DDoS 攻击源追踪算法综述

黄忠厚, 徐 川, 刘宴兵
(重庆邮电大学 网络与计算机研究中心, 重庆 400065)

摘 要: 鉴于因特网出现了越来越多的 DDoS 攻击事件,结合 DDoS 攻击追踪方法的最新研究情况,对 DDoS 攻击追踪算法进行系统分析和研究,对不同的追踪算法进行比较,提出了追踪算法的重点,结合参考文献给出了解决问题的方法和意见,讨论了当前该领域存在的问题及今后研究的方向。

关键词: 分布式拒绝服务; 攻击源追踪; 网络安全; 攻击源定位; 包标记
中图分类号: TP393.08 **文献标志码**: A **文章编号**: 1001-3695(2010)09-3233-04
doi:10.3969/j.issn.1001-3695.2010.09.007

Survey on DDoS attack source tracing algorithm

HUANG Zhong-hou, XU Chuan, LIU Yan-bing
(*Network & Computation Research Center*, *Chongqing University of Posts & Telecommunications*, *Chongqing* 400065, *China*)

Abstract: This paper combined with the latest research on DDoS attack tracking methods, carried on system analysis and research to the DDoS attack tracking algorithm, and compared different algorithms with each other. It proposed the focus of the tracking algorithm and gave some method and advises to solve this problems by references. Finally, discussed the existing problems and the future direction in this field.

Key words: DDoS(distributed denial of service); attack source traceback; network security; attack source finding; packet marking

0 引言

随着互联网的发展,IPv6 将逐渐取代 IPv4 的下一代互联网协议,网络安全问题正日益为人们所关注[1]。分布式拒绝服务(DDoS)攻击是当今网络安全领域最难解决的问题之一,而且造成的破坏也越来越大。其中重要的原因在于网络上存在大量不安全的机器。由于 IP 协议在网络层不能提供认证服务以确保数据包的源地址是真实的,攻击者很容易伪造源地址对网络或网络中的节点实施洪流攻击,而通过数据包的源地址不能确定数据包的真正来源,需要更加有效的方法来定位攻击包的真正来源。

攻击源的追踪就是当攻击者正在进行时或已经结束后,按现有的能够获得的信息来确定攻击者的位置,按照准确度的逐渐提高,可分为追踪到发起攻击的网络、主机、进程和用户。目前针对 DDoS 攻击源追踪的研究工作主要是追踪这个主要的方向,国内外很多学者、研究机构更多的是从对 DDoS 攻击防御技术的研究和 DDoS 攻击追踪技术的研究,对 DDoS 攻击从防御、检测、响应到追踪形成一个体系。近年来,研究人员在 DDoS 攻击追踪领域中做了很多有意义的工作,提出了一些较成熟的算法并开发了相应的程序。研究人员根据不同的研究对象采用了不同的研究方法,也由此提出多种多样的 DDoS 攻击源追踪算法。文献[2]首先提出将路由器的地址信息标记到数据包中,使得追踪风暴型拒绝服务攻击的来源成为可能;在此基础上,文献[3]提出了概率包标记的方案,也称为基本包标记方案;文献[4]进一步提出了高级包标记和带认证的包标记方案来减少在 DDoS 攻击情况下误报严重的问题。近年来,包标记[3,5]一直是研究的热点,在国内也是广泛研究的课题,中国科学院一些研究者们提出了改进后的分块包标记方案[6]、有序包标记方案[7]等,在误报率、路径重构速度和包数量方面均有明显的改善。

当前 DDoS 攻击追踪的研究开展得如火如荼,但对该领域现有成果的综合评述尚未见全面报道。文献[8]提出通过利用 TTL 域对原 PPM 方案进行改进,优点是减少了路径重构所需的数据包数量,提高了路径重构的效率。文献[9]提出了基于 hash 编码的单个分组追踪方案(HBTB 算法)。该算法核心思想是:通过计算和存储每个分组的 32 位 hash 摘要来实现,使用 Bloom filter 的特别数据结构来尽量减少分组摘要的存储空间,因此,Snoren 等人提出了 SPIE 系统的技术模型,目的是用于对攻击后的数据进行追踪;其缺点是需要大量的存储空间,计算量大,需要读取分组中的代表性信息作为 hash 函数的输入,导致破坏隐私完整性的原则。如何快速、有效地确定攻击源,并及时采取相应的防御措施,就成了 DDoS 攻击防御研究中的关键课题。因此,研究 DDoS 攻击及其对策是非常重要的。

本文对 DDoS 攻击中已有的追踪算法的研究成果进行了总结和分析,分析了各种算法的性能和有待解决的问题,并归

收稿日期: 2010-03-26; **修回日期**: 2010-04-26

作者简介:黄忠厚(1984-),男,福建泉州人,硕士研究生,主要研究方向为网络安全(rwx840531@163.com);徐川(1980-),男,重庆人,博士研究生,主要研究方向为高速网络测量;刘宴兵(1971-),男,四川遂宁人,教授,博士,主要研究方向为网络路由、网格计算.

纳出 DDoS 攻击追踪算法所应具备的特征,指明了下一步研究中的重点及其今后的发展趋势。

1　DDoS 攻击追踪算法分类

就通常意义下的 DDoS 攻击追踪而言,其攻击追踪的共性目标主要包括降低运算复杂度、降低重构路径的差错率,以及收敛速度快、路径计算负荷小、误报率低等。

按照攻击源追踪算法分类可以分为两个大类:一类是在攻击完成后的追踪,主要有日志记录算法、数据包标记算法、ICMP 追踪算法等;另一类是在攻击追踪过程中的算法,主要有 IPSec、链路测试、逐跳追踪算法等,如图 1 所示。

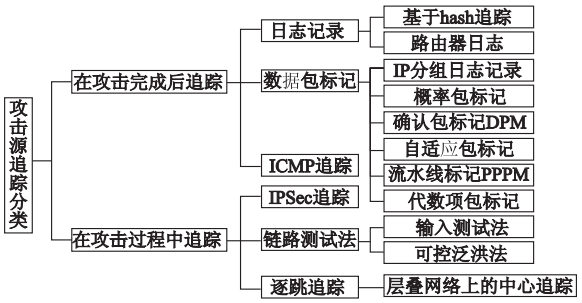

图1　DDoS攻击源追踪的分类

a)按数据包标记的算法有 IP 分组路径记录法、概率标记法、确定包标记法、入口包标记法、流水线标记法以及基于代数包标记算法等。

b)按日志记录的算法有路由器日志算法、基于 hash 追踪。

c)基于标记的追踪包括随机包标记(PPM)、固定包标记(DPM)、ICMP 追踪技术。

d)分组标记算法也有很多种分类,按照是否给每个分组都标记可以分为确定分组标记(DPM)和概率分组标记(PPM);按照分组标记中携带的地址信息可以分为节点标记、边标记和多边标记等。

e)按链路测试的追踪算法包括输入测试法、可控泛洪法等。

2　DDoS 攻击追踪相关算法

2.1　在攻击完成后追踪及攻击源定位

1)随机数据包标记算法

文献[10]中提出了采用基于概率数据包标记方案。该算法的核心思想是:数据包到达路由器时,路由器以某种概率来标记数据包的部分路径信息,当被攻击主机收到一定数量的带有标记信息的数据包时,可通过分析得到足够信息来恢复与构建完整的攻击路径。算法优点是实现事后追踪、无须耗用额外的网络带宽、无须占用路由器的存储空间,且不会增加所传送数据包的大小;缺点是标记效率有待提高、重复标记的问题、数据包标记容易被攻击者利用、PPM 算法追踪到的仅是边缘路由器或者攻击者侵入并控制的机器。

2)可变概率包标记算法

文献[10]中提出的概率包标记方案就是为了追踪攻击者,改变路由器的标记概率,使用可变的概率来进行标记,可以减少重构攻击路径所需的数据包数。可变概率包标记算法思想是:离攻击者越近,标记概率应越大。当取这样的变概率时,受害端重构攻击路径所需的包的数目最少,这样可以更好地追踪到攻击者。可变概率包标记算法优点很多,但是也有不足之处,三大关键问题总结如下:

a)在实现过程中,关键的问题是 i 值的确定,如何确定才能够提高算法的准确度。

b)路由器越靠近攻击者,其标记数据包的概率越大,边界路由器要对所有转发的数据包进行标记,产生的关键问题是将导致路由器的负担过重。

c)当 i 较大时,P_i 会很小,结果产生受害者需要得到更多来自于攻击者的数据包才能重构出攻击路径,耗时过长,易被攻击者利用。

文献[11]提升的改进的概率包标记算法方案是基于文献[10]提出的概率包标记方案并作出了改进,其只需四个有效分片就可以重构攻击路径,减少了 IP 的有效性验证次数,提高了重构攻击路径的速度。经过理论分析与实验结果证明了该方案的有效性。

3)ICMP 追踪算法

ICMP 追踪算法的思想是:当一个数据包通过一个路由器时,该路由器可以一定的概率同时向数据包的发送方和接收方发送带认证的 ICMP 追踪信息[12]。受害者可以挑出与那些攻击性数据包对应的追踪消息,根据其中的相关信息重组一个以受害者自己为目的地的攻击路径,攻击路径的另一端或攻击图的零入度节点处即为离攻击发出端最近的路由器。该算法的缺点有两点:a)被追踪的数据包与追踪包是分开的,它们可能因为路由策略或防火墙策略而使其中一个被丢弃,另一个被传输到受害者,从而使得追踪出现误差;b)在受害者获得攻击路径中所有路由器的追踪信息之前,其需要从攻击者处接收大量的包。

4)数据包日志记录法

数据包日志记录的基本思想是:在数据包的传输路途中,路由器将数据包的信息记录下来。这种方法最大的问题是要求很大的存储空间。Snoren 等人[9]提出了一种基于摘要(hash)的 IP 追踪方法。其优点在于检测到攻击后可以根据攻击数据分组在相关的 hash 值域内查询路由器,构造出给定数据包的路径;缺点在于要求所有的路由器都保存其转发过的数据包的部分信息的摘要。

5)基于代数编码的包标记算法

与基本包标记不同的是,基于代数编码的包标记不是直接将 IP 地址的分块标记到数据包中,而是在数据包中以代数编码的方式标记路由器的地址。文献[13]提出将路径上每个路由器地址对若干特定素数求余,结果分别作为多项式系数,多项式值再对一特定素数求余,依次累加,将结果写入 IP 首部的 ID 域后传送至下一站。这样,攻击数据包将含有路径上所有路由器地址的浓缩信息,受害主机收到大量数据包后,可通过解多项式和中国余数定理求出路径上的路由器。

该算法有两个缺点:a)由于没有距离标志,在路径重构的

过程中需要进行复杂的函数运算,其路径重构的结果不是很可靠;b)重构攻击路径所需的数据包的数量较大,这将延迟受害者对攻击路径的重构,从而有可能延误受害者处的其他防御措施,给受害者带来较大的损失。

2.2　在攻击过程中追踪及攻击源定位算法

1)基于 IPSec 安全关联的报文追踪算法

文献[14]提出了一种基于 IPSec 安全关联的报文追踪算法。当被攻击网络端的边界路由器发现遭到 DDoS 攻击时,向可能的位于攻击路径上的上游路由器请求建立隧道模式 IPSec 安全关联;上游路由器对所有发往 Rv 的网络报文进行认证,边界路由器根据被认证的报文中是否有攻击报文,来确定上游路由器是否真正位于攻击路径上。该算法的缺点是需要对所有可能的报文都进行认证,而且需要管理员的相互协作。

2)链路测试算法

链路测试算法是反应回溯技术的一种,一般是从离被攻击者最近的路由器开始检查,逐级回溯到离攻击者最近的路由器。其优点在于与现有协议兼容,并且与现有的路由器和网络设施兼容,可以逐步地实现;缺点在于要成功溯源需要攻击持续时间足够长,而且不适合应对 DDoS 攻击,多个网络服务提供商之间的协调较困难。

2.3　其他经典攻击追踪算法

1)实时追踪攻击源追踪算法

实时追踪攻击源追踪算法的核心思想是:基于自治域系统(AS)的概率标记算法将攻击源确定在某些 AS 中,然后在 AS 自治域范围内再使用随机数标记算法(RNPM)精确定位攻击源位置。该算法的优点是收敛速度快、路径计算负荷小以及较低的误报率等,适合实现对 DDoS 攻击的实时追踪。

2)追踪部署的着色包标记算法

追踪部署的着色包标记算法的核心思想是:基于追踪部署的相关理论和着色包标记算法,提出一种基于追踪部署的 IP 回溯算法。该算法是以贪心算法为基础,利用 K-剪枝算法[15]在网络拓扑图中找出一些关键的路由器。该算法的优点是减少了重构路径所需的数据包数,降低了路径误报率,提高了追踪到攻击者的速度,而且大大减轻了路由器标记的负担,从而能够迅速准确地找到攻击源。

如何防御分布式拒绝服务攻击是当今最重要的网络安全问题之一,IP 追踪和攻击源定位技术在 DDoS 攻击防御研究中具有重要意义。

3)基于认证的反射 DDoS 源追踪算法

基于认证的反射 DDoS 源追踪算法的核心思想是:利用基于密钥集序列的消息认证码理论,采用新的动态概率序列和基于密钥集序列的 HMAC 算法对标记信息进行认证,防止攻击者修改已有的标记信息,达到较高的安全性和抗干扰性。其优点是能有效地过滤掉攻击者伪造的垃圾数据包,同时对反射 DDoS 攻击也能进行有效的追踪;缺点是计算量大,本方案与 IPv6 协议的兼容性不好,需要优化反射标记算法的数据结构。

文献[16]在对概率包标记算法分析的基础上给出了一种动态概率包标记 ADPPM。该算法核心思想是:利用标记域中的 distance 域确定标记概率,并采用点标记与边标记的结合对封包进行标记,提升了封包来源回溯的效率,让受害者用最少的反查信息就可以重建攻击路径。该算法的优点是减少了反查所需的封包数和时间,可以更有效地发现攻击源。文献[17]分析攻击路径距离、路由器节点流量统计对标记概率的影响,改进了一种复合包标记方法。该算法的优点是优化算法收敛性,降低运算复杂度和重构路径的差错率,使受害者在最短时间内推测出主要攻击路径,能够很好地应用于多个 DDoS 攻击的攻击源追踪中。

3　现有追踪算法性能比较

由于 DDoS 攻击具有应用密切相关性,DDoS 攻击追踪算法表现出多样性的特点,难以直接评判孰优孰劣。为便于说明,本文采用列表对上述的几种经典算法和普通算法进行比较。

表 1 对包括如收敛性、算法开销、路由器开销、管理开销、网络开销、兼容性等几项算法属性进行了对比。

表 2 从存储耗费、计算开销、协议兼容、自身安全等属性进行对比,可以更好地分析比较 DDoS 攻击追踪算法的追踪效果。

表 1　几种追踪算法不同类型的比较

追踪算法	收敛时间	算法开销	路由开销	管理开销	网络开销	兼容性
采样边标记	较快	较大	较小	小	小	较差
入口包标记	较快	小	小	小	小	差
ICMP 定位报文法	慢	较小	小	小	小	差
采样节点标记法	慢	较小	较小	小	小	较差
覆盖网络中心追踪	较快	较大	较大	大	较大	差

表 2　几种追踪算法不同类型的比较

类型	存储耗费	计算开销	DDoS 攻击	协议兼容	自身安全
节点附加	无/低	低/低	不可容忍	部分兼容	差
节点取样	无/很高	低/很高	不可容忍	否	差
边取样	无/高	中/中	不可容忍	否	好
压缩边取样	无/高	中/很高	有漏洞	部分兼容	差
即时概率多边	无/低	低/低	可以容忍	是	好
确定分组标记	低/高	低/低	可以容忍	是	好
代数多项式	低/高	低/高	不可容忍	部分兼容	差
概率流水分组	中/中	中/中	可以容忍	否	一般
Huffman 编码	中/低	中/低	可以容忍	是	一般

4　进一步的工作和总结

DDoS 攻击源追踪的研究是推动追踪效率进一步发展的关键问题。DDoS 攻击源追踪算法的评价取决于算法的效率,高效的算法应同时呈现在三个方面,即收敛速度快、路径计算负荷小和误报率较低。从上述的分析可知,实质上 DDoS 攻击追踪能够使受害者在最短时间内推测出主要攻击路径、重构路

径,并在此基础上还需考虑算法本身实现的代价、现实环境中网络的不可预知性等方面。从研究方法看,针对特定环境及需求建立相应数学模型,随后寻找一种能够被数据包标记所允许的近似解法,此种方法成为一种较合适的算法设计方式。

本文从以下 10 个方面归纳了今后 DDoS 攻击源追踪需关注和研究的问题:

a)追踪仅仅是 DDoS 攻击解决方案中的一个步骤,如何将追踪与其他防范措施相结合尽可能兼顾算法的实现代价和部署性,必将成为 DDoS 攻击追踪研究中的重点和难点之一。

b)方案的移植问题。因为 IPv6 没有标志字段,现有的数据包标记方法在向 IPv6 移植时存在困难,有待进一步研究。

c)如何利用少量路由器存储空间和带内信息传输方式来提高标记信息利用率,如何避免占用额外的网络通信宽带。

d)如何使受害者通过很少的数据包来完成攻击路径重构工作,如何在重构路径时有效地降低攻击路径的误报率,从而能够在更短的时间内找出攻击路径并发现攻击源。

e)如何提高 DDoS 攻击追踪效率,比较有效地过滤掉攻击者伪造的数据包,如何增强包标记追踪的安全性和可靠性。

f)基于主动防御模型的 IP 反向追踪方法利用 hash 函数序列的相关性和概率标记信息,通过主覆盖路由器的协同工作,实现对攻击源快速定位。其特征是在标记时对摘要信息进行拆分,保证了对大范围 DDoS 攻击的准确定位。因此 IP 反向追踪也将是 DDoS 攻击追踪研究中的关注点之一。

g)今后,高速网、无线网和 IPv6 上的 IP 溯源技术将是研究的热点。目前 IP 溯源技术的研究集中在 IPv4 上,但随着高速网、无线网的增多以及 IPv6 技术的推广,针对它们的 DDoS 攻击肯定会多起来,所以这方面的研究应该起步。

h)如何使得一个节点的标记信息不被其他节点所伪造,而又不能让攻击主机重构出攻击路径。在包标记算法上,如何通过减少收敛时间以更优化的标记概率来实现。

i)如何比较有效地减少伪造标记数据包产生的攻击路径漏报和误报,提高追踪效率,如何将 DDoS 攻击追踪技术与其他技术结合起来(如 pushback 技术[18]、filtering 技术[19])共同应对网络攻击。DDoS 攻击追踪技术有着明显的优缺点,与其他技术结合,取长补短是可行之举。

j)如何能够追踪到具体的攻击主机而不仅仅是追踪到它所在的网段,以及对下一代网络的支持问题。另外,进一步提高溯源路径精度和准确度也将是一个重要的课题。由于溯源精度和溯源开销的矛盾始终存在,如何在其间选择一个平衡点是需要权衡的问题。

现有的 DDoS 攻击追踪问题主要着眼于包标记算法的研究,通过重构路径、减少收敛时间、提高追踪效率、实现追踪,这也是笔者未来研究工作的重点之一。本文中的经典算法在理论和模拟的环境中实现具有可行性,但实际效果还需要在进一步的工作中不断完善和深入,因此在下一步工作中具有继续研究的必要和意义。

参考文献:

[1] MA Ting. A link signature based DDoS attacker tracing algorithm under IPv6[J]. International Journal of Security and Its Applications, 2009, 3(2): 27-36.

[2] BURCH H, CHESWICK B. Tracing anonymous packets to their approximate source[C]//Proc of the 14th Conference on Systems Administration. [S.l.]: USENIX Association, 2000: 319-327.

[3] SAVAGE S, WETHERALL D, KARLIN A, *et al*. Network support for IP traceback[J]. ACM/IEEE Trans on Networking, 2001, 9(3): 226-237.

[4] SONG D X, PERRIG A. Advanced and authenticated marking schemes for IP traceback[C]//Proc of the 20th Annual Joint Conference on IEEE Computer and Communications Societies. [S.l.]: IEEE Press, 2001: 878-886.

[5] PARK K, LEE H. A proactive approach to distributed DoS attack prevention using route based packet filtering, CSD00-017[R]. West Lafayette: Purdue University, 2000.

[6] 曲海鹏, 李德全, 苏璞睿, 等. 一种分块包标记的 IP 追踪方案[J]. 计算机研究与发展, 2005, 42(12): 2084-2092.

[7] 曲海鹏, 冯登国, 苏璞睿. 基于有序标记的 IP 包追踪方案[J]. 电子学报, 2006, 34(1): 173-176.

[8] 陈星星,徐红云. IP 追踪中 PPM 算法的改进研究[J]. 计算机工程, 2006, 32(21): 164-166.

[9] SNOREN A C, PARTRIDGE C, SANCHEZ L A. Hash-based IP traceback[C]//Proc of Conference on Application, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM, 2001: 3-14.

[10] SAVAGE S, WETHHERALL D, KARLIN A, *et al*. Practical network support for IP traceback[J]. ACM SIGCOMM Computer Commuaication Review, 2000, 30(4): 295-306.

[11] 张立莉,曹天杰,汤丽娟. 一种改进的概率包标记方案[J]. 计算机工程, 2008, 34(7): 148-150.

[12] BELLOVIN S, LEECH M, LOR T T. ICMP traceback messages[EB/OL]. (2003-08). http://www.cs.columbia.edu/~smb/papers/draft-bellovin-itrace-00.txt.

[13] DEAN D, FRANKLIN M, STUBBLEFIELD A. An algebraic approach to IP traceback[J]. ACM Information and Trans on System Security, 2002, 5(2): 119-137.

[14] CHANG H Y. On real-time intrusion detection and source identification[D]. Raleigh: North Carolina State University, 2000.

[15] 刘渊, 陈彦, 李秀珍. 基于追踪部署的着色包标记算法的研究[J]. 计算机应用研究, 2008, 25(10): 3102-3104.

[16] 奚洋, 夏洪山. IP 追踪中概率包标记算法研究[J]. 计算机工程与设计, 2009, 30(12): 2899-2901.

[17] 高大鹏, 於时才, 闫文芝. 复合包标记 IP 追踪算法研究[J]. 计算机工程, 2009, 35(10): 115-117.

[18] IOANNIDIS J, BELLOVIN S M. Implementing pushback: router-based defense against DDoS attacks[C]//Proc of the 9th Annual Symposium on Network and Distributed System Security. 2002: 100-108.

[19] SUNG M H, XU J. IP traceback-based intelligent packet filtering: a novel technique for defending against Internet DDoS attacks[J]. IEEE Trans on Parallel and Distributed Systems, 2003, 14(9): 861-872.