CN 43-1258/TP
ISSN 1007-130X

计算机工程与科学
Computer Engineering & Science

第 47 卷第 8 期 2025 年 8 月
Vol. 47, No. 8, Aug. 2025

文章编号:1007-130X(2025)08-1391-08

大规模 IPv6 网络 IP-ID 类型测量*

黄峰元,杨轶帆,喻　波,杨振中,蔡志平,侯冰楠

(国防科技大学计算机学院,湖南 长沙 410073)

摘　要:IPv6 网络中,用于为网络层数据报提供分片和重组支持的 IP-ID 字段不再作为固定字段出现,而是被放入扩展头部中,以供灵活使用。近年来有利用 IPv6 分片机制引发 IPv6 目标主机生成 IP-ID,并进行别名前缀解析等的工作,说明在 IPv6 网络中 IP-ID 字段仍然存在信息泄露等问题,存在一定的安全风险。由于现有的 IP-ID 利用方法都是使用简单、可预测的 IP-ID 类型,因此探测互联网 IPv6 设备的 IP-ID 类型是否可预测,对 IPv6 网络安全和资产评估有重大意义。因此提出一种方法对互联网的 IPv6 设备进行探测,并且根据探测得到的结果对该设备生成 IP-ID 的方式进行分类。在得到的近 500 万个 IPv6 地址返回的 IP-ID 结果中,仍然有 41.1%的地址使用可预测的 IP-ID 类型。探测结果表明 IPv6 网络并非免疫于基于分片和 IP-ID 的攻击,IPv6 网络中仍然有相当多的设备使用存在高安全风险的可预测 IP-ID 类型。

关键词:IPv6 协议;IP-ID 字段;网络测量;网络安全

中图分类号:TP393.021　　**文献标志码**:A

doi:10.3969/j.issn.1007-130X.2025.08.006

A large-scale scan of IPv6 IP-ID

HUANG Fengyuan, YANG Yifan, YU Bo, YANG Zhenzhong, CAI Zhiping, HOU Bingnan

(College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China)

Abstract: In IPv6 networks, the Internet protocol identification (IP-ID) fields, which are used to support fragmentation and reassembly of network-layer datagrams, no longer appear as fixed fields but are instead placed in the extension header for flexible use. In recent years, researchers have exploited the IPv6 fragmentation mechanism to induce IPv6 target hosts to generate IP-IDs and perform tasks such as alias prefix resolution, demonstrating that the IP-ID field in IPv6 networks can still leak information and pose certain security risks. Since existing IP-ID exploitation methods rely on simple, predictable IP-ID types, probing whether the IP-ID types of IPv6 devices on the internet are predictable hold significant importance for IPv6 network security and asset assessment. This paper proposes a method to detect IPv6 devices on the Internet, and classifies them into different types. Among the nearly 5 million IPv6 addresses returned, 41.1% of the addresses still used predictable IP-ID, indicating that IPv6 networks are not immune to fragment and IP-ID based attacks. There are still a considerable number of devices in IPv6 network using predictable IP-ID which are of high security risk.

Key words: IPv6 protocol; IP identification (IP-ID) field; network measurement; network security

1　引言

IP-ID (Internet Protocol IDentification) 是 IPv4/v6(Internet Protocol version 4/version 6)报文头中的一个 16 bit/32 bit 字段,配合 offset 字段共同使用。其可以支撑网络层对超过物理链路所

* **收稿日期**:2024-10-18;**修回日期**:2024-11-01
基金项目:国家自然科学基金(62472434)
通信作者:蔡志平(zpcai@nudt.edu.cn)
通信地址:410073 湖南省长沙市国防科技大学计算机学院
Address: College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, Hunan, P. R. China

能承载的数据报进行分片,并且在目标主机处进行重组,以保证上层的过大数据成功传输[1,2]。在IPv4网络中,初始的网络层设计要求RFC(Request For Comments)791[1]以及后续的更新[3-5]均要求为不同三元组(即源地址、目的地址和上层协议)的数据报填充至少一个最大数据报生存时间内不重复的IP-ID值,未对IP-ID的生成方式做出明确规定,因此一些系统仅实现了简单的(如单调递增)IP-ID生成方式,导致设备发送数据报的IP-ID值存在一定相关性,从而造成设备隐藏信息泄露等问题。IPv4 IP-ID也一直被研究人员用于如端口扫描[6,7]、目标流量统计[8]以及别名解析[9-11]等任务,为网络设备的信息安全带来巨大威胁。

在IPv6网络中,从安全和性能等方面考虑,IP-ID字段作为一种扩展头部的形式存在,仅当需要对数据报进行分片时填充。并且路径最大传输单元MTU(Maximum Transmission Unit)发现[12]等机制也在尽力避免数据报分片的发生。在这些条件下,IPv6中IP-ID字段的出现频率远小于IPv4中IP-ID字段的。然而这并不代表IPv6设备可以免受基于IP-ID的攻击,相同的问题依然存在。2013~2017年,BEVERLY等人[13-16]利用第六版互联网控制信息协议ICMPv6(Internet Control Message Protocol version 6)中的Packet Too Big错误报文和过大报文引发IPv6目标主机对回应报文进行分片并且填充IP-ID值,并利用此机制对目标设备进行别名解析[13,14]和设备状态识别[15,16]的工作,说明可预测的IPv6 IP-ID也容易导致设备的信息泄露。

在IPv6设备面临上述威胁并且大多数攻击都利用可预测IP-ID生成方式的基础上,本文对互联网的IPv6设备进行大规模探测,旨在研究IPv6网络下设备的IP-ID生成行为以及发展趋势,做好网络的安全风险评估。本文的主要工作可以归纳为以下几点:

1)设计并实现了一种能够获取互联网IPv6设备的IP-ID生成结果,并且可以对不同IP-ID类型进行有效分类的探测方法。

2)将所提方法应用于900多万个IPv6地址,得到了大量互联网IPv6设备的IP-ID类型,对现今IPv6网络安全从IP-ID角度进行一定的了解分析。

3)为了推进IPv6网络安全的发展,将相应的代码和IP-ID分类结果进行公开,对应网址:https://github.com/Huangfengyuan/IP-ID/。

本文的探测与分类结果显示,尽管早在2011年RFC6274[17]中就已经推荐使用Random IP-ID类型,但是仍有9.2%和31.9%的地址使用相对可预测的Global IP-ID和Local IP-ID(各IP-ID类型特征描述见3.1)。通过对这些IPv6地址的地域分布进行分析,发现不同国家使用可预测IP-ID类型的比例相差较大,最高可达72%。本文还对不同操作系统的实现方式进行探究,发现Windows和Solaris系统仍在使用Local类型和Global类型。

2 相关工作

IPv4网络中,IP-ID作为固定字段存在,即使在没有分片需求的情况下也需要对其进行填充,因此研究人员通过直接发送数据包就可以获取目标设备当前IP-ID值并用于各种任务。而在IPv6中,直到2013年,BEVERLY等人[13]才提出TBT(Too Big Trick)方法来引导目标地址对回应报文进行分片。其分3步,首先发送1 300 Byte的ICMPv6 Echo Request报文,并且得到回应;然后发送ICMPv6 Too Big报文设置MTU为1 280并且填充第1步中回应报文的部分内容;最后再次发送1 300 Byte的ICMPv6 Echo Request报文,即可得到分片的回应结果。

2.1 IP-ID探测

2006年,RFC4413[18]将IPv4 IP-ID生成类型分为了3类,分别是Sequential jump、Random和Sequential(大多数文献也使用Global、Random和Local/per-destination代替),并且指出Sequential jump(Global)为最普遍使用的一类IP-ID生成类型。2005年的一项研究[8]中表明约有38%的主机使用Global的IP-ID生成方式。

2018年,SALUTARI等人[19]对公开的地址集的1 600万个IPv4地址进行探测,探测目标覆盖了所有24位前缀,并且根据响应报文的IP-ID序列将目标地址分为5类,分别是Constant,Global,Local,Random和Odd。在其258万个探测结果中,这些生成类型的占比分别为34%,18%,39%,2%和7%。说明即使信息泄露风险高的Global类型的使用在减少,但是最为安全的Random类型却依然没有在IPv4中得到广泛使用。BEVERLY等人[15]对2万多个路由器IPv6地址进行探测,得到的响应结果中约40%地址是Random类型,表明IPv4和IPv6中IP-ID使用和分布

有很大不同。由于其探测数量少、探测目标集中,因此需要对 IPv6 网络的 IP-ID 类型与行为特征进行全面且深入的探究。

2.2 IP-ID 的利用

可预测的 IP-ID 生成方式一直被广大研究人员使用以进行各项任务。其中 Global IP-ID 的信息泄露风险程度是最高的。攻击方式包括多种,首先是侧信道攻击,攻击者可以通过自身的 IP-ID 探测结果推测出目标主机在探测时间段内的与其他主机的通信情况。如 ANTIREZ[6] 于 1998 年提出空闲扫描(Idle Scan)的端口扫描方式,该方式控制 Global IP-ID 的僵尸主机,通过僵尸主机的 IP-ID 增长值来推测目标主机是否对僵尸主机发送回应报文,以得知目标主机端口是否开放,从而达到隐藏攻击者的目的。利用相似的原理,Global IP-ID 还被利用于对目标设备进行流量统计[8]、别名解析[8-10]、NAT(Network Address Translation)后主机数量统计[20]和源地址验证部署检测[21]等任务。其次是分片攻击,如域名系统 DNS(Domain Name System)缓存毒化攻击,攻击者可以发送与受害者缓冲区中 IP-ID 相同的 DNS 分片报文,当该分片报文被重组并记录时,就会污染受害者的 DNS 记录[22,23]。而在 IPv6 网络中,部分相似的攻击手段同样适用。如 BEVERLY 等人[13,14] 于 2013 年提出通过引导目标主机生成 IP-ID 的方式来获得目标地址的 IP-ID 值,并且同样利用其以进行路由器别名解析的工作。MORBITZER[7] 也同样提出在 IPv6 网络下的 Idle Scan。

除了 Global IP-ID 外,Local IP-ID 生成类型同样具有一定威胁。2014 年,Linux 3.16 版本内核[24]抛弃了 Global IP-ID 转而使用 Local IP-ID。其维护 2 048 位的计数器,将数据报的目的地址经过哈希以后选择其中一个计数器进行使用。ZHANG 等人[25]利用哈希碰撞的机制,找到与目标地址使用同一计数器的地址,以此达到 Global IP-ID 的效果并且同样用于端口扫描。其假设攻击者有 1 万个 IPv4 地址,计算出其中至少存在一个地址与任一目标地址发生碰撞的概率为 99.24%。后续也有诸多使用该碰撞机制的工作,如 ALEXANDER 等人[26] 利用返回的 RST(ReSeT the connection)报文中的 IP-ID 信息推测 2 个主机之间是否存在 TCP(Transmission Control Protocol)连接;FENG 等人[27]使用碰撞地址来推测目标地址当前 TCP 连接使用的序列号和确认号,并以此劫持目标地址的 TCP 连接。而在 2021 年,ERIC[28] 修改计数器的个数,由固定的 2 048 个变为根据内存进行调整,在 4 GiB 大小的内存下,计数器个数为 65 536 个,此时 1 万个 IPv4 地址与任一其他地址发生碰撞的概率约为 14.15%,大大降低了信息泄露风险。

包括但不限于 Global 和 Local 类型,一旦 IP-ID 序列表现出规律性,其就有可能被利用。BEVERLY 等人[15,16]通过周期性探测路由器 IPv6 地址并且根据其 IP-ID 序列变化周期来推测目标路由器的正常运行时间,并且根据得到的路由器中断情况来分析其对各自治系统 AS(Autonomous System)的影响。

3 探测与分类方法

为了准确得到目标地址的 IP-ID 类型,本节首先对不同的 IP-ID 序列特征进行分析,将不同 IP-ID 序列的行为表现分为 4 种。其次,由于 IPv6 默认不填充 IP-ID,本节对不同的引导方式进行对比并且选择了一个高效引发目标地址分片的方式。最后,本节设计了一种探测方法,以获取足以对不同 IP-ID 类型进行区分的 IP-ID 序列值并且使用决策树算法进行有效分类。

3.1 IP-ID 分类

与文献[19]对 IPv4 网络的探测结果分类相似,本文将 IPv6 设备的 IP-ID 行为类型分为图 1 所示的 4 类。其中 4 个散点图代表在主机 A、B 依次交替探测目标地址的情况下所可能得到的探测结果示例。从图 1 中可以看到,不同的 IP-ID 行为拥有不同的特征,分别是:

1)Global:该设备对所有目的地址维护相同计数器,每发一个包,计数器增加,增加幅度不定。因此主机 A、B 的 IP-ID 序列值交替增长,总体看上去是单调递增序列。

2)Local:该地址对于不同连接或不同目的维护不同计数器,每发一个包,计数器增加,增加的幅度不定。因此主机 A、B 的 IP-ID 序列值分别增长,彼此互不相关。

3)Random:使用随机数生成器生成 IP-ID,导致 IP-ID 值无规律可循,其也是最安全的类型。

4)Odd:极少数情况,IP-ID 序列怪异,统一划分为 Odd 类型。图 1d 中表示的是其中一种:主机在不同时间段对所有目的地址维护不同计数器,导致其在不同时间段内呈递增状态。还有其他实现方式或者由 bug、错误配置和乱序导致错误等情况。

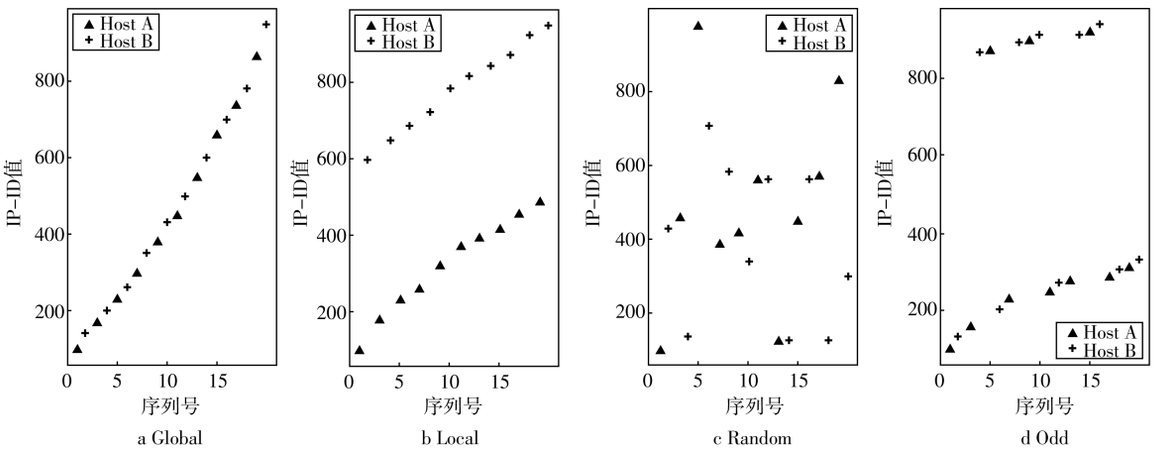

Figure 1 Examples of IP-ID behavior

图 1 IP-ID 行为示例

3.2 目标分片引导

IPv6 网络仅支持端节点对数据报进行分片，禁止中间路由设备的分片行为。如果中间路由器在转发报文时发现出口链路不能承载数据报的大小，该路由器会发送一个 ICMPv6 Too Big 消息，填写出口链路的 MTU 大小并发送给报文的源地址。源地址设备收到 ICMPv6 Too Big 消息后，会更新路径 MTU 大小，然后由上层协议重新分配数据块大小或者由网络层对数据报进行分片。

通过分析以上机制可以发现使用不同探测方式都可能引导目标地址进行报文分片。如直接发送过大的 ICMPv6 Request 报文；或者如 BEVERLY 等人[13]提出的 TBT 方法。但是，由于安全策略和系统配置等原因，过大的 ICMPv6 Echo 报文或者 ICMPv6 错误报文有可能被拦截并过滤，导致不同方式得到目标地址分片的成功率不同。为了验证不同引导方式的效率，本文取 GASSER 等人[29]对 IPv6 地址的活性探测结果的 20 万个地址进行对照试验，引导方法和探测结果如表 1 所示。其中 $MTU=x$ 代表发送设置 MTU 为 x 的 ICMPv6 Too Big 报文；$Pkt\_Size=y$ 代表发送大小为 y 的 ICMPv6 Request 报文。可以看到 TBT 方法能够引导最多的地址对返回报文进行切片，后续主动探测获取目标地址返回的 IP-ID 序列也基于此进行。

Table 1 Probing results of different inducing methods

表 1 不同引导方式得到的探测结果

%

分片方法	响应且分片	响应无分片	无响应
TBT 方法	27.4	36.3	36.2
$MTU=1\ 100$，$Pkt\_Size=1\ 200$	3.0	62.7	34.3
直接发送 $Pkt\_Size=1\ 300$	3.5	62.0	34.3
直接发送 $Pkt\_Size=1\ 600$	3.0	3.0	94.0

3.3 探测与分类方法

为了获取目标地址的 IP-ID 序列，并且对不同的 IP-ID 行为进行区分，本文设计了一种主动探测方法，其使用 2 个探测主机 A 和 B 依序向目标节点发送探针序列并且接收处理返回报文。其步骤如图 2 所示。

第 1 步，如图 2a，探测主机 A 以本机地址和伪造的主机 B 的地址为源地址分别向 X 发送一个 1 300 Byte 的 ICMPv6 请求，得到相应的回应报文。第 2 步，如图 2b，探测主机 A 以本机地址和伪造的主机 B 的地址为源地址分别向 X 发送一个 ICMPv6 Too Big 消息，其中设置 $MTU=1\ 280$，并且填充部分第 1 步中回应报文的数据。知道报文大小为

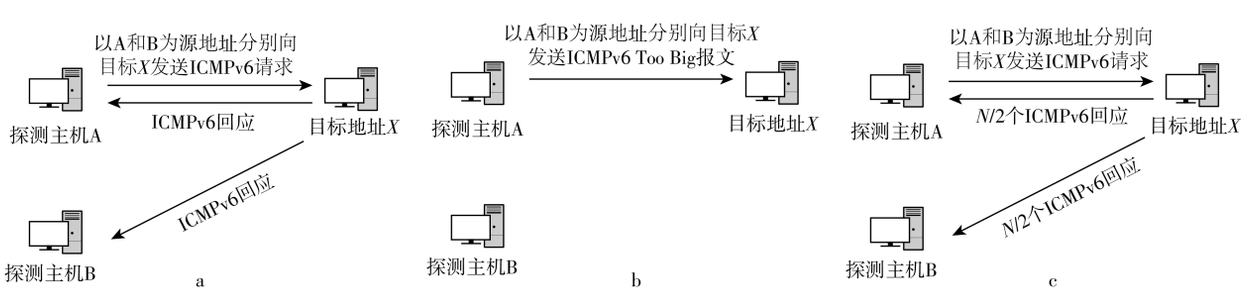

Figure 2 Probing method for acquiring IP-ID sequence

图 2 获取目标地址 IP-ID 序列的探测方法

1 280 Byte,这是因为 RFC 4443[30] 中表示 ICMPv6 错误报文内应该尽可能填充引起这次错误的报文的内容。第 3 步,如图 2c,探测主机 A 以本机地址和伪造的主机 B 的地址为源地址分别交替向 X 发送 $N/2$ 个 ICMPv6 请求,共得到 N 个回应的 IP-ID 序列 $S_T=\{x_1,x_2,x_3,\cdots,x_{(N-1)},x_N\}$,其中 A 处为 $S_A=\{x_1,x_3,\cdots,x_{(N-1)}\}$,B 处为 $S_B=\{x_2,x_4,\cdots,x_N\}$。$S'_T$、$S'_A$、$S'_B$ 中每个元素代表计算对应序列中的相邻元素的差值(如 $S'_T[i]=S_T[i+1]-S_T[i]$)。注意主机 B 应该发送的报文一直被主机 A 以伪造地址的形式实现,这是因为不同主机之间的同步过程过于繁杂,这种方式能够直接设置探针的发送顺序,更好地控制 A 和 B 发送报文的交替进行。并且 A、B 主机应该在同一子网下,以尽量减小被源地址验证机制过滤的可能性。

通过分析 A 和 B 处的 2 组 IP-ID 序列,即可得到目的地址 X 的 IP-ID 生成类型。本文使用决策树算法来创建并训练一个树形的决策过程,每个节点通过不同的特征信息进行判断,其使用的最优化分属性为信息熵,输入的特征为 S_T,S'_T,S_A,S'_A 的排列熵、一阶矩和二阶矩,分别用函数 $H(\cdot)$,$E(\cdot)$和 $\sigma(\cdot)$表示。

在模型的训练方面,本文从 4.1 节中的有效探测结果中人为挑选了 2 000 个 IP-ID 序列并且为其打上标签作为训练集,输入模型进行训练。然后随机选择了 8 000 个 IP-ID 序列,打上标签作为验证集,用于验证模型的效果,得到其准确率为 98.5%。验证得到的混淆矩阵如图 3 所示。

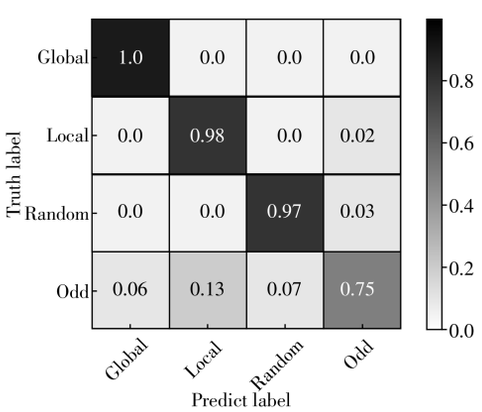

Figure 3 Confusion matrix of IP-ID validation

图 3 IP-ID 验证结果的混淆矩阵

观察图 3 可以发现,相比于其他 3 类,Odd 类的分类效果较差,因为此类中包含多种不同且样本数小的 IP-ID 生成类型,导致其 IP-ID 序列中包含的特征也不完全相同,并且这也导致了其他 3 类更容易被划分到 Odd 类中。图 4 展示了重要程度最高的 6 个特征,可以看到 S_T,S_A 的排列熵的重要程度明显更高。

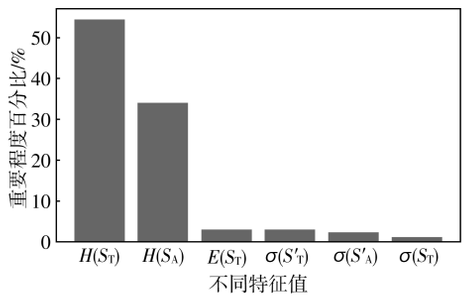

Figure 4 Importance of different features

图 4 不同特征的重要程度

4 探测结果与分析

4.1 主动探测与分类

本文将第 3 节所述的探测和分类方法应用于大规模的互联网 IPv6 设备探测并且对这些设备的 IP-ID 类型进行分析。选择同一子网内的不同主机 A、B 作为探针发送设备。对于每个探测地址,设置发送的探针数 N 为 30。不同于 IPv4 探测[19],本文需要设置 ICMPv6 Request 包大小超出 MTU 限制,大量的探针数会导致探测流数据过大,增加探测时间。

在探测目标选择方面,本文依赖于公开的 IPv6 地址探活数据集。对 Oliver Gasser 发布的 IPv6 Hitlists 数据集进行探测,其中包括 9 222 263 个 IPv6 地址[29]。在得到返回结果后,将未收到 ICMPv6 Reply 的地址分类为无响应;收到 ICMPv6 Replay 但分片的报文数小于 18(60%)的地址分类为响应无分片;分片报文数大于 18(60%)的地址分类为响应且分片,被视为有效的结果,进行下一步 IP-ID 分类。本文注意到响应无分片结果的形成原因多样,可能是高丢包率导致 ICMPv6 Too Big 消息或者大量 Request 报文丢包,也可能是安全策略阻挡导致 ICMPv6 Too Big 消息被阻拦。为了得到更加准确的结果,本文进行 3 轮探测,每轮将上一轮的响应无分片结果重新探测。

3 轮探测结束后,得到最终的探测结果。在总的探测结果中,有 21.2%的地址是无响应类型,有可能是该地址不存活或者 ICMPv6 Echo Request 报文被拦截;有 25.3%的地址是响应无分片结果,剩余的 53.5%则是有效的响应且分片结果。

本文将所有的响应且分片结果输入到决策树算法模型中,得到的结果如图 5 所示。在有效的 IP-ID 结果中,Global,Local,Random 和 Odd 的占

比分别是9.2%,31.9%,55.1%和3.8%。可以看到有规律增长的IP-ID类型(Global和Local)占比达到41.1%,表明相当一部分的IPv6设备仍然使用可预测的IP-ID类型,可能存在一定的安全风险。还有55.1%的设备使用了Random类型的IP-ID,相较于IPv4中该类型2%的占比得到了巨幅提升。

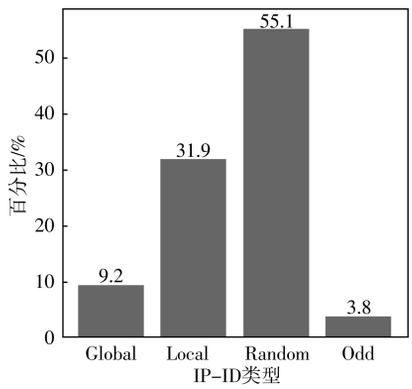

Figure 5 Classification results of IP-ID sequence

图5 IP-ID分类结果

本文对IP-ID序列的数值大小进行分析,发现在Global类型中有64.8%的地址其第1个IP-ID值小于10 000,16.5%的地址更是小于10,这表明这些主机很少甚至没有对IPv6数据包进行过分片操作,同时表明了IPv6网络中数据包分片和重组的操作并不频繁。而在Local和Random类型中,本文统计IP-ID值小于65 535(即$2^{16}-1$)的数量,发现其占比分别为6.0%和0.1%,可能是由于这些地址直接使用IPv4的IP-ID值填充IPv6的IP-ID字段导致的。

4.2 地理位置和操作系统分析

为了进一步探究IP-ID类型分布规律和发展现状,本文对MaxMind IPv6地理位置数据库[31]进行查询,根据结果对IPv6 Hitlist数据集中IPv6地址最多的前10个国家进行取样,计算这些国家IPv6地址中Global和Local类型的占比,得到的结果如表2所示。可以看到部分国家使用可预测的IP-ID占比很少,最少的法国只有6.5%的比例,而占比最高的日本甚至有72%的IPv6地址使用Global或者Local类型。

本文同时也通过阅读源码或主动探测的方式获取一些流行的和较新的操作系统版本的IP-ID类型,结果如表3所示。可以看到除了Windows家族和Solaris以外,大部分操作系统都使用安全的Random方式。结合二者,本文推测可能是不同的操作系统使用习惯和版本更新习惯的差异导致不同国家IPv6设备IP-ID类型使用的不同。

Table 2 Percentages of Global and Local IP-ID in different countries

表2 不同国家探活地址中Global和Local类型占比 %

国家	百分比
法国	6.5
德国	9.1
美国	22.0
中国	25.4
巴西	40.0
日本	72.0
荷兰	15.0
英国	15.6
瑞士	62.6
新加坡	15.2

Table 3 IP-ID types of different OSs

表3 不同操作系统的IP-ID类型

操作系统版本	IP-ID类型
Linux 4.15.0	Random
Linux 6.5.0	Random
Windows 11 64 bit	Local,每包增长2
Windows 10 64 bit	Local,每包增长2
Windows Server 2022	Local,每包增长2
OpenBSD 7.2	Random
FreeBSD 13.1	Random
NetBSD 9.3	Random
Solaris 11.4	Global,每包增长1

4.3 IPv4与IPv6的IP-ID类型对比

IPv4网络中,仅有2%的地址使用Random IP-ID,远远低于IPv6网络,说明IPv4网络和IPv6网络之间具有一定差距。以Linux为例,IPv6 IP-ID直接调用随机数生成器生成;而IPv4 IP-ID仍是使用Local IP-ID,其存在哈希冲突的问题,即使在增加哈希表大小后[28]仍然存在一定碰撞的可能性。

本文猜测IPv4设备更少使用Random IP-ID原因可能有2点:1)IPv4中有Constant类型可以为IPv4设备提供保护,其在数据包不分片时为IP-ID字段填充一个常数,并且该类型实现相较于Random类型更为简单。但是本文认为Constant类型相较于Random类型是不可靠的,因为攻击者同样可以通过分片诱导的方式使目标主机必须生成唯一的IP-ID。2)性能考虑。由于IPv4网络需要对每个数据报的IP-ID都填充一个值,在高流量负载的情况下,使用随机数生成器对IP-ID进行赋值导致更高的计算成本,从而造成性能下降。

而在 IPv6 中,没有以上 2 点问题。首先,IPv6 IP-ID 不可能使用 Constant 类型;其次,由于中间路由器不允许对报文进行分片,IPv6 分片的情况较 IPv4 更少。而在本文的探测结果中,有 64.76%的 Global 类型的地址,其 IP-ID 值平均数小于 1 万,说明这些设备很少发送分片报文,从而使计数器的值较小,这也从侧面说明 IPv6 网络中分片事件发生较少。从使用角度看,RFC 8900[17] 列举了几种可能依赖 IPv6 分片机制的应用,如一些用户数据报协议 UDP(User Datagram Protocol)应用,开放式最短路径优先 OSPF(Open Shortest Path First)协议和隧道封装等。更多应用则无需依赖分片机制,如传输控制协议 TCP(Transmission Control Protocol)可以通过调整数据块大小来避免分片的发生。

因此综合性能和安全 2 个角度,IPv6 网络的 Random IP-ID 较其他类型更加有优势。在对性能方面的考虑下,使用 Random IP-ID 可以大大减少 IP-ID 被预测的可能性,从而保护 IPv6 用户的安全。

5 结束语

本文实现了一种能够获取互联网 IPv6 设备的 IP-ID 生成结果并且对不同 IP-ID 类型进行有效分类的探测方法并进行了互联网环境下的 IP-ID 探测,是首个对互联网 IPv6 设备的 IP-ID 类型进行系统探测、分类的研究。探测结果表明,在 IP-ID 被滥用的情况下,已经有 55.1%的 IPv6 地址使用安全的 Random IP-ID,但还是有 41.1%的地址使用可预测的 IP-ID 类型,尤其是信息泄露风险最高的 Global IP-ID 占比 9.2%,说明 IP-ID 攻击对 IPv6 网络仍然存在威胁。进一步的研究表明一些主流的操作系统较新版本大多使用 Random IP-ID。使用更安全的操作系统和及时更新版本能够很好地维护网络安全,因此操作系统应该使用更加不可预测的 IP-ID 类型并且用户也应该及时更新操作系统版本以保证信息安全。

在本文研究的基础上,未来的工作计划从以下 3 个方面展开。首先是优化探测算法,由于 ICMP 速率限制等因素,在短时间向同一个网络发送流量可能导致丢包的概率增大,因此如何使发送的流量变得均匀是需要考虑的一个问题。其次是探测范围,即增大探测目标的丰富度,对如路由器地址、服务器地址等不同类设备进行不同的特征研究。最后是 IPv6 IP-ID 的利用,从已有的文献中可以看到 IPv6 IP-ID 的利用情况相较于 IPv4 仍然较少。尽管由于 IPv6 IP-ID 不再成为固定字段,一些如目标流量统计类的工作很难在 IPv6 设备上进行,但是一些任务仍有实施的可能性。在进一步的工作中将探究 IPv6 IP-ID 的可利用性,以判断可预测的 IP-ID 类型对 IPv6 网络的威胁程度。

参考文献:

[1] POSTEL J. Internet protocol: RFC 791[S]. Arlington: Defense Advanced Research Projects Agency, 1981:9.

[2] DEERING S, HINDEN R. Internet protocol, version 6 (IPv6) specification: RFC 2460[S]. Reston: Internet Engineering Task Force, 1998:12.

[3] ALMQUIST P. Type of service in the internet protocol suite: RFC 1349[S]. Reston: Internet Engineering Task Force, 1992:7.

[4] NICHOLS K, BLAKE S, BAKER F, et al. Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers: RFC 2474[S]. Reston: Internet Engineering Task Force, 1998:12.

[5] TOUCH J. Updated specification of the IPv4 ID field: RFC 6864[S]. Reston: Internet Engineering Task Force, 2013:2.

[6] New TCP scan method[EB/OL]. (1998-12-18)[2024-10-01]. https://seclists.org/bugtraq/1998/Dec/79.

[7] MORBITZER M. TCP idle scans in IPv6 [D]. Nijmegen: Radbound University, 2013.

[8] CHEN W F, HUANG Y, RIBEIRO B F, et al. Exploiting the IPID field to infer network path and end-system characteristics[C]//Proceedings of the 6th International Workshop on Passive and Active Network Measurement, 2005:108-120.

[9] SPRING N, MAHAJAN R, WETHERALL D, et al. Measuring ISP topologies with Rocketfuel[J]. IEEE/ACM Transactions on Networking, 2004, 12(1):2-16.

[10] BENDER A, SHERWOOD R, SPRING N. Fixing ally's growing pains with velocity modeling[C]// Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, 2008:337-342.

[11] KEYS K, HYUN Y, LUCKIE M, et al. Internet-scale IPv4 alias resolution with MIDAR[J]. IEEE/ACM Transactions on Networking, 2013, 21(2): 383-399.

[12] MCCANN J, DEERING S, MOGUL J, et al. Path MTU discovery for IP version 6: RFC 8201[S]. Reston: Internet Engineering Task Force, 2017:7.

[13] BEVERLY R, BRINKMEYER W, LUCKIE M, et al. IPv6 alias resolution via induced fragmentation

[C]//Proceedings of the 14th International Conference on Passive and Active Measurement,2013:155-165.

[14] LUCKIE M, BEVERLY R, BRINKMEYER W, et al. Speedtrap:Internet-scale IPv6 alias resolution[C]//Proceedings of the 2013 Conference on Internet Measurement Conference,2013:119-126.

[15] BEVERLY R,LUCKIE M,MOSLEY L,et al. Measuring and characterizing IPv6 router availability[C]//Proceedings of the 16th International Conference on Passive and Active Measurement,2015:123-135.

[16] LUCKIE M, BEVERLY R. The impact of router outages on the AS-level Internet[C]//Proceedings of the Conference of the 2017 ACM Special Interest Group on Data Communication,2017:488-501.

[17] BONICA R, BAKER F, HUSTON G, et al. IP fragmentation considered fragile: RFC 8900[S]. Reston: Internet Engineering Task Force,2020:9.

[18] WEST M,MCCANN S. TCP/IP field behavior:RFC 4413[S]. Reston:Internet Engineering Task Force, 2006:3.

[19] SALUTARI F,CICALESE D,ROSSI D J. A closer look at IP-ID behavior in the wild[C]//Proceedings of the 19th International Conference on Passive and Active Measurement,2018:243-254.

[20] MONGKOLLUKSAMEE S, FUKUDA K, PONGPAIBOOL P. Counting NATted hosts by observing TCP/IP field behaviors[C]//Proceedings of the 2012 IEEE International Conference on Communications, 2012:1265-1270.

[21] SCHULMANN H,ZHAO S J. Insights into SAV implementations in the Internet[C]//Proceedings of the 25th International Conference on Passive and Active Network Measurement,2024:69-87.

[22] HERZBERG A,SHULMAN H. Vulnerable delegation of DNS resolution[C]//Proceedings of the 18th European Symposium on Research in Computer Security, 2013:219-236.

[23] ZHENG X F, LU C Y, PENG J, et al. Poison over troubled forwarders:A cache poisoning attack targeting DNS forwarding devices[C]//Proceedings of the 29th USENIX Security Symposium,2020:577-593.

[24] Inetpeer:get rid of ip_id_count[EB/OL]. (2014-06-02)[2024-10-01]. https://git.kernel.org/pub/scm/linux/kernel/git/stable/linux-stable.git/commit/?id=73f156a6e8c1074ac6327e0abd1169e95eb66463.

[25] ZHANG X,KNOCKEL J,CRANDALL J R. ONIS:Inferring TCP/IP-based trust relationships completely off-path[C]//Proceedings of the 2018 IEEE Conference on Computer Communications, 2018: 2069-2077.

[26] ALEXANDER G,ESPINOZA A M,CRANDALL J R. Detecting TCP/IP connections via IPID hash collisions [J]. Proceedings on Privacy Enhancing Technologies, 2019(4):311-328.

[27] FENG X W,FU C P,Li Q,et al. Off-path TCP exploits of the mixed IPID assignment[C]//Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security,2020:1323-1335.

[28] Inet:Use bigger hash table for IP ID generation[EB/OL]. (2021-03-21)[2024-10-01]. https://git.kernel.org/pub/scm/linux/kernel/git/stable/linux.git/commit/?id=aa6dd211e4b1dde9d5dc25d699d35f789ae7eeba.

[29] GASSER O, SCHEITLE Q, FOREMSKI P, et al. Clusters in the expanse:Understanding and unbiasing IPv6 hitlists[C]//Proceedings of the Internet Measurement Conference,2018:364-378.

[30] CONTA A,DEERING S,GUPTA M. Internet control message protocol (ICMPv6) for the Internet protocol version 6 (IPv6) specification:RFC 4443[S]. Reston: Internet Engineering Task Force,2006:3.

作者简介:

黄峰元(2000—),男,江西高安人,硕士生,研究方向为网络测量。**E-mail**:huangfengyuan@nudt.edu.cn

HUANG Fengyuan, born in 2000, MS candidate, his research interest includes network measurement.

杨轶帆(2001—),男,湖南益阳人,博士生,研究方向为网络测量。**E-mail**:yangyifanyyf@nudt.edu.cn

YANG Yifan, born in 2001, PhD candidate, his research interest includes network measurement.

喻波(1985—),男,湖南宁乡人,博士,研究员,CCF 杰出会员(36538D),研究方向为软件安全与系统安全。**E-mail**:yubo0615@nudt.edu.cn

YU Bo, born in 1985, PhD, research fellow, CCF distinguished member(36538D), his research interest includes software security & system security.