

# 基于合作过滤机制的 ICMP 回溯 DDoS 攻击源方法

罗淇方<sup>1</sup>, 钟 诚<sup>1</sup>, 李 智<sup>1,2</sup>, 杨 锋<sup>1</sup>, 冯艳华<sup>1</sup>

(1. 广西大学 计算机与电子信息学院, 广西南宁 530004;

2. 广西科技信息网络中心, 广西南宁 530012)

**摘 要:**提出一种改进的 ICMP 回溯方法。此方法是基于合作过滤机制的, 采用合作过滤机制能够使产生的 ICMP 回溯包更有效并在尽可能靠近 DDoS 攻击源的地方过滤攻击包和保护合法包, 改进后的 ICMP 方法对引发 ITrace 包的 IP 包从靠近攻击源的地方同步跟踪到受害者, 提高了重构攻击路径的速度和准确性。

**关键词:**ICMP 回溯; 合作过滤机制; DDoS

**中图分类号:**TP393.08

**文献标识码:**A

**文章编号:**1673-629X(2006)09-0240-03

## An ICMP Traceback Method of DDoS Attack Source Based on Cooperative Filtering Mechanism

LUO Qi-fang<sup>1</sup>, ZHONG Cheng<sup>1</sup>, LI Zhi<sup>1,2</sup>, YANG Feng<sup>1</sup>, FENG Yan-hua<sup>1</sup>

(1. School of Computer and Electronics and Information, Guangxi University, Nanning 530004, China;

2. Guangxi Science and Technology Information Network Center, Nanning 530012, China)

**Abstract:** Presents an improved ICMP traceback method based on cooperative filtering mechanism. The cooperative filtering mechanism can generate ICMP traceback packets more efficiently and filter attack packets and protect legal packets the location of DDoS attack source as closely as possible. The improved ICMP traceback method tracks victim synchronously from the location near the attack source by the IP packets which evokes ITrace packets, and it enhances the speed and accuracy of reconstructive traceback path.

**Key words:** ICMP traceback; cooperative filtering mechanism; DDoS

### 1 分布式拒绝服务攻击回溯方法研究现状

IP 包记录(IP Logging)方法<sup>[1,2]</sup>是通过关键路由器记录经过互联网的数据包, 然后利用数据挖掘技术提取攻击源的信息。虽然这种方法表面上看起来可以准确地分析攻击流量, 但是它有个很大的缺点就是路由器需要大量的处理和存储能力。同时, ISPs(Internet Service Providers)之间需要共享这些信息。就当前的链接速度来看, 数据包记录会增长到无法管理的地步。即使采用随机抽样数据包记录并采用压缩技术来减少资源的需求, 但是需求量仍然很大。文献[3]联合相邻路由器的攻击包的采样建立攻击树来减少采样和记录包的数量。

IP 包标记(IP Marking)方法<sup>[4,5]</sup>的思想是由从互联网到目的主机的数据包所经过的路由器把回溯数据记入 IP 包中, 然后主机可以利用每个包里的标记信息来推断

出信息量经过的路径。该方案的不足之处是需要在数据包首部增加空间, 使得路由器的负载过高, 并可能导致数据包分片, 以及与 MTU(Maximum Transmission Unit)技术的不兼容, 随着分布式攻击路径的增加, 重构路径的计算量会呈指数级增长。文献[6]提出了使用回推机制进行 IP 回溯的方法。

ICMP 回溯(ICMP Traceback)方法<sup>[7]</sup>在路由器中增加跟踪机制来实现路由追踪, 称为 ITrace 路由器。一个 ITrace 路由器以 1/20000 的概率发送对数据包的拷贝, 该拷贝是一种特殊类型的 ICMP 数据包, 其中记录发送它的路由器的 IP 地址, 以及其前一跳和后一跳路由器的 IP 地址, 还包括诱发它的数据包的信息。ITrace 路由器向源或目的地址都转发该 ICMP 数据包。受害者收集足够多的 ITrace 数据包, 就可以找出攻击路由。文献[8]采用累积路径的 ICMP 回溯方法, ITrace 包存储了相应 IP 包经过的完整路径, 这样重构攻击路径相对容易, 但 ITrace 路由器要存储 IP 包来对照接收到的 ITrace 包。文献[9]根据路由器到接收者的距离使用动态调节的概率产生 ITrace 包来代替原来的固定概率。

由于 IP 网络路由结构<sup>[10]</sup>没有对 IP 包的源地址的真实性进行认证, 目的主机只是用源地址来回复源信息, 通

收稿日期: 2005-12-08

基金项目: 广西科学基金(桂科自 0339008); 广西科技信息网络中心和广西大学博士科研基金(B0309031)

作者简介: 罗淇方(1976-), 女, 广西北海人, 硕士研究生, 研究方向为网络信息安全; 钟 诚, 教授, 硕士生导师, 研究方向为网络信息安全与并行计算。

常也没有实体对源地址的正确性进行响应。结果,IP协议和它相关机制设置使它很难识别包的真正来源。恶意用户就利用网络的这个特点来隐藏它们的源地址和身份。

尽管IP包头中的源地址可能是虚假的,但每个IP包都必须经过从攻击方到受害者之间的路由器转发,如果能够记录下这些路由器,就可以重构出攻击所经过的路径,这也正是攻击路由反向追踪的基本思路<sup>[11]</sup>。

接收端倘若没有受到攻击,这时路由器产生ITrace信息是没有必要的,同时它们的产生还降低了路由器的追踪性能。当利用ITrace信息构造攻击路径时<sup>[7]</sup>,由于每条ITrace信息只带有全部路径的一或两个连接,这些片断都是对应不同的IP包,将大量的片断组合起来,这项任务在分布式拒绝服务攻击(DDoS)中较难实现。如果在一次攻击刚刚开始就产生ITrace信息包,DDoS检测系统就能够有充足的时间处理攻击事务,此时针对攻击包产生的ITrace信息才是最有价值。文中提出的基于合作过滤机制的ICMP回溯方法,采用合作过滤机制并改进文献<sup>[7]</sup>的ICMP回溯方法,通过具有过滤检测追踪功能的路由器向上游也具有同样功能的路由器发送消息来回溯到攻击源,在最接近攻击源的地方对检测出的攻击包产生ITrace包并标记该IP包,对有标记的IP包在它经过的每一跳都产生ITrace包,ITrace包中包含有相应IP包的32位哈希值,将具有相同哈希值的ITrace包中的路径提取出来连接起来就是攻击路径,该方法提高了攻击路径重构速度和准确性,由于它是通过路由器的流量丢包率和接收攻击包特征检测自身流量决定标记攻击包来引发相应的ITrace包,与以一定概率发送ICMP回溯包的方法相比不会有明显的ITrace的流量增加。

## 2 基于合作过滤机制的ICMP回溯的结构

假设所有的路由器都加载了跟踪机制,既可以发送ICMP回溯(ICMP traceback,简称为ITrace)包,这里称为ITrace路由器,又在网络中关键节点路由器(例如:边界路由器,有多个分支的路由器等)上加载过滤机制,并将其称为Filter路由器,Filter路由器不仅能发送ITrace包还能向上游的Filter路由器发送Filter消息。Filter路由器分布在网络之中,下游路由器可以向上游路由器发送消息,通过这种相互合作找到攻击源。这里的ITrace包也是如文献<sup>[7]</sup>一样是对相应IP包的一个特殊拷贝,但其中加载的是IP头的116位(包括4位版本号、8位服务类型、16位总长度、16位标识、8位协议、32位源IP地址、32位目的IP地址)不变部分的32位哈希值、当前路由器地址以及下游路由器的地址的XOR值。这里的过滤机制通过检测数据包的特征来阻止DDoS攻击,它由检测模块和流量控制模块组成。Filter路由器的过滤回溯结构如图1所示。

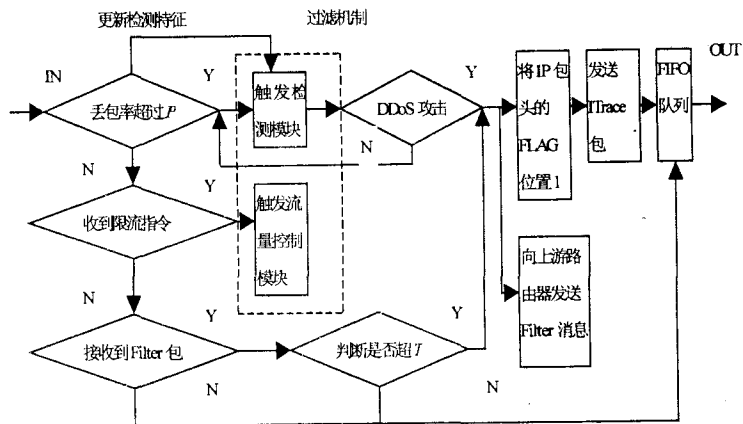


图1 Filter路由器的过滤回溯结构

将Filter路由器配置在网络中的各关键结点中,因为Filter路由器对硬件的要求比较高,为了节约网络开销本文只把它们布置在网络中的关键节点上。当Filter路由器丢包率超过阈值 $P$ 时即有可能发生DDoS攻击时才触发检测,这样可以节约路由器资源。Filter消息里带有DDoS攻击包的特征,Filter路由器提取出这些特征对通过的流量进行检测,如果只检测到一个具有DDoS攻击特征的IP包就发送ITrace消息,Filter路由器显得过于敏感,因此文中给定一个阈值 $T$ ,当检测到的DDoS流量超过这个值时才发送ITrace消息。Filter路由器通过向上游的Filter路由器发送Filter消息来回溯到攻击源,这样也许发现的不是真正的攻击源但至少也能发现发送DDoS攻击的傀儡机的区域。采用的合作过滤机制和回推相似,但回推并不能确认DDoS攻击流量,它会在过滤DDoS攻击流量的同时把合法流量也过滤掉。文中的方法在检测出攻击包时有针对性地对攻击包进行过滤,尽最大可能保护合法流量。

## 3 基于合作过滤机制的ICMP反向回溯过程和回溯路径重构

### 3.1 基于合作过滤机制的ICMP反向回溯过程

基于合作过滤机制的ICMP反向回溯过程描述如下:

(1)Filter路由器根据丢包历史周期性地计算丢包率,如果超过 $P$ 则报警,触发过滤机制的检测模块进行检测,如检测到DDoS攻击,则提取出攻击包的特征,向上游的Filter路由器发送带有攻击包特征的Filter消息,检测模块继续检测,直到丢包率小于 $P$ 值。同时它向相应攻击包的地址发送一个ITrace包,对于产生ITrace包的IP包路由器将IP包头的三位标志域(FLAG)的第一位置1(这个域现在还没有用),以表示这个IP包已经产生过ITrace包。

(2)当Filter路由器接收到一个Filter消息时,它将接收到的特征与经过本路由器的数据包进行对比,如果有此特征的数据包超过给定阈值 $T$ 时,它向相应IP包的地址发送一个ITrace包,同样对于产生ITrace包的IP包路由器将IP包头的三位标志域(FLAG)的第一位置1,该

路由器继续向上游 Filter 路由器发送相同的攻击包特征的消息包。

(3) 每个 ITrace 路由器接收到标记过的 IP 包后, 产生一个相应的 ITrace 包。

(4) 受害者根据收到的 ITrace 包, 重构攻击路径。

(5) 受害者向重构路径中包含的 Filter 路由器发送限流信息。

(6) 路由器启动过滤机制的流量控制模块过滤攻击包。

在文献[7]的方法中每个 ITrace 路由器采用固定的概率产生 ITrace 包, 发出 ITrace 包的路由器可能并没有攻击包流过, 这样的 ITrace 包对接收端来说是无用的, 构造出的攻击路径就会包括很多非攻击路径。文中的 ICMP 回溯方法则是对可能的攻击包从攻击源一直追踪到受害者端, 接收端收集到的 ITrace 包包含攻击包经过的所有路径信息, 构造出的攻击路径将会更准确。

IP 头的不变部分包括 4 位版本号、8 位服务类型、16 位总长度、16 位标识、8 位协议、32 位源 IP 地址、32 位目的 IP 地址共 116 位可以惟一标识一个 IP 包, 这里的 ITrace 包包括 116 位的 IP 包惟一标识、32 位当前路由器 IP 地址、32 位下一跳路由器 IP 地址的信息。文中将 116 位的 IP 包惟一标识通过哈希函数转换成 32 位的哈希值, 此哈希函数的选取应尽量使输出函数结果不相同以保持输入值的惟一性, 相同的 IP 包产生的 ITrace 包中的哈希值是相同的, 接收者将这些包中的路径提取出来经过 XOR 操作就可以重构出完整的攻击路径。

### 3.2 回溯路径重构

回溯路径重构方法是:

(1) 通常网络的最大跳数是 32, 包传输的距离能通过 IP 包头 8 位的 TTL(生存时间)值的低 6 位计算出来, 这里路由器提取出攻击包  $I$  的 TTL 值的低 6 位保存到相应产生的 ICMP 回溯包中 8 位的 TOS(服务类型)中, 设 TOS 的值为  $S$ ,  $S = I$  中的  $TTL \wedge 00111111$ 。

(2) 从攻击包  $I$  中提取出 116 位的  $M$ ,  $M$  是 IP 包的惟一识别特征信息。将  $M$  通过哈希函数转换成 32 位的哈希值  $H$ , 32 位的哈希值产生冲突的概率非常小。

(3) 产生 32 位的  $AX'$ ,  $AX' = AY \oplus AX$ 。 $AX$  是对应 IP 包的当前路由器 IP 地址,  $AY$  是此 IP 包传送到下一个路由器的 IP 地址。

(4) 通过 ICMP 回溯消息来发送  $H$  和  $AX'$ 。

(5) 接收者从 ICMP 回溯消息中提取哈希值和路由器地址信息。

从受害者处接收到的 ITrace 包具有相同  $H$  值的包即为同一个 IP 包经过的 ITrace 路由器发出的 ITrace 包, 通过 ITrace 包中的 TOS 值的大小来判断路由器离受害者的距离, 从而决定提取出的路径 XOR 操作的先后顺序。在离受害者最近的路由器发出的 ITrace 包必然只包含有一个路由器地址  $V$ , 能通过  $AX = AX' \oplus V$  (因为  $A = B \oplus A$

$\oplus B$ ) 来得到离受害者次远的路由器的地址, 重复这一操作, 受害者就可以计算出攻击包所经过的传输路径。

上述方法受害者不需要知道网络拓扑结构, 只需要通过收集到的 ITrace 包就可以计算出攻击路径。对具有相同哈希值的 ITrace 包分别进行路径重构要比对所有收集到的 ITrace 包一起进行路径重构速度要快, 且 ITrace 包是由攻击包产生的, 标记过的攻击包在每一跳都产生 ITrace 包, 因此收集到的 ITrace 包包含有相应攻击包的经过的所有路由器的信息, 构建出的路径更具准确性。

### 4 结束语

采用合作过滤机制能够使产生的 ICMP 回溯包更有效并在尽可能靠近 DDOS 攻击源的地方过滤攻击包和保护合法包, 改进后的 ICMP 方法对引发 ITrace 包的 IP 包从靠近攻击源的地方同步跟踪到受害者, 提高了重构攻击路径的速度和准确性。该方法虽然对标记过的 IP 包在每一跳都发送 ITrace 包, 但它是通过路由器的流量丢包率和接收攻击包特征检测自身流量决定标记攻击包来引发相应的 ITrace 包, 与以一定概率发送 ICMP 回溯包的方法相比不会有明显的 ITrace 的流量增加, 在大规模多攻击路径的 DDOS 攻击中该方法更具有优势。但标记 IP 包头会产生新的安全隐患, 如果攻击者人为地将所有 IP 包的标志域的第 1 位置 1, 那么所有的包都会在它经过的路由器处产生 ITrace 消息, 这样必然会增加网络流量, 加重 DDOS 攻击的危害。为了防止恶意用户伪造和篡改 ITrace 包和标记过的 IP 包, 可以采用加密技术来保证这些信息的可靠性。

### 参考文献:

- [1] Snoeren A C, Partridge C, Sanchez L A, et al. Single - Packet IP Traceback[J]. IEEE/ACM Trans. Networking, 2002, 10 (6): 721 - 734.
- [2] Baba T, Matsuda S. Tracing Network Attacks to Their Sources [J]. IEEE Internet Computing, 2002, 6(2): 20 - 26.
- [3] Li Jun, Sung Minho, Xu Jun, et al. Large - Scale IP Traceback in High - Speed Internet: Practical Techniques and Theoretical Foundation[A]. Proceedings of the 2004 IEEE Symposium on Security and Privacy[C]. New York: IEEE Press, 2004. 115 - 129.
- [4] Savage S. Practical Network Support for IP Traceback[A]. Proc of ACM SIGCOMM 2000[C]. New York: ACM Press, 2000. 295 - 306.
- [5] Song D, Perrig A. Advanced and Authenticated Marking Schemes for IP Traceback[A]. Proc of IEEE INFOCOM 2001 [C]. New York: IEEE Press, 2001. 878 - 886.
- [6] Lee Hyung - Woo. Advanced Packet Marking Mechanism with Pushback for IP Traceback[A]. Proc of the Second International Conference on Applied Cryptography and Network Security

(下转封三)

### 2.3 分类算法

分类在数据挖掘中是一项非常重要的任务,其目的是学会一个分类函数或者分类模型(也称为分类器),该模型能把数据集中的记录映射到给定类别中的某一个。

给定一训练数据集  $T$ ,  $T$  中的记录由若干属性描述(例如表 1 中的网络连接记录)。所有属性中有且仅有一个称作类别(class label)的属性。属性集合用向量  $X = (X_1, X_2, \dots, X_n)$  表示,其中  $X_i (1 \leq i \leq n)$  对应各非类别属性,可具有不同的值域。用  $C$  表示类别属性,  $C = \{c_1, c_2, \dots, c_k\}$ 。  $T$  隐含确定了一个从向量  $X$  到类别函数  $H: f(X) \rightarrow C$ , 分类的目的就是要将这个隐含关系  $H$  表示出来。

入侵检测从数据分析的观点来看,可以看作是一个分类的过程。可以把各种连接记录看成是正常(normal)或者某种类型的攻击(attack)。对于一个带有类标签的记录集合,分类算法利用最具区别性的特征值来描述每一条记录。分类模型的精确性直接取决于训练数据集中所提供的属性集(特征集合),选择一个合适的属性集合是形成一个有效分类器的关键。在构建入侵检测分类模型的时候,一般是先需进行关联分析和时序分析,挖掘出关联规则和频繁时序模式,然后以此来指导特征提取及连接记录的瞬时统计特征的构建工作。常用的分类算法有 ID3, C4.5, RIPPER<sup>[3]</sup> 等。

### 3 数据挖掘的入侵检测模型

哥伦比亚大学的 Wenke Lee 等人最早将数据挖掘技术应用到入侵检测领域,提出了基于数据挖掘的入侵检测系统框架<sup>[4,5]</sup>,同时进行了大量仿真实验,取得了较好的实验结果,证实了数据挖掘应用在入侵检测领域的可行性和有效性。图 1 为一个基于数据挖掘的入侵检测模型。

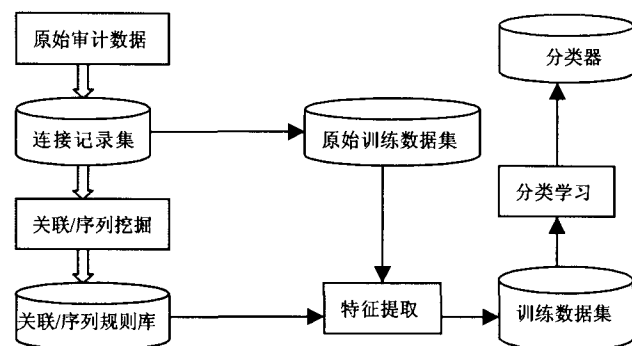


图 1 一个基于数据挖掘的入侵检测模型

从模型中可以看出,构建基于数据挖掘的入侵检测系统,需经过以下步骤:

- (1) 收集原始数据(如用 tcpdump 收集),对数据进行转换处理形成连接记录集合  $R$ 。
- (2) 对集合  $R$  应用关联规则算法和序列规则算法,生成关联/序列规则。
- (3) 利用(2)对训练数据集进行特征提取(如形成一些时间统计特征等),最终形成适用于分类的数据集  $T$ 。
- (4) 对集合  $T$  应用分类算法(如 RIPPER 算法),生成分类器。

### 4 结束语

入侵检测可以被看作是一个分类问题,将数据挖掘应用到入侵检测上,发挥了数据挖掘在进行海量数据的处理和分析上的优势。利用数据挖掘中的关联分析和序列分析算法可以分别找出属性之间和记录之间的关联,从而可以被用来指导构建分类模型。文中主要对从宏观上去构建一个模型进行了研究,但一些细节在模型构件过程中的某一部分可能是非常重要的,例如在模型构建过程中数据的预处理部分,包含了数据清洗、数据变换、数据离散化等一系列工作,它是一切后续工作的基础。数据挖掘应用到入侵检测领域目前还处于研究阶段,以后要做的工作是对数据挖掘中的有关算法进行改进,使这个检测模型实现。

### 参考文献:

- [1] Agrawal R, Srikant R. Fast algorithms for mining association rules in large database[R]. In Research Report RJ 9839, San Jose, CA: IBM Almaden Research Center, 1994.
- [2] Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases[A]. Proceedings of the ACM SIGMOD Conference on Management of data[C]. [s.l.]: [s.n.], 1993. 207 - 216.
- [3] Cohem W W. Fast Effective Rule Induction[A]. In Proceedings of the Twelfth International Conference on Machine Learning (ICML - 95)[C]. Lake Tahoe, CA: Morgan Kaufman, 1995. 115 - 123.
- [4] Lee W. A Data Mining Framework for Constructing Features and Models for Intrusion Detection Systems[D]. New York: Columbia University, 1999.
- [5] Lee W, Stolfo S J. A Framework for Constructing Features and Models for Intrusion Detection Systems[J]. ACM Trans on Inform and System Security, 2000, 3(4): 227 - 261.

(上接第 242 页)

- ity, LNCS, 3089/2004[C]. Berlin: Springer - Verlag, 2004. 426 - 438.
- [7] Bellovin S, Leech M, Taylor T. ICMP Traceback messages[Z]. IETF Internet Draft "draft - ietf - itrace - 04. txt, Work in progress, 2003.
- [8] Lee H C J, Thing V L L, Xu Yi, et al. ICMP Traceback with Cumulative Path, an Efficient Solution for IP Traceback, Information and Communications Security[A]. LNCS[C]. [s.

l. ]: Springer - Verlag, 2003. 124 - 135.

- [9] Thing V L L, Lee H C J, Sloman M, et al. Enhanced ICMP Traceback with Cumulative Path[A]. Proc 61st IEEE[C]. Washington: Vehicular Technology Society Press, 2005.
- [10] Postel J. Internet Protocol[Z]. Request for Comments 0791, Internet Engineering Task Force, 1981.
- [11] Elliot J. Distributed Denial of Service Attack and the Zombie Ant Effect[J]. IT Professional, 2000, 2(2): 55 - 57.