

基于 SSDP 和 DNS-SD 协议的双栈主机发现方法及其安全分析

施凡, 钟瑶, 薛鹏飞, 许成喜

(国防科技大学电子对抗学院, 安徽 合肥 230037)

摘要: 随着全球互联网 IPv4 地址分配耗尽, IPv6 开始加速推广和部署。双栈技术允许设备同时启用 IPv4 和 IPv6 栈, 这意味着用户暴露了双倍的安全风险。尽管现有的工作可实现对部分双栈服务器的识别和测量, 但仍存在以下问题。首先, 双栈主机识别需要对主机服务进行深层协议识别, 然而这种方式会消耗过高的扫描资源。其次, 实际的网络服务提供商有可能在分布式主机上提供一致的服务, 使得通过服务指纹进行双栈主机判别的准确性难以保证。针对此问题, 利用局域网服务发现协议将主机服务与 IP 地址绑定的协议特性, 提出了基于 SSDP 和 DNS-SD 协议的双栈主机发现方法: 在 IPv4 网络环境下, 通过 SSDP 诱导目标主机主动向构建的 IPv6 服务器发送请求, 然后从服务器日志中提取 IPv6 地址; 或通过 DNS-SD 协议枚举目标主机的服务列表及其对应的 AAAA 记录, 获取目标主机 IPv6 地址, 实现双栈地址对的发现。该方法直接从 IPv4 主机获取其 IPv6 地址, 确保了发现的双栈主机的准确性, 同时, 在发现过程中只需要针对特定协议构造请求数据包, 极大地节约了扫描资源。基于该方法, 对全球 IPv4 网络意外暴露的 SSDP 主机和 DNS-SD 主机进行了测量, 共收集到 158 000 个不重复的 IPv6 地址, 其中 55 000 个为拥有全球可达 IPv6 地址的双栈主机地址对。与现有工作将测量目标聚焦于双栈服务器不同, 该方法主要针对终端用户和客户端设备, 构建了迄今为止尚未探索过的活跃 IPv6 设备独特集合及双栈主机地址对集合。通过对获取的 IPv6 地址编址类型的分析, 随机生成已成为当前 IPv6 地址分配的主要方式, 这一方式大大降低了 IPv6 主机被扫描发现的可能性。特别地, 通过对双栈主机的开放端口和服务测量, 发现双栈主机在不同协议栈上的安全策略差异: IPv6 栈上暴露了更多高风险服务, 扩大了主机的攻击面。研究表明, IPv6 地址空间遍历扫描的不可行性缓解了 IPv6 的安全风险, 但错误的网络配置大幅增加了这些高风险的 IPv6 主机被发现的可能性, 用户应该重新审视双栈主机上的 IPv6 安全策略。

关键词: 双栈主机; 简单服务发现协议; DNS 服务发现; 网络测量

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2023003

Dual-stack host discovery method based on SSDP and DNS-SD protocol

SHI Fan, ZHONG Yao, XUE Pengfei, XU Chengxi

College of Electronic Engineering, National University of Defense Technology, Hefei 230037, China

收稿日期: 2022-05-27; **修回日期:** 2022-09-04

通信作者: 钟瑶, zhongyao@nudt.edu.cn

基金项目: 国家重点研发计划 (2021YFB3100500)

Foundation Item: The National Key R&D Program of China (2021YFB3100500)

引用格式: 施凡, 钟瑶, 薛鹏飞, 等. 基于 SSDP 和 DNS-SD 协议的双栈主机发现方法及其安全分析[J]. 网络与信息安全学报, 2023, 9(1): 56-66.

Citation Format: SHI F, ZHONG Y, XUE P F, et al. Dual-stack host discovery method based on SSDP and DNS-SD protocol[J]. Chinese Journal of Network and Information Security, 2023, 9(1): 56-66.

Abstract: With the exhaustion of the IPv4 addresses, the promotion and deployment of IPv6 has been accelerating. Dual-stack technology allows devices to enable both IPv4 and IPv6 protocols, which means that users are facing more security risks. Although the existing work can realize the identification and measurement of some dual-stack servers, the following problems still exist. Dual-stack host identification requires deep protocol identification of host services, but this method consumes too much scanning resources. Besides, network service providers may provide consistent services on distributed hosts, making it difficult to guarantee the accuracy of dual-stack host identification through service fingerprints. To solve these problems, the LAN service discovery protocol was used to bind host services to IP addresses, and a dual-stack host discovery method based on SSDP and DNS-SD protocols was proposed. In IPv4 network environment, the target host was induced to actively send a request to the constructed IPv6 server through SSDP protocol, and then the IPv6 address was extracted from the server's log. Or the service list of the target host and its corresponding AAAA record was enumerated through the DNS-SD protocol and the IPv6 address of the target host was obtained, in order to realize the discovery of the dual stack address pairs. With this method, IPv6 addresses was obtained directly from the IPv4 host, which ensured the accuracy of the discovered dual-stack host. At the same time, only request packets for specific protocols were needed during the discovery process, which greatly saved scanning resources. Based on this method, the SSDP hosts and DNS-SD hosts accidentally exposed to the global IPv4 network were measured. A total number of 158k unique IPv6 addresses were collected, of which 55k were dual-stack host address pairs with globally reachable IPv6 addresses. Unlike existing work that focused on dual-stack servers, this method mainly targeted end-users and client devices, and built a unique set of active IPv6 devices and dual-stack host address pairs that have not been explored so far. Through the analysis of the obtained IPv6 address addressing type, it shows that IPv6 address is mainly generated in a random manner, which greatly reduces the possibility of IPv6 hosts being discovered by scanning. In particular, by measuring the ports and services of dual-stack hosts, we found that the security policy differences of dual-stack hosts on different protocol stacks. Especially, IPv6 protocol stack exposes more high-risk services, expanding the attack surface of hosts. The research results also show that the infeasibility of IPv6 address space traversal scanning mitigates the security risks of IPv6, but incorrect network configuration greatly increases the possibility of these high-risk IPv6 hosts being discovered and users should revisit IPv6 security strategy on dual-stack hosts.

Keywords: dual-stack host, SSDP, DNS-SD, network measurement

0 引言

近年来,随着物联网和移动互联网的飞速发展,全球联网设备数量逐年高速增长,对 IP 地址的需求不断扩大,基于 IPv4 的全球互联网面临网络地址消耗殆尽、服务质量难以保证等制约性问题,IPv6 成为全球公认的下一代互联网商业应用解决方案。截至 2022 年 4 月,根据 APNIC 的统计数据,全球 IPv6 部署率已达到 30.72%^[1]。从服务端的角度看,越来越多的网站开始支持 IPv6。Alexa Top 1M 网站的 IPv6 采用率已达到 20.8%^[2],从用户端的角度看,谷歌已有 37% 的用户通过 IPv6 访问其服务^[3]。

IPv6 极大地增加了地址空间,互联网终端用

户可以从网络运营商处获取全球可达的 IPv6 地址前缀,并通过无状态配置方式获取接口标识 IID。IPv6 地址分配方式改变了传统 IPv4 空间借助 NAT 进行互联网通信的方式,提供了更好的端到端连接。但是对于互联网运营商和设备供应商,全面支持 IPv6 仍存在不小的技术障碍,IPv6 全面取代 IPv4 还需要经历较长的过渡阶段。目前,对 IPv6 的过渡仍处于 IPv4/IPv6 双栈改造阶段,大量 IPv6 主机实际处于双栈环境下。

根据 IETF 发布的 IPv6 安全指南^[4],从操作安全性的角度来看,双栈技术意味着用户暴露了双倍的风险,需要同时考虑保护这两种协议;从安全策略的角度来看,双栈网络的 IPv6 部分至少应与 IPv4 保持一致,以避免暴露更多风险服务,

随着时间的推移,保持安全策略之间的一致性具有挑战性。因此,对双栈主机的扫描与安全评估具有十分重要的价值。

目前的网络空间测绘^[5]工作主要关注 IPv4 空间,针对双栈主机的测量研究较少,且已有方法存在对双栈主机的误判和消耗扫描资源过多等问题。本文致力于使用尽可能少的扫描资源实现更加准确的双栈主机发现。

局域网服务发现协议将主机上开放的服务与 IP 地址绑定,网络的错误配置可能导致公网上的主机也能发现这些服务和地址。本文选择了简单服务发现协议(SSDP, simple service discovery protocol)和 DNS 服务发现(DNS-SD, DNS-based service discovery)协议这两种局域网服务发现协议,在 IPv4 网络环境下,通过 SSDP 诱导目标主机主动向构建的 IPv6 服务器发送请求,然后从服务器访问日志中提取 IPv6 地址;或通过 DNS-SD 协议枚举目标主机的服务列表及其对应的 AAAA 记录,获取目标主机的 IPv6 地址,从而构建双栈主机地址集,并对这些双栈主机的端口开放情况进行扫描,以初步分析主机在不同协议栈上表现出的服务开放情况和安全策略。本文共发现了 158 000 个不重复的 IPv6 地址,其中 55 000 个为拥有全球可达 IPv6 地址的双栈主机地址对,该地址集分布与传统的 FDNS、IPv6 Hitlist 等数据集的地址分布有较为明显的差异,更多地覆盖了终端用户和客户端设备,提供了一部分尚未被探索过的 IPv6 设备独特子集。同时,通过对获取的 IPv6 地址编址类型的分析,IPv6 接口标识 IID 的生成方式中,随机生成已占据较高比例,大大降低了 IPv6 主机被扫描发现的可能性。通过对这些双栈主机的开放端口和服务进行测量发现:大部分双栈主机在 IPv4 栈上构建了有效的安全防护策略,使得非授权用户无法访问主机的高风险服务,但这些策略没有覆盖 IPv6 栈,并暴露了 SMB、SSH 和 RDP 等服务,扩大了主机攻击面,使得主机面临更大的安全风险。

1 相关工作

IPv4 地址数量有限,可利用 Zmap 等工具进

行快速遍历式扫描,但 IPv6 地址空间巨大,难以使用传统的暴力扫描方法实现地址发现与探测。因此,对双栈主机的测量发现必须解决 IPv6 主机发现的问题。

针对此问题,学术界针对双栈主机发现的研究工作主要可分为两类:一类需要预先获取活跃的 IPv6 地址,构建潜在的双栈主机地址集,然后通过主动测量的方式获取主机信息,进行主机匹配;另一类则直接诱导 IPv4 主机通过其可能存在的 IPv6 网络发送请求,研究人员通过被动收集的方式获取双栈主机的信息。

1.1 双栈主机的主动测量发现方式

双栈主机的主动测量方式需要提前获取活跃的 IPv6 地址集合,近年来,许多研究人员提出了有效的 IPv6 地址预测和收集手段,包括通过 DNS 收集 IPv6 地址^[6-8],通过 traceroute 发现 IPv6 活跃地址及拓扑数据^[9-11],通过 P2P 协议进行 IPv6 对等实体发现^[12]和通过聚类算法进行 IPv6 地址预测生成^[13-15]等。

通过上述方式获取的 IPv6 地址中,有一部分采用过渡方案,这一类 IPv6 地址嵌入了 IPv4 地址或端口^[16],可直接用于双栈主机的发现和验证。

文献[17]和文献[18]通过 DNS 记录构建潜在的双栈主机列表,改进了之前通过 TCP 时间戳进行远程时钟偏差估计的工作,将其用于主动探测 IPv4/IPv6 主机并进行双栈主机判别,但并非所有的 TCP 实现都支持时间戳选项。

Czyz 等^[19]通过过滤 FDNS 数据集中同时具有 A 记录和 AAAA 记录的域名建立双栈主机的初始列表,然后对列表中主机的 HTTPS、SNMP、NTP、SSH、MySQL 等服务进行深层协议交互,通过综合比较多个协议抓取的指纹是否一致进行双栈主机判别。这一工作提供了有效的双栈主机判别方式,但实际的网络服务提供者可能在多个分布式主机上提供完全一致的服务,该方法无法保证判别的准确性,同时对多种协议进行深度扫描可能会耗费大量扫描资源。

1.2 双栈主机的被动测量发现方式

DNS 协议也被尝试用来进行双栈主机的被动测量,并成功发现了部分双栈 DNS 客户端。

Berger 等^[20]在来自 IPv4 的 DNS 查询响应中,

通过添加 1~2 条指向 IPv6 名称服务器的 NS 记录,在 DNS 解析器具有可用的 IPv6 网络时,收集可能存在的双栈地址对。但 IPv6 的 NS 记录只作为正常 DNS 服务的附加内容,DNS 解析器的选择顺序通常取决于其历史负载和性能特征,即使 IPv6 网络可用,也可能不会作为首选。

Al-Dalky 等^[21]诱导递归解析器向权威名称服务器发送 DNS 查询请求,通过被动收集 DNS 日志的方式发现成对的协作解析器,并认为它们是双栈主机。

这两类方法仅适用于 IPv4 网络上的 DNS 客户端或开放递归解析器,并需要 DNS 服务器后端基础设施,实现大规模测量的成本较高。

2 双栈主机发现技术

传统的双栈主机被动收集方式局限性较强,而主动测量方式基于 TCP 选项的一致性或多个端口的服务指纹一致性,受限于预先收集的 IPv6 地址集的质量和服务提供商在 IPv4/IPv6 栈上服务配置的一致性,存在准确率不高、资源消耗较大等不足。本文从不同的角度考虑这个问题,将扫描的重点转移到将双栈主机地址关联的局域网服务发现协议,并提出了一种新的双栈主机发现技术,可用于快速、准确的双栈主机发现。

2.1 局域网服务发现协议

一台连入局域网的终端设备往往具有与局域网中其他设备通信的需求,为实现局域网的网络邻居发现,不同操作系统提供了众多设备或服务发现协议,如 LLNMR、NetBIOS、AFP、SSDP 和 DNS-SD 等。随着互联网的发展,其中多数协议已经不再建议使用。

目前,SSDP 和 DNS-SD 协议已逐渐成为应用最广泛的基于 UDP 的内网主机和服务发现协议,基于这两种协议,提供服务的设备可以将自己的 IP 地址通过协议规定的多播方式发送到局域网中,被对该服务感兴趣的设备找到,这两种协议均可工作在 IPv4 栈和 IPv6 栈上。按照设计,它们应该只工作于本地网络,然而事实上,许多设备错误地将 SSDP 端口与 DNS-SD 端口暴露给

互联网。过去,这些暴露的设备通常被广泛用于 IPv4 网络中的 UDP 反射放大攻击^[22-23],通过类似的请求包,也可获取与服务绑定的可用的 IPv6 地址,实现对双栈主机的发现。

开启局域网服务发现协议的主机可以作为服务的使用方或提供方,并根据协议设计与局域网中其他的服务提供方或使用方进行通信。在利用局域网服务发现协议进行双栈主机发现的过程中,当确认目标主机开放相应的局域网服务发现协议后,可以向目标主机发布自己的服务或请求目标主机的服务。在 IPv4 网络环境中,当向目标主机发布自己服务的 IPv6 地址时,目标主机将使用 IPv6 网络访问设备,从而可以获取目标主机的 IPv6 地址;而当请求目标主机的服务时,则可以直接获取目标主机上与服务相关的 IP 地址信息。根据协议实现,与服务绑定的多个主机地址位于同一块网卡接口上,确保发现的 IPv6 地址与 IPv4 地址存在于同一主机。

2.2 基于 SSDP 的双栈主机发现

SSDP 是一种应用层协议,使用 HTTPU 和 HTTPMU 规范,基于固定的多播地址和 UDP 端口进行通信。SSDP 也是构成通用即插即用(UPnP, universal plug and play)协议的核心协议之一,用于实现 UPnP 设备之间的交互。基于 SSDP 的双栈主机发现,将目标主机作为服务的使用方,通过发布服务的方式诱导目标主机使用 IPv6 栈发送请求,从而获取双栈地址对。

2.2.1 SSDP 服务发现流程

SSDP 包含两种基本的消息类型,分别是查询请求和服务通知^[24],查询请求用于控制点(客户端)主动在网络中查询指定类型的设备和服务,服务通知则是由根设备(提供服务的设备)主动向网络中的所有设备声明自己的存在。SSDP 的服务发现过程基于这两种消息类型,可以由控制点主动请求或被动监听发现,如图 1 所示,第一组消息为根设备通过多播方式向网络中所有控制点发布服务,第二组消息为控制点通过多播方式寻找感兴趣的服务,提供相应服务的根设备直接通过单播方式响应控制点请求。通知消息和响应消息中都包含了一个指向设备详细信息的 URL。

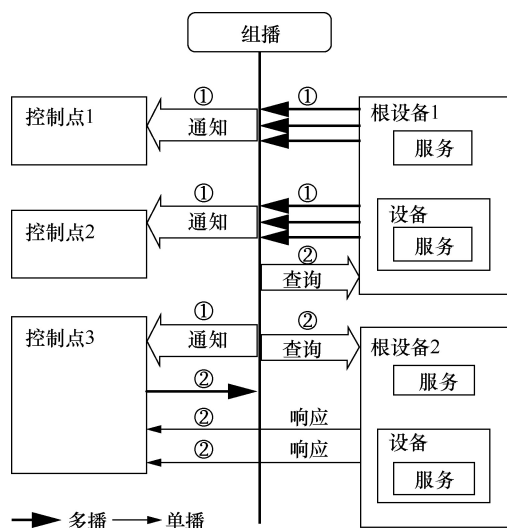


图1 SSDP 服务发现流程
Figure 1 SSDP service discovery process

2.2.2 UPnP 设备交互

在传统 IPv4 网络中, 私有网络的主机有被访问的需求时, 必须借助 NAT 网关设备进行端口映射, 将外部主机访问该端口的数据包转发给内网的设备, 从而实现内外网设备的通信。

UPnP 协议使支持 UPnP 的 NAT 设备可以自动为内网主机或服务分配端口并进行端口映射, 这一技术标准的出现可以使得家庭网络或小型企业网络在连接到互联网或不受管理的网络时更加灵活和易于使用^[25]。

UPnP 体系结构可以基于 SSDP 实现对来自各种厂商的各种设备类别的自动发现, 然后获取设备的描述信息并实现控制。控制点在通过 SSDP 发现根设备后, 往往需要了解更多关于设备功能的信息, 或者与设备交互, 此时, 控制点会访问发现消息中提供的 URL, 检索设备的 XML 描述, 如图 2 所示。最后, 控制点可以根据设备描述中提供的控制消息向设备的服务发送操作^[26]。

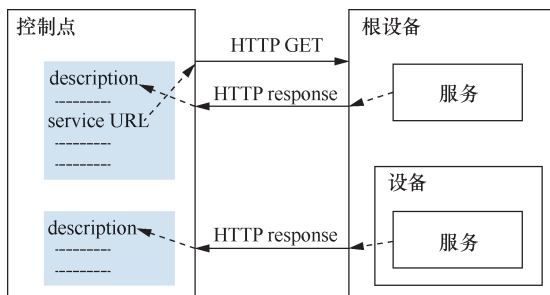


图2 控制点获取根设备描述信息
Figure 2 Control point obtains the root device description

2.2.3 基于 SSDP 的 IPv6 地址发现模型

按照 UPnP 体系的设计, 根设备提供的 LOCATION 字段中指向的 XML 文件应由根设备自身提供, 也就是 SSDP 服务通知数据包中指向 Web 服务的主机地址应与 SSDP 包本身的 IP 地址一致, 同时, UPnP 架构被设计用于私有网络的设备发现和控制, 为保证安全性, 根设备不应该直接连接到公网, Web 服务的主机地址应该是一个私有网络的 IP 地址。

在实际的设备发现过程中, 控制点既不校验 Web 服务的主机地址是否与宣称提供服务的设备的 IP 地址一致, 也不校验这个地址是否为私有网络以外的地址。当控制点设备收到一个指向公网地址的服务通知时, 它将会向公网地址发起数次 HTTP 请求, 黑客经常利用暴露 SSDP 端口的设备对某一公网设备发起 DDoS 攻击^[23]。如果控制点设备具有可用的 IPv6 网络, 当目标地址填入 IPv6 地址时, 它将向公网的 IPv6 服务器发送 HTTP 请求, 服务器的日志文件将会记录这次请求, 基于 SSDP 的地址发现模型如图 3 所示。

对于基于 SSDP 的双栈地址发现模型, 网络中至少包含 3 个设备, 用于地址发现的两个设备分别部署在 IPv4 和 IPv6 网络环境下, 而目标设备需要同时具备可用的 IPv4 和 IPv6 网络。针对活跃的 SSDP 目标主机, 通过工作在 IPv4 网络中的主机向其发送 NOTIFY 数据包, 数据包中 LOCATION 字段指向预先配置好的 IPv6 Web 服务器, 目标主机将通过同一网络接口上可用的 IPv6 网络发起 HTTP 请求, 即从 IPv6 Web 服务器的日志文件中提取访客 IPv6 地址。

为进一步从日志信息中找到原始的 IPv4 地址, 在构造的 LOCATION 字段中嵌入原始 IPv4 地址作为参数^[27], 即可在日志文件中的 GET 请求中提取 IPv4 地址, 从而获得一一对应的 IPv4/IPv6 地址对, 也就是双栈主机地址对。

2.3 基于 DNS-SD 的双栈主机发现

DNS-SD 基于多播域名系统 (mDNS, multicast DNS) 协议实现, 应用程序可以使用 DNS-SD 来查找网络服务并执行主机名解析。基于 DNS-SD 的双栈主机发现方法将目标主机作为服务的提供方, 根据协议设计, 请求目标主机上

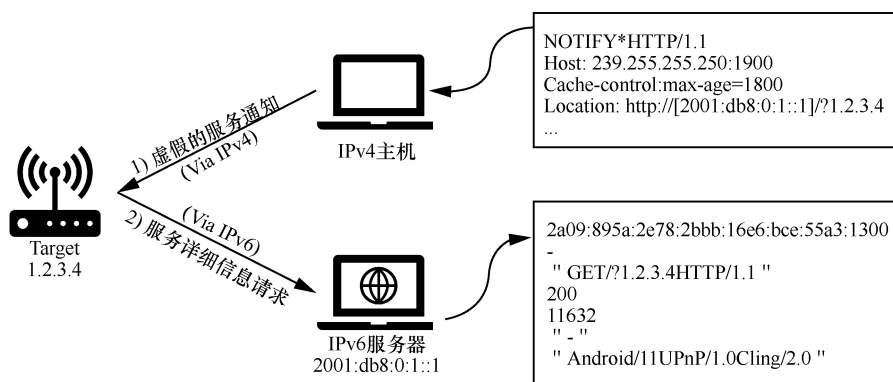


图 3 基于 SSDP 的地址发现模型
Figure 3 Address discovery models based on SSDP

的服务及相关的 IPv4/IPv6 地址。

2.3.1 基于 mDNS 的局域网主机发现

mDNS 遵循 DNS 协议，使用现有的 DNS 信息结构、语法和部分资源记录类型，实现了局域网内没有传统 DNS 服务器情况下的主机发现^[28]。

mDNS 协议允许局域网内的主机使用以 .local 结尾的域名，通常可以是“hostname.local”的形式，除非在局域网内产生名称冲突。当主机 HOST 被新添加到网络时，会向网络中的所有主机发送多播消息，消息包含 HOST.local 域名中可能存在的所有资源记录；当另一台新主机加入网络并尝试与 HOST.local 通信时，同样会向网络中的所有主机发送多播消息，询问 HOST.local 的 A 记录或 AAAA 记录，主机 HOST 根据 DNS 请求包中是否指定单播，进行单播或多播响应。

2.3.2 基于 DNS-SD 的局域网主机服务发现

mDNS 的实现中仅定义了局域网主机的域名解析，大多数情况下，应用程序对主机上提供的服务更感兴趣，因此，DNS-SD 在 mDNS 的基础上为主机提供的服务定义了可解析的域名形式。

当局域网中某个设备对某个类型的服务感兴趣时，首先发送一个查询 PTR 记录的数据包到多播组，查询的服务格式为 _service._ transport.local；接着对应服务的设备会响应一系列本设备上所具有的服务实例。主机获得 PTR 响应记录中的实例名称后，可以通过继续请求 SRV 和 TXT 记录来获取进一步的信息。其中，SRV 记录中有该服务对应的主机名和端口号，TXT 记录中有该服务的其他附加信息^[29]。

2.3.3 基于 DNS-SD 的 IPv6 地址发现模型

基于 DNS-SD 的 IPv6 地址发现过程中，客户端只关心特定的服务，但是对于网络管理员进行网络管理和问题诊断来说，获取主机上所有已发布服务类型的列表可能是有价值的。为此，DNS-SD 协议标准中定义了一个特殊的元查询，即名为 _services._dns-sd._udp.<DOMAIN> 的 PTR 查询。发送这一查询将得到一组 PTR 记录，即开放在该主机上的所有使用 DNS-SD 协议的服务记录，每个 PTR 记录的 rdata 是一个以 .<DOMAIN> 为结尾的双标签服务名称，如 _http._tcp.<DOMAIN>，代表一个运行在 TCP 上的 HTTP 服务。可以使用这些双标签服务类型构造后续的服务实例枚举 PTR 查询，以发现该服务类型的实例。

在局域网中，.<DOMAIN> 被定义为 .local，且 .local 作用域应该仅限于局域网中，DNS-SD 不应响应来自公网的 .local 请求。然而，协议本身并不对网络边界进行校验，当 UDP 的 5353 端口暴露在公网上时，它对任意向它发起查询的主机提供响应，此时，利用这一查询即可枚举主机的所有服务实例，并进一步获取其可能存在的 AAAA 记录，基于 DNS-SD 的 IPv6 地址发现模型如图 4 所示。

对于基于 DNS-SD 的双栈地址发现模型，用于地址发现的设备可以部署在纯 IPv4 网络环境中，目标主机启用了 IPv6，但不一定具有可用的 IPv6 网络。针对活跃的 DNS-SD 目标主机，首先发送用于服务列表枚举的 DNS 查询请求，根据响应的服务信息列表，进一步请求某一个具体服务实例的 PTR 记录，目标主机将响应服务实例的

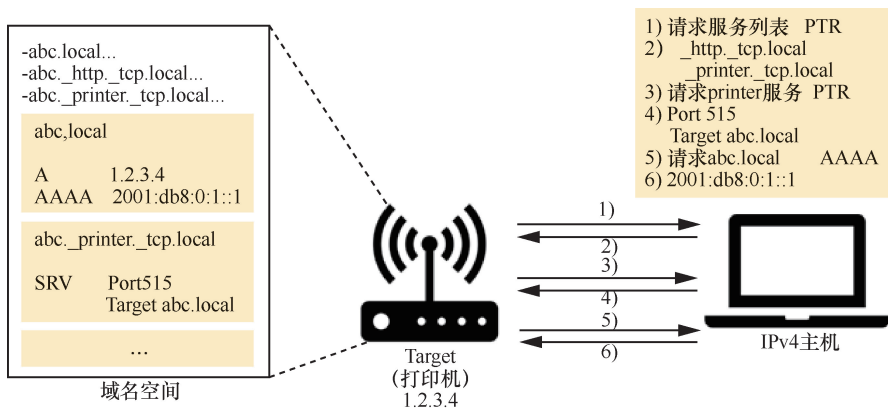


图 4 基于 DNS-SD 的 IPv6 地址发现模型
Figure 4 IPv6 address discovery model based on DNS-SD

SRV 记录, 包括服务实例的开放端口和绑定的主机在局域网内的域名, 继续查询该域名的 A 记录和 AAAA 记录, 即得到目标主机的 IPv4 地址和 IPv6 地址。当域名存在 AAAA 记录时, 目标主机为双栈主机。

3 双栈主机测量与分析

本节进行双栈主机发现与测量, 验证所提方法的有效性。本节利用 ZMap 对全球 IPv4 地址空间进行为期一周的扫描, 共发现了 1 320 000 个主机响应 SSDP M-SEARCH 数据包和 280 000 个主机响应 DNS-SD 查询请求, 在此基础上进行进一步的实验, 共收获了 55 000 个双栈主机。除此之外, 结合 IID 和应用层协议信息分析其安全特性。

3.1 扫描策略

双栈主机发现方式需要目标主机开启 SSDP 或 DNS-SD, 同时存在错误配置的网络边界, 使得这两类网络服务可以在公网访问。首先使用 ZMap 进行端口存活性识别, 向目标主机的 1900 端口和 5353 端口发送如图 5 所示的 UDP 探针包; 然后对进行响应的 SSDP 主机和 DNS-SD 主机, 分别发送伪造的 NOTIFY 数据包和 DNS 请求数据包, 并从 Web 服务器日志文件和 A/AAAA 记录对中获取潜在在开放的双栈主机; 最后分别在两个协议栈上执行 ICMP/ICMPv6 扫描和互联网 TOP 100 端口扫描。

在双栈主机发现过程中, 端口存活性识别只需要向目标主机发送单一的 UDP 数据包, 对于开放对应端口的主机, 后续的发现过程根据协议具

体实现, 需要再发送 1~3 个 UDP 数据包, 只消耗极少的扫描资源。

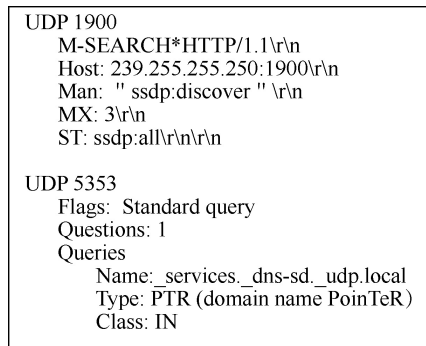


图 5 UDP 探针包
Figure 5 UDP probe packer

为避免对被测试主机的正常运行造成负担, 本文遵循网络测量研究的伦理准则^[30], 以尽可能减少潜在的风险。在测量中, 在源探测 IPv4 地址和 IPv6 Web 服务器上开放了网页, 以表明探测意图, 并显示了联系方式; 同时, 限制了探测速率, 以减轻探测对上游网络提供商和被测试主机带来的负面影响; 最后, 只进行了端口的无状态扫描, 避免了深层次的协议交互识别消耗目标主机过多资源。

3.2 IPv6 地址编址类型分析

IPv6 地址发现结果如表 1 所示, 通过对两类协议的扫描, 共提取到 158 000 个 IPv6 地址, 分布在 42 000 个不重复的/64 子网前缀下, 其中重复 IPv6 地址只占到 0.049%, 且重复的 IPv6 地址对应完全相同的 IPv4 地址, 在一定程度上验证了地址发现方法的准确性。其中, 通过 DNS-SD 查

表 1 IPv6 地址发现结果
Table 1 IPv6 address discovery results

协议	地区	地址数量	EUI-64	%	Random	%
SSDP	CN	9 041	701	7.75	8 034	88.86
	TH	4 220	182	4.31	4 036	95.64
	VN	3 958	3 782	95.55	175	4.42
	PK	3 196	78	2.44	3 116	97.50
	其他地区	7 223	5 048	69.89	934	12.93
DNS-SD	US	5 985	1 823	30.46	2 916	48.72
	FR	4 499	3 550	78.91	272	6.05
	DE	2 081	670	32.20	462	22.20
	CN	2 442	1 211	49.59	692	28.34
	其他地区	12 448	7 505	60.29	2 612	20.98
	Link-Scoped (私有地址)	100 111	83 662	83.57	15 868	15.85
	Unique Local (私有地址)	3 070	2 349	76.51	449	14.63
总计		158 274	110 561	69.85	39 566	25.00

询得到具有 AAAA 记录的主机中, 约有 81.6% 的主机网卡启用了 IPv6, 但不具备可用的公网 IPv6 地址, 仅存在唯一的本地单播地址和链路范围单播地址, 只能在局域网内进行访问。

利用 addr6 工具, 本节检查收集的地址中 IID 的分布, 分析了最具代表性的几类 IID 的数量和占比。

- ① EUI64: 嵌入主机网卡 MAC 地址;
- ② Low-byte: 高比特固定 (通常为 0), 只分配低比特位;
- ③ Embed-IPv4: 嵌入主机 IPv4 地址;
- ④ Byte-pattern: 具有某些固定模式;
- ⑤ Random: 遵循安全建议^[31]的随机 IID。

表 1 提取了两种发现方式下 IPv6 地址分布最多的 4 个国家, 分析了包括链路单播地址在内的所有 IPv6 地址的最典型的编址类型。表 2 和表 3 分别给出了两种协议发现的 IPv6 地址 IID 分析。

表 2 基于 SSDP 发现的 IPv6 地址 IID 分析
Table 2 IID analysis based on SSDP

类型	数量	IID 占比	类型	数量	IID 占比
EUI-64	9 793	35.43	Byte-pattern	58	0.21
Low-byte	605	2.19	Random	16 299	58.97
Embed-IPv4	883	3.20	Total	27 638	100.0

表 3 基于 DNS-SD 发现的 IPv6 地址 IID 分析
Table 3 IID analysis based on DNS-SD

类型	数量	IID 占比	类型	数量	IID 占比
EUI-64	14 817	53.97	Byte-pattern	323	1.18
Low-byte	4 215	15.35	Random	7 009	25.53
Embed-IPv4	1 091	3.97	Total	27 455	100.0

110 000 个(69.85%)的 IPv6 地址基于 EUI-64 格式生成, 这些地址暴露了主机的网卡地址, EUI-64 地址类型成为 IPv6 地址分配的最主要方式。当研究人员掌握运营商提供的/64 前缀池后, 这部分地址空间将被大幅压缩, 地址猜测变得十分容易, 设备的暴露风险也将随之增大。此外, 由 SLAAC 为终端主机设备生成的随机地址也是极具代表性的, 这类地址难以被预测和扫描, 保障了终端主机在互联网上不被攻击者发现。

3.3 双栈主机设备类型分布

日志文件和枚举的 DNS PTR 记录中有着丰富的字段信息, IPv6 地址本身也有潜在的语义信息, 充分挖掘这些信息有助于推断发现的双栈主机设备类型。

从一个 IPv4 主机可能发现多个活跃的 IPv6 地址, 这是因为在 IPv4 网络中, 多个设备通过 NAT 网络共用一个出口地址, 而在 IPv6 设备中, 它们拥有自己独特的地址, 且绝大多数分布在不同的/64

前缀下。进一步利用开源的 IPv6 别名前缀地址集进行验证,确保所获取的 IPv6 地址均分布在非别名前缀下^[32]。因此,使用这些独特的 IPv6 地址标识双栈主机,IPv6 地址总数即双栈主机总数。

3.3.1 双栈主机的网卡设备制造商

EUI-64 格式的 IID 通常由 48 bit 的硬件地址得到,即向 IEEE 指定的 24 bit 组织唯一标识符(OUI, organizationally unique identifier)和制造商为产品指定的 24 bit 中间嵌入 0xfffe 得到。

从 EUI-64 格式的地址提取 MAC 地址的 OUI 来识别地址设备制造商分布。通过 SSDP 发现的地址设备制造商相对集中,海康威视、华为和宇视科技的设备占设备总数的 89.8%,这与通过 SSDP 发现的 IPv6 地址主要分布在亚洲的结论一致。而通过 DNS-SD 发现的地址中,涉及的设备厂商众多,除群晖占到总设备的 49.6%外,其他设备分布相对均匀,证明将 DNS-SD 开放在公网上的错误普遍存在各类型用户终端设备的网络配置中。

3.3.2 双栈主机的 UPnP 客户端实现

大多数 UPnP 控制点在访问根设备描述信息时,使用 User-Agent 字段标明控制点的操作系统及版本、使用的 UPnP 协议版本和 UPnP 客户端实现方式及版本,只有不到 5%的设备隐藏了自己的设备信息。

在所有获取到的双栈主机中,有 24.6%的设备使用了多个不同的 User-Agent 字段发出 HTTP 请求,全部使用遵循 UPnP/1.1 协议的 MiniUPnPc 客户端,99.8%的 User-Agent 条目属于相同设备或相同制造商生产的相似设备,如表 4 所示,因此表现出相似的请求行为和相同的操作系统版本。

表 4 多客户端双栈主机的用户代理
Table 4 User-Agent for multi-client double-stack hosts

操作系统	版本	UPnP 客户端	版本	数量
Ubuntu	16.04	MiniUPnPc	1.9	7 046
Debian	8.11	MiniUPnPc	2.1	7 030
Debian	bullseye/sid	MiniUPnPc	2.1	4 355
Ubuntu	12.04	MiniUPnPc	2.1	3 582
Linux	3.10.0-957.el7.x86_64	MiniUPnPc	2.2.3	2 980

在剩余的使用固定客户端发出 HTTP 请求的设备中,有超过 97%的设备为嵌入式 Linux 设备,其中约 2%为安卓设备。除此之外,还有少量 Window 设备和苹果设备。

在 UPnP 的实现方面,95%的设备暴露了其客户端版本,包括 LibUPnP、MiniUPnPc、jUPnP、Cling 和 Java 多种实现方式。LibUPnP 为最主要的实现方式,其中超过 17 000 条使用了过时的客户端版本。根据 CVE 官方网站,发现双栈主机 LibUPnP 的多个过时版本中,存在 7 个已知漏洞,全部为高危漏洞,很容易被攻击者利用,作为进一步渗透内网的跳板。

3.4 双栈主机的端口开放情况

统计两种类型双栈主机在流行端口上的开放情况,总体来说,双栈主机在 IPv6 栈上开放了更多的端口和服务。基于 SSDP 发现的双栈主机中,通过 IPv6 扫描获得更多端口的主机占主机总数的 9.1%,而基于 DNS-SD 发现的主机中,这一比例则为 9.8%。更危险的是,在基于 SSDP 发现的双栈主机中,5.7%的主机屏蔽了所有来自 IPv4 的 TCP 连接请求,但是在 IPv6 栈上开放了不止一个服务;基于 DNS-SD 发现的主机中,这一比例较低,但仍然有 1.4%的主机表现出类似性质。

图 6、图 7 显示了双栈主机端口开放情况在 IPv4 和 IPv6 之间的差异,对于每个端口,展示了 IPv4 和 IPv6 的开放比例,仅 IPv4 或 IPv6 的开放比例,以及 IPv6 与 IPv4 的开放差异。

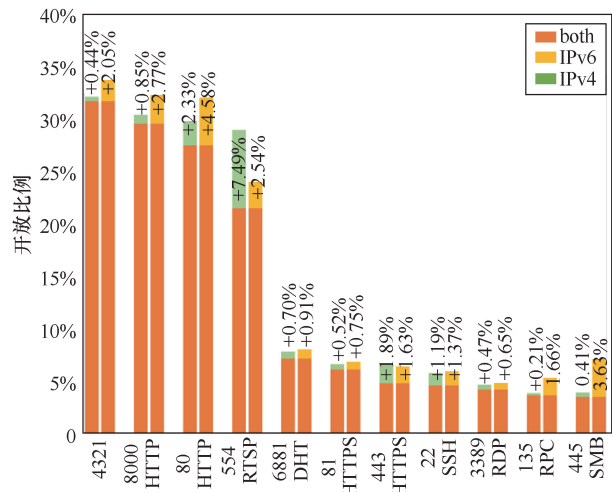


图 6 基于 SSDP 的双栈主机端口流行度
Figure 6 Percent of population open based on SSDP

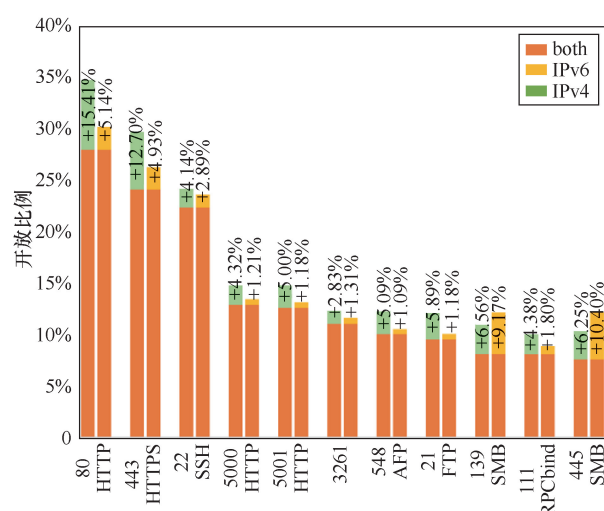


图7 基于 DNS-SD 的双栈主机端口流行度
Figure 7 Percent of population open based on DNS-SD

在基于 SSDP 的双栈主机中，开放比例最高的 11 个端口中有 9 个端口在 IPv6 上更开放。特别地，一些高价值的服务端口在 IPv6 中更开放，如 22、3389、135、445，这些端口在 IPv6 中的开放程度分别比 IPv4 高出 103%、104%、138% 和 183%。由于 SSH 和 RDP 允许远程控制设备，如果通过暴力密码尝试、收集管理员在其他受损网站或主机上使用的密码，将可能导致隐秘的大规模攻击，而 MS-RPC 和 SMB 协议则都包含多个公开的漏洞，极易受到攻击。

在基于 DNS-SD 发现的双栈主机中，11 个最开放的端口中有 2 个端口在 IPv6 上更开放，分别是 139、445 端口，这些端口在 IPv6 中的开放程度分别比 IPv4 高出 111%、118%。139 端口和 445 端口通常用于 Windows 上 SMB 协议的实现，是 Windows 系统上的高危端口。

4 结束语

本文基于 SSDP 和 DNS-SD 对 IPv4 主机进行了大规模的测量研究，以发现意外暴露的双栈主机和 IPv6 安全服务。这一双栈主机发现方式相较于之前的工作，使用了更少的扫描资源，且能够获得更加精准的双栈主机地址对。本文将发现的 IPv6 地址与互联网上公开的 IPv6 地址集进行比较，重合率不到 0.1%，这意味着虽然发现的双栈主机数据量较少，但它代表了迄今为止尚未探索的活动 IPv6 设备的独特集合。此外，双栈主机上

IPv6 地址普遍采用了 SLAAC 分配方式，保证了地址分布的随机性，使得主机的 IPv6 地址难以被发现，缓解了 IPv6 的安全风险。但是，意外暴露的双栈主机在实际中并没有保持在双栈上的安全策略一致性，面临潜在的安全威胁。随着 IPv6 技术的发展和 IPv6 主机数量的增加，这些暴露的双栈主机也会增加，意味着目前通过 NAT 技术隐藏在私有 IPv4 空间的配置或维护不良的设备极有可能被滥用，用户应该确保在双栈上维护相同的安全策略或防火墙配置，以提升设备的整体安全性。

参考文献：

- [1] APNIC. IPv6 capable rate by country (%) [EB].
- [2] W3Techs. Usage statistics of IPv6 for websites [EB].
- [3] Google. IPv6 adoption statistics [EB].
- [4] VYNCKE É, CHITTIMANENI K, KAELO M, et al. Operational security considerations for IPv6 networks[S]. IETF RFC 9099, 2021.
- [5] 赵帆, 罗向阳, 刘粉林. 网络空间测绘技术研究[J]. 网络与信息安全学报, 2016, 2(9): 1-11.
ZHAO F, LUO X Y, LIU F L. Research on cyberspace surveying and mapping technology[J]. Chinese Journal of Network and Information Security, 2016, 2(9): 1-11.
- [6] STEPHEN D S. Bootstrapping active IPv6 measurement with IPv4 and public DNS[J]. arXiv preprint arXiv:1710.08536, 2017.
- [7] FIEBIG T, BORGOLTE K, HAO S, et al. Something from nothing (there): collecting global IPv6 datasets from DNS[C]// International Conference on Passive and Active Network Measurement. 2017: 30-43.
- [8] FIEBIG T, BORGOLTE K, HAO S, et al. In rDNS we trust: revisiting a common data-source's reliability[C]//International Conference on Passive and Active Network Measurement. 2018: 131-145.
- [9] PLONKA D, BERGER A. Temporal and spatial classification of active IPv6 addresses[C]//Proceedings of the 2015 Internet Measurement Conference. 2015: 509-522.
- [10] GASSER O, SCHEITL Q, GEBHARD S, et al. Scanning the IPv6 internet: towards a comprehensive hitlist[J]. arXiv preprint arXiv:1607.05179, 2016.
- [11] BEVERLY R, DURAIRAJAN R, PLONKA D, et al. In the IP of the beholder: strategies for active IPv6 topology discovery[C]// Proceedings of the Internet Measurement Conference. 2018: 308-321.
- [12] MILESCU G, BARDAC M. Ipv6 peer availability in BitTorrent distributed hash table[C]//2013 19th International Conference on Control Systems and Computer Science. 2013: 532-536.
- [13] MURDOCK A, LI F, BRAMSEN P, et al. Target generation for internet-wide IPv6 scanning[C]//Proceedings of the 2017 Internet Measurement Conference. 2017: 242-253.
- [14] GASSER O, SCHEITL Q, FOREMSKI P, et al. Clusters in the expanse: understanding and unbiasing IPv6 hitlists[C]//Proceedings of the Internet Measurement Conference. 2018: 364-378.
- [15] SONG G, HE L, WANG Z, et al. Towards the construction of global

- IPv6 hitlist and efficient probing of IPv6 address space[C]//2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS). 2020: 1-10.
- [16] 张千里, 姜彩萍, 王继龙, 等. IPv6 地址结构标准化研究综述[J]. 计算机学报, 2019, 42(6): 1384-1405.
- ZHANG Q L, JIANG C P, WANG J L, et al. A survey on IPv6 address structure standardization research[J]. Chinese Journal of Computers, 2019, 42(6): 1384-1405.
- [17] BEVERLY R, BERGER A. Server siblings: Identifying shared IPv4/IPv6 infrastructure via active fingerprinting[C]//International Conference on Passive and Active Network Measurement. 2015: 149-161.
- [18] SCHEITL Q, GASSER O, ROUHI M, et al. Large-scale classification of IPv6-IPv4 siblings with variable clock skew[C]//2017 Network Traffic Measurement and Analysis Conference (TMA). 2017: 1-9.
- [19] CZYZ J, LUCKIE M, ALLMAN M, et al. Don't forget to lock the back door! a characterization of IPv6 network security policy[C]//Network and Distributed Systems Security (NDSS). 2016.
- [20] BERGER A, WEAVER N, BEVERLY R, et al. Internet name server IPv4 and IPv6 address relationships[C]//Proceedings of the 2013 Internet Measurement Conference. 2013: 91-104.
- [21] AL-DALKY R, SCHOMP K. Characterization of collaborative resolution in recursive DNS resolvers[C]//International Conference on Passive and Active Network Measurement. 2018: 146-157.
- [22] LIN H, DONG S K, BERGMANN N W. SAGA: secure auto-configurable gateway architecture for smart home[C]//2019 IEEE 24th Pacific Rim International Symposium on Dependable Computing (PRDC). 2019: 11-1109.
- [23] 韩虹莹. 基于 SSDP 的 DDoS 攻击防御技术的研究[D]. 长春: 吉林大学, 2019.
- HAN H Y. Research on DDoS attack defense technology based on SSDP[D]. Changchun: Jilin University, 2019.
- [24] GOLAND Y Y, CAI T, LEACH P. Simple service discovery protocol/1.0[S]. IETF DRAFT, 1999.
- [25] BOUCADAIR M, PENNO R, PENNO R, et al. Universal plug and play (UPnP) Internet gateway device-port control protocol interworking function (IGD-PCP IWF)[S]. IETF RFC 6970, 2013.
- [26] HAKIM M A, AKSU H, ULUAGAC A S, et al. U-pot: a honeypot framework for UPnP-based IoT devices[C]//2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC). 2018: 1-8.
- [27] MARTIN Z, ALEKSANDAR N. IPv6 unmasking via UPnP [EB].
- [28] CHESHIRE S, KROCHMAL M. Multicast DNS[S]. IETF RFC 6762, 2013.
- [29] CHESHIRE S, KROCHMAL M. DNS-based service discovery[S]. IETF RFC 6763, 2013.
- [30] BAILEY M, DITTRICH D, KENNEALLY E, et al. The menlo report[J]. IEEE Security & Privacy. 2012, 10(2):71-75.
- [31] GONT F, COOPER A, THALER D, et al. Recommendation on stable IPv6 interface[S]. IETF RFC 8064, 2017.
- [32] SONG G, YANG J, WANG Z, et al. DET: enabling efficient probing of IPv6 active addresses[J]. IEEE/ACM Transactions on Networking, 2022, 30(4):1629-1643.

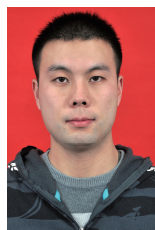
[作者简介]



施凡 (1983—), 男, 安徽合肥人, 国防科技大学副教授, 主要研究方向为网络空间测绘、网络安全知识图谱构建。



钟瑶 (1998—), 女, 山西忻州人, 国防科技大学硕士生, 主要研究方向为 IPv6 网络空间测绘。



薛鹏飞 (1989—), 男, 安徽合肥人, 国防科技大学讲师, 主要研究方向为网络空间测绘与知识图谱构建。



许成喜 (1989—), 男, 安徽合肥人, 国防科技大学讲师, 主要研究方向为网络空间测绘、互联网基础设施测绘、网络空间测绘数据分析。