (1969-), 男, 天津人, 天津理工大学副教授, 主要研究方向为网络安全和信息安全。



宫良一 (1987-), 男, 山东烟台人, 中国科学院计算机网络信息中心高级工程师, 主要研究方向为互联网安全、物联网安全。



吕昆嫻 (2002-), 女, 吉林通化人, 中国科学院计算机网络信息中心硕士生, 主要研究方向为域名安全。



左鹏 (1985-), 男, 安徽池州人, 中国互联网络信息中心高级工程师, 主要研究方向为网络安全、互联网基础资源关键技术。