文章编号:1001-5132（2008）04-0501-04

防御 DDoS 攻击的新过滤 PHF 模型

李　渊，金　光，张会展，陈　征，钱江波

（宁波大学 信息科学与工程学院，浙江 宁波 315211）

摘要：在分布式拒绝服务攻击对网络安全的危害日益严重的情况下，在众多的攻击防御技术中，采取路径标识是一种能有效对抗 DDoS 攻击的技术. 而为更有效地防御 DDoS 攻击，利用 Pi 方案中，受害主机使用 Pi 标记对收到的数据包进行过滤的方式，提出了结合 Pi 标记与跳数的新过滤模型，即受害主机采用<Pi，HC>元组识别和过滤攻击包方式. 并通过基于真实因特网拓扑的实验，证明 PHF 模型的防御效果明显优于 Pi 方案.

关键词：因特网安全；分布式拒绝服务攻击；路径标识；跳数

中图分类号：TP393.08　　文献标识码：A

分布式拒绝服务(Distributed Denial of Service，DDoS)攻击是目前乃至今后很长一段时间内因特网面临的严重安全威胁. 在典型的 DDoS 攻击中，攻击者先控制相当数量的傀儡机，然后利用它们向受害主机发送大量数据包，严重消耗受害主机的资源，使得合法用户无法访问受害主机[1]. 近年来，涌现了许多防御技术方案，而 Yaar 等人[2,3]提出的路径标识(Path Identification，Pi)及其改进方案堆栈路径标识(Stack Path Identification，StackPi)能迅速、直接、有效地在受害主机端实现防御.

本文在对原有 Pi 方案进行深入分析的基础上，对其过滤机制进行彻底的改进，提出了结合路径标识与跳数(Hop Count，HC)的新过滤模型(Pi-to-HC Filter，PHF). 原方案提出受害主机利用 Pi 标记识别和过滤攻击包，而我们建议受害主机不仅搜集和记录数据包的 Pi 信息，而且记录包的跳数信息，从而在对包进行过滤时，利用<Pi，HC>元组区分合法包与攻击包. 本文并且对数百万条因特网真实路径进行了仿真实验，而结果表明 PHF 模型比原 Pi 方案在性能上有 20%~44%的提高.

1　Pi 方案及其他相关方案

Yaar 等人[2,3]提出利用路径信息防御 DDoS 攻击的 Pi 方案，在其 Pi 方案中，路由器收到 1 个包之后，先把包头的标识(Identification，ID)域左移 1 或 2 位，然后将自己 IP 地址的 MD5 散列值的最后 1 或 2 位插入到 ID 域最右端. 这样，每个包到达受害主机时就带有它所经过的路径信息——Pi 标记. 由于攻击者所发包与合法用户发送包到达受害主机经过的路径不同，因此这些包带有的 Pi 标记也不同. 受害主机就可根据 Pi 标记区分合法包与攻

收稿日期：2008-03-04.　　宁波大学学报（理工版）网址：http://3xb.nbu.edu.cn
基金项目：浙江省自然科学基金（Y106023）；浙江省教育厅科研项目（20070978）；宁波市自然科学基金（2006A610014，2007A610007）；宁波大学人才基金（XR0710004）.
第一作者：李　渊（1983－），男，浙江宁波人，在读硕士研究生，主要研究方向：计算机网络及安全. E-mail: g06b08120302@email.nbu.edu.cn

击包，从而选择接收合法包或者丢弃攻击包.

在 DDoS 攻击中，攻击者为了防止被受害主机追踪，常常伪造 IP 源地址来隐藏自己的位置信息. Jin 等[4]提出了跳数过滤(Hop-Count Filtering，HCF)方案，防御利用了 IP 地址伪造技术的 DDoS 攻击. 因为攻击者无法准确猜测他所伪造的 IP 地址到受害主机之间的跳数，所以 HCF 建议受害主机记录收到的包的 IP 源地址和跳数值，并建立两者的映射表，然后使用这个表来识别收到的包是合法包还是攻击包.

2 PHF 的基本思想

IP 包头部有 16 位的 ID 域，Savage 等人[5]提出可以利用该 16 位记录信息为新的增强网络安全的方案服务，此种思想已被越来越多的方案所采纳[6]. Pi 方案正是用 ID 域存储包所经过的路径信息，根据不同发送者发送包经过的路径不同而导致 Pi 标记的不同来区分合法包与攻击包. 为增强过滤效果，Yaar 等人提出通过建立<Pi，IP>元组数据表来过滤收到的数据包[3]. 但是，每个元组需要 48 位的空间，这对于受害主机来说是个沉重的负担. 所以本文提出 PHF 模型，以跳数代替 IP 与 Pi 组成元组. 受害主机根据<Pi，HC>进行过滤，这样既提高了过滤效果，受害主机也不会承受太大的负担. 且跳数值可以很容易地通过存活时间(Time-to-Live，TTL)域的初值减去终值得到，因为大多数现代主流的操作系统只使用特定的几个 TTL 初值[7].

IP 包头部有 8 位的 TTL 域，它规定数据包经过路由器个数的上限. 包每经过 1 个路由器，它的 TTL 值就被减 1，如果 TTL 值为 0，该包就被丢弃. 可以看出，TTL 终值也反映了包所经过的路径信息. 如果说 Pi 标记反映路由器的地址信息的话，那么 TTL 终值就间接地反映了路径上路由器的个数信息. 2 个反映路径信息的值都存在于 IP 包的头部，因此如果能把这两者结合起来区分合法包与攻击包，防御效果必将得到提高.

在 PHF 模型中，路由器部署 Pi 机制，对经过它的每个包插入地址信息. 当包到达受害主机时，它的 ID 域就带有路由器的地址信息，而路由器的个数信息可以通过 TTL 初值减去终值得到. 受害主机提取包头的这 2 部分信息，建立 1 张<Pi，HC>的映射表. 当遭受攻击时，受害主机就可利用这张表判断包的合法性，从而执行接收或者丢弃的操作. 在本文方案的映射表中，每条信息需要 21 位(16+5)的空间. 即<Pi，HC>映射表占用受害主机的内存空间容量为 5.25 MB，这对于受害主机来说是个完全可以承受的存储开销.

3 实验仿真与结果分析

以下将通过实验对 PHF 模型和 StackPi 方案进行比较，选用的性能衡量标准是接收率之差，即合法包的接收率与攻击包的接收率之差，接收率差越大表示防御效果越好.

3.1 因特网数据集

为更好地仿真 PHF 模型在因特网真实环境下的性能，本文使用 CAIDA 提供的原始数据集作为实验的网络拓扑进行仿真测试[8]. 原始数据集通过某特定主机向因特网上的许多其他主机运行 traceroute 命令，然后记录每条路径上路由器的 IP 地址获得. 因实验需要，将运行 traceroute 命令的主机作为受害主机，把路径上目标端主机作为合法用户或攻击者. 在处理掉所有不完整的路径后，数据集共含有 5 252 850 条路径，其跳数分布如图 1 所示.

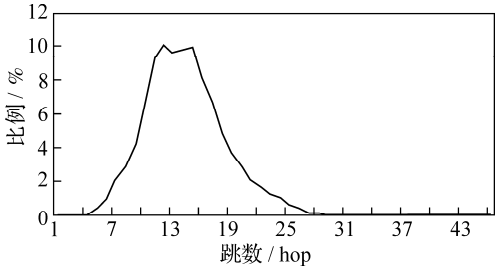


图 1 原始数据集的跳数分布

然后，在这些路径中选取两两不重复的 15 000 条路径，再以 1:2 比例随机确定其为合法或攻击，最后得到 5 000 条合法路径和 10 000 条攻击路径，作为仿真实验的实验数据集.

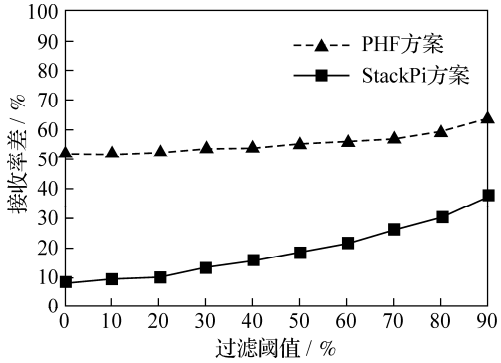
3.2　实验设计

在 StackPi 方案中，路由器有标记 1 位或标记 2 位的 2 种标记方法. 如果每个路由器标记 2 位地址信息，就至少需要 8 个路由器才能完全填充 ID 域，也就是 Pi 标记最多只能记录离受害主机最近的 8 个路由器的信息. 但是这将增加区分合法路径与攻击路径的难度，进而增加区分合法包与攻击包的难度，因为它们极有可能在到达受害主机之前汇聚，但此问题可以通过延长标记的路径长度来解决. 我们假设每条路径的最后 3 个路由器属于受害主机的自治系统，而自治系统被认为是安全的，所以规定路径上最后 3 个路由器不参与标记. 这样，Pi 标记就可以记录离受害主机最近的 11 个路由器的信息. 而如果路由器标记 1 位地址信息，则需要至少 16 个路由器才能完全填充 ID 域. 如果延长标记长度，就需要更多的路由器完成填充. 但是图 1 所示的跳数分布显示将会增加 ID 域不被路由器的信息完全填充的可能性，因此此时路径上的所有路由器都需参与标记.

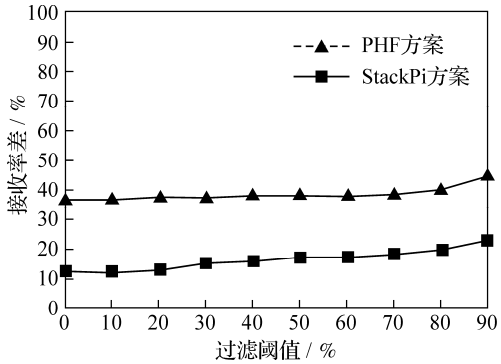
与 StackPi 方案一样，本文分学习阶段和过滤阶段来模拟 DDoS 攻击. 在学习阶段中，受害主机通过把数据包提交到应用层来鉴别包的合法性，并把攻击包的 Pi 标记或<Pi，HC>元组记录下来. 到了过滤阶段，受害主机就利用前阶段记录的信息，对收到的数据包进行过滤. 如果某个 Pi 标记或<Pi，HC>元组为攻击包与合法包所共有，本方案就采用阈值过滤的方法. 即对于某个 Pi 标记或<Pi，HC>元组；如果含有它的攻击包数量与含有它的所有包(攻击包加上合法包)数量的比值大于给定阈值，则丢弃含有该 Pi 标记或<Pi，HC>元组的所有包，反之，则接收所有的包. 本文设计的 DDoS 攻击规模是学习阶段的每个合法用户和攻击者分别发送 3 个包和 30 个包；过滤阶段的每个合法用户和攻击者分别发送 20 个包和 200 个包.

3.3　实验分析比较

对路由器的 2 种标记方法都进行实验. 如前所述的实验分为学习阶段和过滤阶段，实验结果全部来自过滤阶段. 在过滤阶段中，阈值过滤操作使用的阈值从 0～90%，防御的效果以接收率差来表示，实验的结果如图 2(a)和图 2(b)所示.



(a) 每个路由器标记 1 位



(b) 每个路由器标记 2 位

图 2　2 种方案接收率差的比较

从图 2 的实验结果可以看出，受害主机在 Pi 的基础上，结合跳数值对收到的包进行识别和过滤，不管路由器标记 1 位还是 2 位地址信息，都取得了比 StackPi 方案更加理想的防御效果. 表 1 则显示了在不同的过滤阈值下，PHF 的防御效果与 StackPi 比较的提高程度. 通过上述的实验验证，可

表 1　PHF 相比 StackPi 防御效果提高的程度

提高程度	n=1	n=2
最多提高	0.443 9	0.246 5
最少提高	0.265 3	0.205 1

以得出的结论是 PHF 模型能比 StackPi 方案更有效地防御 DDoS 攻击.

4 结语

本文提出的 PHF 模型对 Pi 方案进行了彻底改进，并建议利用 Pi 标记与跳数的结合来识别和过滤攻击包. 实验证明：PHF 模型相比原方案取得了更好的防御效果，而且新模型并未增加路由器的负担，仅借助于 IP 协议中规定的 TTL 操作；而在受害主机端，每个元组 21 位的空间要求也是受害主机完全可承受的存储开销. 因此在目前的因特网架构下，本方案的可行性较强.

接下来我们将继续深入开展 DDoS 攻击防御技术的研究，力争将本方案和其它技术相结合，以期更好地抑制此类攻击.

参考文献：

[1] 金光, 朱锡雄. 因特网防御 DoS 攻击技术评述[J]. 宁波大学学报: 理工版, 2004, 17(4):460-465.

[2] Yaar A, Perrig A, Song D. Pi: a path identification mechanism to defend against DDoS attacks[C]//In Proceedings of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, USA, 2003: 93-107.

[3] Yaar A, Perrig A, Song D. StackPi: new packet marking and filtering mechanisms for DDoS and IP spoofing defense[J]. IEEE Journal on Selected Areas in Communications, 2006, 24(10):1 853-1 863.

[4] Jin C, Wang H, Shin K G. Hop-count filtering: an effective defense against spoofed DDoS traffic[C]//In Proceedings of 10th ACM Conference on Computer and Communications Security, University of Karlsruhe, Germany, 2003:30-41.

[5] Savage S, Wetherall D, Karlin A, et al. Network support for IP traceback[J]. IEEE/ACM Transactions on Networking, 2001, 9(3):226-273.

[6] 金光, 赵杰煜, 赵一鸣, 等. 还原 DoS 攻击入口的地址元组标记模型[J]. 计算机研究与发展, 2004, 41(7):1 117-1 123.

[7] The Swiss Education and Research Network. Default TTL values in TCP/IP[EB/OL]. (2002-10-15)[2007-06-12]. http://secfr.nerim.net/docs/fingerprint/en/ttl_default.html.

[8] Caida. Skitter[EB/OL]. (2002-12-18)[2007-06-12]. https://sk-data.caida.org:8444/skitter-old/a-root/2002/.

A New Filtering PHF Model to Defend Against DDoS Attacks

LI Yuan, JIN Guang, ZHANG Hui-zhan, CHEN Zheng, QIAN Jiangbo

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: Distributed Denial of Service (DDoS) attacks have been posing more and more threats to cyber security. Among many proposed defending techniques, Path Identification (Pi) is a promising DDoS countermeasure. In the Pi scheme, the victim filters incoming packets based on Pi marks. To defend against DDoS attacks more efficiently, a new filter model is proposed of combining the Pi mark and the hop count (HC). In the proposed model, a victim host identifies and screens out malicious packets based on <Pi, HC> pair. The simulation experiments are conducted on real Internet topology, revealing that the PHF model outperforms the Pi scheme.

Key words: internet security; distributed denial of service; path identification; hop-count

CLC number: TP393.08 **Document code:** A

（责任编辑 章践立）