认证确定包标记算法

宗易安[1], 窦万峰[1,2]

(1. 南京师范大学 数学与计算机科学学院, 江苏 南京 210097;
2. 南京大学 计算机软件新技术国家重点实验室, 江苏 南京 210093)

[摘要] 确定包标记算法只需要边界路由器进行标记, 可以对只使用少量包的拒绝服务攻击进行追踪, 能同时追踪上千个攻击者, 并且易于实现. 针对确定包标记算法中, 被攻击者控制的路由器(边界路由器或中间路由器)修改标记或加入伪造包, 进而妨碍受害者重构入口地址的问题, 提出了新的基于 MAC认证的确定包标记算法. 研究表明, 认证确定包标记算法提供了足够的安全性, 能有效阻止子网内的攻击者或傀儡路由器伪造虚假的标记, 从而保证了受害者端地址重构的准确性.

[关键词] 拒绝服务攻击, IP追踪, 确定包标记, 基于 MAC的认证
[中图分类号] TP393.08 [文献标识码] B [文章编号] 1672-1292(2007)02-0067-05

Authenticated Deterministic Packet Marking Scheme

Zong Yian[1], Dou Wanfeng[1,2]

(1. School of Mathematics and Computer Science, Nanjing Normal University, Nanjing 210097, China;
2. National Important Laboratory of Computer Software New Technology, Nanjing University, Nanjing 210093, China)

Abstract Deterministic packet marking (DPM) algorithm only requires edge routers to perform packet marking and can trace a large number of attackers simultaneously with only a few packets from each attacker. For that, compromised routers, either edge routers or transit routers, can easily forge packet markings to prevent the victim performing reconstruction successfully. For that, a new scheme, namely MAC-based Authenticated DPM (ADPM), is proposed. Researches indicate that ADPM algorithm supplies sufficient security that attackers in subnets or compromised routers cannot forge markings, which assures the veracity of address reconstruction at the victim.

Key words denial of service attack, IP traceback, deterministic packet marking, MAC-based authentication

0 引言

随着 Internet的发展, 对 Internet的攻击也日渐增多. 一种经常被使用的攻击手段——分布式拒绝服务攻击(Distributed Denial of Service, DDoS), 利用相当数量的傀儡机产生巨大的流量来消耗网络设备的资源, 使正常的服务被降低或拒绝. 另一种拒绝服务攻击针对系统的漏洞, 只需发送少量的攻击包就能使系统瘫痪. 为了使攻击者对其行为负责, 在检测到攻击后, 定位攻击源显得尤为重要.

然而, 利用 IP协议的匿名性, 攻击者会在攻击包中使用错误的或虚假的 IP地址, 即使受害者收到攻击包, 也无法发现真实的攻击源. 定位真实攻击源的问题, 称作 IP追踪问题. 到目前为止, 已有多种 IP追踪技术被提出. 入口过滤[1], 边界路由器禁止非法源地址的包进入网络, 缺点在于需要所有的边界路由器都实现该功能. 受控洪泛[2], 洪泛受害者的各个上游链路, 通过观察受害者收到的攻击包的数量变化来确定攻击路径, 这本身就是 DoS攻击. 输入调试[3], 路由器允许管理员在出口上过滤符合攻击特征的包, 并确定这些包的入口. 这种方法需要网管介入, 耗费大量人工, 且需要多个 ISP之间的协调. 日志记录[3,4], 在关键路由器上对所有上游链路的数据包进行日志记录, 攻击发生时或之后, 受害者利用提取的攻击包的特

收稿日期: 2006-09-15.
基金项目: 江苏省高校自然科学基金(04KJD520106)资助项目.
作者简介: 宗易安(1981-), 女, 硕士研究生, 主要从事网络安全方面的学习与研究. E-mail: yianzong@163.com
通讯联系人: 窦万峰(1968-), 副教授, 博士后, 主要从事协同计算方面的教学与研究. E-mail: douwf@email.njnu.edu.cn

征与路由器中的日志信息比较,恢复出攻击包经过的路径.但是,日志信息占用路由器大量的系统资源,且需要大规模数据库和数据挖掘技术的支持. ICMP消息追踪[5],路由器以极低的概率生成和发送含有路由信息的追踪包.因为发送追踪包的概率很小(1/20 000),追踪包可能被攻击流淹没,受害者需要收到追踪包才能重构攻击路径.概率包标记[6,7],路由器以一定概率将路由的边信息写入 IP包头部,该算法需要知道网络的拓扑结构.

Belenky和 Ansari提出的确定包标记(Deterministic Packet Marking DPM)算法[8,9],优势在于只需要边界路由器进行标记,可以对只使用少量包的攻击进行追踪,能同时追踪上千个攻击者,并且易于实现,没有额外的带宽要求,不要求 ISP揭示其内部拓扑结构.但是,当多个攻击者使用具有相同源地址的包进行攻击时,或当攻击者发送的每一个包都使用不同的源地址时,基础 DPM算法[8](一般用 BDPM表示,而用 DPM表示这类算法)失效.于是, Belenky和 Ansari又改进了 BDPM算法[9],称之为基于 Hash的 DPM(Hash-based DPM, HDPM)算法.

以 HDPM为代表的 DPM算法仍有其局限性,归结如下: Hash冲突;数据包分片;路由器的负载;受控路由器.对于前三个问题,已有相关工作[10-12].而对被攻击者控制的路由器(边界路由器或中间路由器)修改标记或加入伪造包,进而妨碍受害者重构入口地址的问题还无人提及.本文引入认证机制,提出了认证确定包标记(Authenticated DPM, ADPM)方案,解决了该类问题.

1 BDPM算法和 HDPM算法

BDPM[8]的基本思想是:当 IP包从子网通过边界路由器进入网络时,边界路由器上与该子网相连的接口对 IP包标记.进行标记的接口称为 DPM接口, DPM接口的地址称为 IP包的入口地址.标记域是 IP包头部的 16位 ID域和 1位 RF域.标记包含 IP包部分入口地址.在 BDPM中,入口地址被分为两段,前 16位是段 0,后 16位是段 1. DPM接口每收到一个包,就随机选择入口地址的前 16位或后 16位填入 IP包的 ID域,包的 RF域写入对应的段号 0或 1.受害者维护一个以源地址为索引的表 IngressTbl.当受害者收到一个攻击包,首先检查该包的源地址是否存在表中,如果不存在,就创建这一项.然后将包内的部分入口地址填入表中对应的项.当受害者得到对应同一个源地址的入口地址的两个分段,就可获得入口地址.

文献[9]指出,当多个攻击者使用具有相同源地址的包进行攻击时, BDPM误检率较高.而当攻击者发送的每一个包都使用不同的源地址时, BDPM完全失效.因此,文献[2]提出了基于 Hash的 DPM(Hash-based DPM, HDPM)算法.在 HDPM中, IP包头部 16位 ID域和 1位 RF域,共 17位标记域被分成 3个部分: d位入口地址摘要域, a位入口地址段域, s位段号域.入口地址被分成 k段, $k=2^s$,并且每段含 a位.对于一个给定的 DPM接口,标记中的入口地址摘要部分是不变的.这样,受害者通过将地址段与摘要相关联就可重构入口地址. k个标记含有相同的入口地址摘要,不同的入口地址段及其对应的段号. DPM接口每收到一个包,就随机选择一个标记写入包头的 ID域和 RF域.

HDPM的重构过程分成两个独立的部分: (1)标记记录,每个可能出现的标记在重构表 RecTbl中能被惟一表示. (2)地址恢复,对有相同摘要的所有入口地址段的排列运用 Hash函数,若其值为摘要值,则该排列为有效入口地址,并将该有效地址移入入口地址表 IngressTbl.

然而,即使 HDPM使用的是理想的 Hash函数,如果攻击者的数目 $N > 2^d$,误检还是不可避免的.

2 基于 MAC的认证确定包标记

2.1 DPM算法的局限性

以 HDPM为代表的 DPM算法仍有其局限性.

2.1.1 Hash冲突

HDPM在标记内加入了入口地址的 Hash摘要,无法避免的 Hash冲突,正是造成 HDPM误检率的主要原因.文献[10]用 5个 Hash函数,产生 8种不同的标记,虽然重构的计算复杂度增加,但误检率明显降低.

2.1.2 数据包分片

虽然分片的情况在 0.25% ~ 0.5%之间[11],但确实存在.使用了 ID域的 DPM算法,可能导致分片在

目的端组装失败, 因为对同一个序列的分片使用不同的标记, 或对不同序列的分片使用相同的标记. 为此文献[11]提出分片持续 DPM, 实现了对属于同一个序列的分片使用相同标记的目的, 但未能解决对不同序列的分片使用相同标记的问题.

2.1.3 路由器的负载

包标记的过程会耗费边界路由器诸如内存、计算能力等资源. 当有大量包到达时, 路由器可能过载. 文献[12]使用了基于流的标记方案. 当路由器负载过高时, 路由器根据流信息有选择地标记包, 效果优于随机标记, 但记录流信息有一定空间耗费.

2.1.4 受控路由器

DPM 的特点是所有包在进入网络前被边界路由器标记, 这使得任何攻击者伪造的标记都将在这个包经过的第一个路由器被正确的标记覆盖. 因此, 文献[8, 9]声称 DPM 阻止了虚假标记. 本文认为 DPM 只是阻止了攻击者在子网内欺骗, 却对被攻击者控制的边界路由器和中间路由器伪造的标记无能为力. 一个受控的路由器 (边界或中间路由器), 可以轻易伪造其它正常边界路由器的标记, 从而阻碍受害者正确恢复地址.

为了解决这个问题, 需对包标记进行认证. 数字签名是认证的有效手段, 但签名和验证的代价太高, 带宽耗费较多. 与之相比, 基于消息认证码 (Message Authentication Code, MAC)的认证更为经济. 因此, 本文在 HDPM 的基础上, 引入认证机制, 提出了基于 MAC的认证 DPM (Authenticated DPM, ADPM).

2.2 基于 MAC的标记认证

MAC常用于双方消息认证, 假定通信双方 A 和 B 共享一个密钥 K, MAC函数 C. 当 A 向 B 发送消息时, A 计算 MAC= $C_K(M)$, 将 MAC与消息 M 一起发送给 B. B 对收到的消息用相同的密钥进行相同的计算得出新的 MAC, 与接收到的 MAC比较. 若两者相等则消息未被篡改, 且来自真正的发送者 A. MAC的计算效率很高, 一个快速工作站每秒可计算大约 300000个 8位的 HMAC-MD5.

假定 DPM 接口 I 与受害者共享一个密钥 K. 令 F 表示 MAC函数 (F 是公开的), F_K 表示以 K 为密钥的 MAC函数. I 对包的源地址 SA、目的地址 DA 以及自己的地址 IA 计算 MAC= $F_K(SA, DA, IA)$, 并将 MAC值附加在标记之后. 受害者用共享密钥验证标记的真伪. 因为受控路由器不知道其它正常的边界路由器 DPM 接口的密钥, 也就无法伪造其它正常边界路由器 DPM 接口的标记. 在这种情况下, 受控路由器仍可以使用相同的源地址来伪造标记, 但是受害者可以阻止来自这些源地址的包来避免攻击.

2.3 随时间变化的密钥链

2.2节假定各个 DPM 接口与每个可能的受害者共享一个密钥, 这是不实际的. 因为标记认证本质上要求非对称性 (所有可能的接收者能验证标记真伪, 但不能制造经过认证的标记), 所以采用密钥延迟公开的方法来实现这种非对称性, ADPM 使用随时间变化的密钥计算 MAC. 类似的方法在文献[7]中用于概率包标记的标记认证. 由于 ADPM 要求受害者和路由器近似时序同步, 我们将在 2.5节讨论近似时序同步的实现.

DPM 接口 I 生成密钥链 $\{K_m\}$, 方法是对随机选择的种子 K_n 运用单向 Hash函数 g, 得到一个密钥链 $K_m = g(K_{m+1})$. 因为 g 是单向函数, 可以向前计算 (时间上向后), 但不能向后计算 (时间上向前), 如图 1所示. 例如, 给定 K_{m+1}, 可以计算出 $K_0, K_1, \cdots K_m$, 但给定 $K_0, K_1, \cdots K_m$, 不能计算出 K_{m+1}.

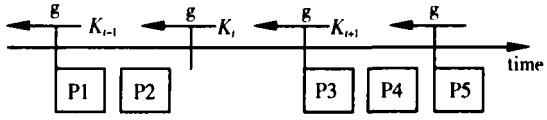

图 1 随时间变化的密钥链

Fig.1 Time-released key chains

路由器通过一个标准承诺协议, 为其每个 DPM 接口的密钥链作承诺. 例如, 用其私钥对密钥链的第一个密钥 K_0 签名 (假定各路由器有一个被证明的公钥), 并在网站上公布承诺.

将时间分片, DPM 接口 I将时间片与其密钥链对应, 每个时间片对应密钥链中一个密钥. 在时间片 t, 接口 I 使用当前时间片的密钥 K_t 来计算 MAC, 标记该时间片内到达的包 (如图 1所示).

接口 I 将在时间片 t 结束时的一个延迟 δ 后公布密钥 K_t. 密钥公布延迟 δ 要大于 Internet任何合理的往返时间加上路由器与受害者之间的同步误差. 密钥可以公布在网站上.

当受害者收到标记包时, 保存包的到达时间. 在重构入口地址前, 受害者下载 DPM 接口已公开的密

钥. 因为使用密钥链, 受害者仅需下载接口 I 最近时间片的密钥, 然后就能计算出所有之前的密钥. 对于各个标记的包, 受害者首先用其到达的时间确定其发送时间. 假定包的到达时间 T_a, 路由器与受害者之间的同步误差是 $\pm\delta_1$, 包的最大传输延时是 δ_2, 则包实际发送时间 T_s 满足 $T_a - \delta_1 - \delta_2 < T_s < T_a + \delta_1$. 因此, 如果时间片长度远大于 $2\delta_1 + \delta_2$, 受害者能以较高的准确率确定标记包的发送时间. 在确定包的发送时间之后, 受害者能使用对应的密钥计算包的 MAC.

2.4 认证确定包标记算法

认证确定包标记(ADPM)算法, 仍使用 IP包头的 16位 ID域和 1位 RF域作为标记域. 17位标记域被分成 3个部分: d 位 MAC域, a 位入口地址段域, s 位段号域. 入口地址被分成 k 段, $k = 2^s$, 并且每段含 a 位. DPM 接口 I 每收到一个 IP包, 用当前时间片对应的密钥 K_t, 对包的源地址 SA、目的地址 DA 以及自己的地址 IA 计算 $MAC = F_K(SA, DA, IA)$, 并与随机选择的入口地址段及其对应段号, 一并写入 IP包头的标记域.

ADPM的重构分为两步: 标记记录和地址恢复.

标记记录: 受害者维护一个重构表 RecTbl, 该表包含到达时间、源地址、MAC、入口地址段、段号、发送时间片共 6个域. 受害者每收到一个 IP包, 就将包的信息填入 RecTbl中对应的域(发送时间片域为空).

地址恢复: 入口地址表 IngressTbl仅含入口地址一个域, 用来存放有效的入口地址. 受害者首先根据 RecTbl中各条记录的到达时间域, 推断出相应的发送时间片, 写入发送时间片域. 然后, 将 RecTbl中所有记录按照相同的发送时间片、源地址、MAC进行划分. 对具有相同的发送时间片、源地址、MAC的记录集合, 得到其密钥, 并对该记录集合中每一个可能的入口地址段的排列, 与源地址和受害者地址一起计算得到新的 MAC, 若新的 MAC与该记录集合的 MAC相同, 则该入口地址段排列为有效入口地址, 将这个有效入口地址写入 IngressTbl.

认证确定包标记算法如下:

```
Let F be the Hash function to compute MAC
Let k be the number of fragments of ingress address
Marking procedure at router R, interface I
    Let IA be the address of interface I
    For each packet w
        Let n be a random integer from [0, k - 1]
        Let f be the fragment of IA at fragment number n
        Let K be the current key
        Let SA be the source address of w
        Let DA be the destination address of w
        w. frag = f
        w. num = n
        w. MAC = F_K (SA, DA, IA)
Recording procedure at victim V:
        Let RecTbl be a table of tuples
(arriving time, source address, MAC, fragment number, sending time interval)
    For each packet w from attacker
        Let T be the arriving time of w
        Let SA be the source address of w
        RecTbl Insert (T, SA, w. MAC, w. frag w. num, NULL)
Recovery procedure at victim V:
    Let IngressTbl be a table of ingress address
    For each record r in RecTbl
        Deduce r. sending-time-interval according to r. arriving-time
    Partition all records in RecTbl with arriving time, source address and MAC
    Let rcd be a record set with same arriving time, source address and MAC
    For each rcd in RecTbl
        Get key K
        For each possible ingress address combination IA in rad
            If F_K (SA, DA, IA) = rcd. MAC
            IngressTbl Insert(IA)
```

2.5 近似时序同步

近似时序同步是 ADPM 的重要前提. 虽然存在精确的时序同步协议, 但管理难度大, 实现复杂, 且有许多功能 ADPM 并不需要, 需要寻求一种简单而安全的近似时序同步协议. ADPM 只要求接收者知道自己与发送者之间的时间差的上限 δ. 如果发送者与接收者之间的真实时间误差是 δ, 对于近似时序同步只要求保证 $\delta' \geqslant \delta$ 文献[13]的直接时序同步和间接时序同步协议恰能满足要求.

近似时序同步协议, 首先由接收方向发送方(边界路由器)发送同步请求, 并记录本地发送时间 t_R, 发送方收到同步请求立即记录本地接收时间 t_S, 并发送包含 t_S 的响应. 接收方收到响应, 计算 $\delta' = t_S - t_R$, 显然 $\delta' \geqslant \delta$ 间接时序同步协议, 接收方和发送方分别与同一个参考时间同步, 将接收方和发送方各自与参考时间的测量误差和作为 δ, 显然这个值大于或等于接收方和发送方之间的真实时间差 δ

3 结 语

认证确定包标记算法引入了 MAC 认证机制, 为 HDPM 算法提供了足够的安全性, 不仅阻止了子网内的虚假标记, 而且能有效阻止受控的路由器(边界或中间路由器)伪造其它正常边界路由器的标记, 从而保证受害者端不被虚假标记误导, 提高了地址恢复的准确性. 同时, 近似的时序同步易于实现; 边界路由器为其每个接口计算密钥链, 虽然在一定程度上增加了边界路由器的负载, 但仍在可接受的范围之内. 遗憾的是, ADPM 算法仍存在不足之处, 与其它 IP 追踪算法一样, 即使在最好情况下也只能追踪到攻击者所在的子网, 而无法定位攻击源. 这个问题单靠 ADPM 算法的发展可能是无法解决的, ADPM 算法与其它追踪技术的结合, 可能是今后 IP 追踪技术发展的方向.

[参考文献] (References)

[1] Ferguson Paul, Senie Daniel. RFC 2827: Network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing[EB/OL]. [2006-09-15]. http://www.ietf.org/rfc/rfc2827.txt may 2000

[2] Burch Hal, Cheswick Bill. Tracing anonymous packets to their approximate source[C]//USENIX LISA. New Orleans: Remy Evard, 2000: 319-327.

[3] Stone Robert. Centertrack: An ip overlay network for tracking dos floods[C]//9th USENIX Security Symposium. Denver, Colorado: Steven Bellovin, 2000: 199-212

[4] Sager Glenn. Security fun with ocxmon and cflowd[R/OL]. [2006-09-15]. Presentation at the Internet 2 Working Group, 1998. http://www.caida.org/. projects/NGI/content/security/1198

[5] Bellovin Steve. The icmp traceback message: draft-bellovin-itrace-00.txt[EB/OL]. [2006-09-15]. http://www.cs.columbia.edu/~smb/papers/draft-bellovin-itrace-00.txt

[6] Savage Stefan, Wetherall David, Karlin Anna, et al. Network support for IP traceback[J]. IEEE/ACM Transactions on Networking, 2001, 9(3): 226-237.

[7] Song Dawn, Perrig Adrian. Advanced and authenticated marking schemes for IP traceback[C]//IEEE INFOCOM'01. Anchorage, Alaska: Bhaskar Sengupta, 2001: 878-886.

[8] Belenky Andrey, Ansari Niwan. IP traceback with deterministic packet marking[J]. IEEE Communications Letters, 2003, 7(4): 162-164

[9] Belenky Andrey, Ansari Niwan. Tracing Multiple Attackers with Deterministic Packet Marking(DPM)[C]//IEEE PACRIM'03. Victoria, Canada: Fayez Gebali, 2003: 49-52

[10] Belenky Andrey, Ansari Niwan. Accommodating fragmentation in deterministic packet marking for IP traceback[C]//IEEE GLOBECOM'03. San Francisco, USA: Terry E F Keen, 2003: 1374-1378.

[11] Lee Tsern-Huei, Huang Tze-Yau William, Lin Iven. A deterministic packet marking scheme for tracing multiple internet attackers[C]//IEEE ICC'05. Seoul, Korea: Yong-Kyung Lee, 2005: 850-854.

[12] Xiang Yang, Zhou Wanlei. A defense system against DDoS attacks by large-scale IP traceback[C]//ICITA'05. Sydney, Australia: Sean He, 2005: 431-436.

[13] Perrig Adrian, Canetti Ran, Song Dawn, et al. Efficient and secure source authentication for multicast[C]//ISOC NDSS'01. San Diego, California: Terry Weigler, 2001: 35-46.

[责任编辑: 严海琳]