一种基于路由器的伪装IP回溯方法

黄克军[1] 王 钰[1] 刘 梅[2] 李毅超[1]
(电子科技大学计算机学院 成都610054)[1]
(曲靖市供电有限责任公司信息中心 云南曲靖655000)[2]

摘 要 在因特网中,攻击者能伪造其IP地址。本文提出了一种新的针对洪泛型攻击(flooding attack)的IP回溯技术——PPL。在该方法中,路由器以一定的概率存储转发分组的信息,然后利用这些信息,从被攻击者开始回溯攻击分组到其源,对网络攻击者起到威慑的作用。仿真实验证明,新方案在回溯洪泛型攻击时,其性能优于文[2]中的方案。

关键词 分布式管理,网络安全,IP回溯,概率分组日志,拒绝服务

A Router-Based Method of IP Traceback

HUANG Ke-Jun[1] WANG Yu[1] LIU Mei[2] LI Yi-Chao[1]
(College of Computer Science and Engineering, UESTC of China ,Chengdu 610054) [1]
(Information Center of Qujing Power Supply Co. Ltd., Qujing 655000)[2]

Abstract This paper descriebes a technique for tracing anonymous packet flooding attacks in the Internet back toward their source. In our approach, forwarding nodes (such as routers) log information with probabilistic about traversing packets on the Internet and then use the log data to trace each packet from its final destination back to its source, hop by hop. The simulation test proves that the performance of PPL is better than method in literature [2].

Keywords Distributed manegement, Network security, IP traceback, Probabilistc packet logging, Denial-of-service

1 引言

IP协议简单有效,却难以保证包含在分组源地址段中的地址就是该分组的源发地址。IP协议本身无法解决这样的安全问题,结果攻击者可在因特网的任何位置发送伪造了其源地址的攻击分组。

DOS(Denial Of Service,拒绝服务)攻击消耗远程主机或网络的资源,使得主机或网络对合法用户的服务被拒绝或降级。这类攻击是最难解决的安全问题之一,因为它们容易实现,难以阻止和回溯。其中,洪泛型攻击是本文研究的对象,它不同于其它类型的DOS攻击,它要对被攻击者有效,须不间断、快速发送分组。

基于网络的攻击检测系统可能检测出一个攻击,通常却不能指出攻击分组来源于何处。可靠地回溯分组到其源,是控制攻击者的第一步,也是重要的一步[6]。在构造回溯系统时,有一些重要问题需考虑,如:回溯哪一个分组,如何保持隐私性(不影响合法用户的隐私)和开销最小化(路由器用在回溯上的时间和用于存储信息的空间开销)。

本文要解决的问题是:回溯伪装分组流。

S. Savage等提出了一种概率分组标记方案(PPM, Probabilistic Packet Marking)[1]来对付洪泛型DOS攻击。在这种方案中,当分组在网络中传输时,它以一定的概率存储其所经过的每一个路由器的信息,即分组被这些信息标记。标记分组的接收方可利用分组中的路由信息构造出到分组源的路径。路由器在标记分组的同时,不影响正常分组转发过程。

PPM的优点是:对网络和路由器引入的开销比较低、支持逐增配置(incremental deployment);其缺点是:编码和重构路径的计算量很大,对付DDOS(分布式拒绝服务攻击)时,误报率很高[3]。

最近Tatsuya Baba等提出了一种IP回溯的方法[2]。因特网的中间结点记录所转发的分组信息,在需要时可利用这些信息回溯一个分组从终点到源的路径。该方法还采用了一种分布式管理结构来实现跨越不同网络的回溯。

与PPM相比较,这种方法计算量低且容易实现,而且AMN的引入能解决在不同网络间回溯的问题。但在对付洪泛型DOS时有一些缺陷。路由器要记录所转发的所有分组的信息,这需要大量的存储空间,而且数据容易被覆盖。但实际上,只要大分组流中有一个分组驻留在路由器中用于回溯就可以了。据此,我们提出了概率分组日志方案:路由器以一定的概率来记录分组信息。我们的仿真实验显示,存储量的增长率比文[2]中低得多。这意味着分组的驻留时间延长了。

2 新方案PPL

下面将描述PPL方法的原理。图1是以受害方V为根的一颗树,V可以是受攻击的单个主机或者一个网络边缘设备如防火墙、攻击检测系统等。A是潜在的攻击源,是树叶,路由器R是A到V路径上的中间结点。攻击路径由A到V间的路由器序列唯一标识。例如:产生于A2的攻击要到达V必须经过的路径为(R6,R3,R2,R1),如图中粗线所示。

来自A2的攻击分组被R6转发,R6以概率P来记录分组的输入接口和源IP地址。当V的攻击检测系统探测到攻击分组,V发送包含攻击分组信息的消息给上游路由器R1,R1

黄克军 讲师,主要研究方向:计算机网络及信息安全;王 钰 硕士研究生,主要研究方向:计算机网络及信息安全;刘 梅 工程师,主要研究方向:计算机网络及信息安全;李毅超 副教授,主要研究方向:计算机网络及信息安全。

在其缓冲器中搜索与分组信息相匹配的分组，从而找到攻击分组的输入接口，知其上游路由器为 R2。R1将结果返回给 V，V 则发送一个消息给 R2，R2搜索其缓冲区得知攻击分组来自 R3，R2将结果返给 V，V 又向 R3发送回溯消息。这个过程不断地在上游路由器中重复直到找到攻击源。最后，V 便知道攻击路径为(R6，R3，R2，R1)。

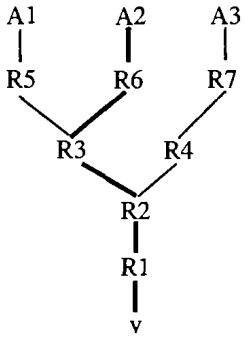

图1　从攻击的受害者所看到的网络

我们的方案中有这样一些基本假设：①攻击者能够产生任何分组；②攻击者可能意识到正在被回溯；③分组可能丢失或重新排序；④攻击者可发送大量的分组；⑤攻击者和受害方间的路由器相当稳定；⑥路由器都有 CPU 和内存两方面的限制；⑦路由器未广泛受侵。

前三个假设是对当今攻击者和网络的局限性的保守评价。在因特网环境中设计一个回溯系统是极具挑战的。几乎没有可信任的实体。特别是攻击者有能力产生任意分组，这一情况限制了许多潜在的解决方法。当路由器接受到一个分组，它无法识别这个分组是被上游路由器还是被攻击者伪装了。事实上，我们唯一可依赖的是：分组从攻击者到受害者所经过的路由器序列。

后面四个假设反映了我们设计的基础。洪泛型 DOS 攻击由成千上百万的分组构成。它们侵占了受害方的资源。我们的方案正是依赖这个特性。路由器以一概率 P 来转发分组信息以保证有少量攻击分组被存储在路由器中用于回溯。该假设不支持单包引发的攻击。

下面是路由器记录分组的算法

```
logging procedure at router R:
  For each packet w
    Let x be a random number from[0..1)
      If x<p then
      Log information about w
```

3　体系结构和回溯过程

3.1　体系结构

记录分组、攻击检测、回溯处理和构造攻击路径的功能被分配给 PPL 系统的独立模块中。图2中表明了体系结构中的三个模块：tracer、sensor、manager。每个附加 PPL 功能的路由器有个 tracer 模块。根据所考虑的路由器的类型，tracer 可作为软件代理实现和配置或做成一接口卡插入连接总线上或者一独立的附加盒通过附加接口连到路由器上。

Sensor 配置在每个目标网络，监视网络中的分组。当 sensor 检测到一个攻击，它会发送一个回溯请求给 manager。manager 响应 sensor 请求，控制 tracer 并管理整个回溯过程。tracer 保持被转发分组的日志记录，比较回溯分组和日志数据，从而发现回溯路径。

因特网之大，集中管理整个回溯过程和回溯信息是不可能的。而且，配置不同接入策略的网络，没有任何限制地回溯来自其它网络的分组，这是很困难的。文[2]引入了一种分布式的管理方法 AMN(autonomous management network，自治网)在一组网络中控制回溯过程和回溯信息。我们的方法中也采用了 AMN 的思想。

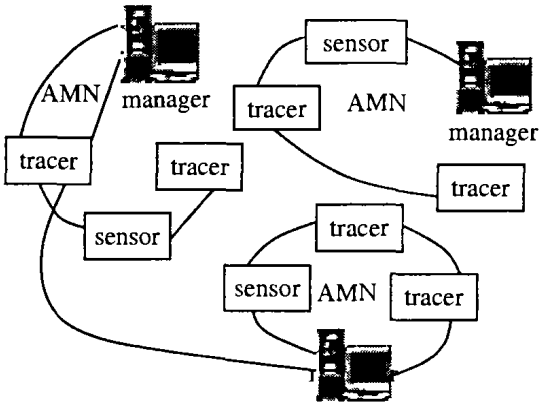

图2　本文的体系结构

3.2　回溯过程

当 sensor 检测到一个攻击，它产生包含攻击分组特征的数据并向配置在 AMN 中的 manager 发出回溯请求。manager 命令 AMN 中的 tracer 对分组进行回溯。基于返回的结果，上述过程会一直继续下去直到 tracer 识别到攻击源。如果回溯过程超出了 AMN 的范围，回溯过程被传递给相邻 AMN 的 manager。每个 AMN 中的 manager 在其 AMN 的范围内进行回溯，把结果返回给发起回溯请求的 manager，该 manager 将最终结果即攻击源返回给提请回溯的 sensor。

3.3　仿真实验结果

在我们的仿真实验中，攻击者以100bit/s 的速率发送分组，链路的速率是9600bit/s。图4是仿真实验模型，图3显示了转发路由器在 p=0和 p=0.5两种情况下缓冲器的增长率。表1显示的是不同跳数下，PPL 所需的回溯时间。

我们看到当 P=0.5时，存储量的增长率比 P=1时低得多。我们能很容易得出结论：P 值越小，存储量的增长率越低。而文[2]中的方案，其 P 值为1，对存储的要求很高，易引起分组丢失。

表1　回溯时间

跳数	2	3	4
回溯时间	1.531	1.804	2.710

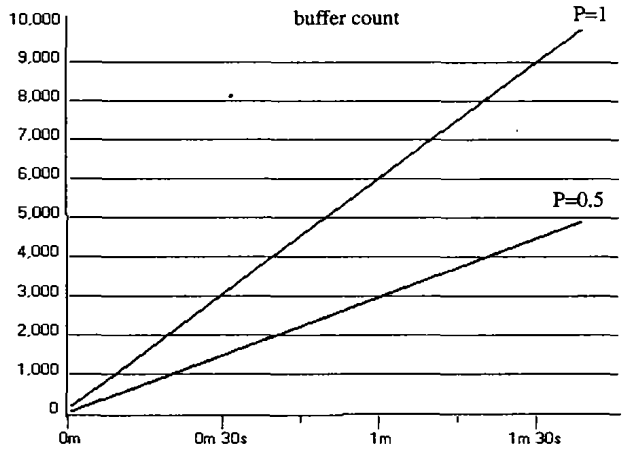

图3　当 P=1和 P=0.5时缓存的变化

(下转第112页)

Intl. Conf. on System Sciences - 2001
2 1SoftwareInc. Webtrends. http://www.webtrends.com,1995
3 OpenMarketInc. OpenmarketWebreporter. http://www.openmarket.com,1996
4 NetGenesisCorp. Netanalysisdesktop. http://www.netgen.com, 1996
5 Chen M S, Park J S, Yu P S. Data mining For Path traversal patterns in a Web environment. In: Proc. of the 16th Int'l Conf. on Distributed Computing Systems. HongKong, 1996. 385～392
6 Zaiane O R, Xin M, Han J. Discovering Web access patterns and trends by applying OLAP and datamining technology on Weblogs. In: Proc. of Advances in Digital Libraries Conf. Santa Barbara, CA, 1998. 19～29
7 Cooley R, Mobasher B, Srivastava J. Grouping Web page references into transactions formining World Wide Web browsing patterns. Department of Computer Science, University of Minnesota: [Tech Rep: TR97-021]. 1997
8 Mobasher B, Jian N, Han E, et al. Web mining: Pattern discovery from World Wide Web transactions. Department of Computer Science, University of Minnesota: [Tech Rep: TR96-050]. 1996
9 Berendt B, Spiliopoulou M. Analysis of navigation behaviour in web sites integrating multiple information systems. The VLDB Journal, 2000, 9: 56～75
10 Giannotti F, Gozzi C. Characterizing Web User Accesses: A Transactional Approach to Web Log Clustering. In: Proc. of the Intl. Conf. on Information Technology: Coding and Computing (ITCC. 02)
11 Adamic L A, Bernardo. The Nature of Market in the World Wide Web, Xeror Research Certer, May 1999
12 Zhong Su1, Qiang Yang. Correlation-Based Web Document Clustering for Adaptive Web Interface Design. Knowledge and Information Systems, 2002, 4: 151～167
13 Ozmutlu H C, Amanda Spink B. Analysis of large data logs: an application of Poisson sampling on excite web queries. Information Processing and Management, 2002, 38: 473～490
14 Tauscher L, Greenberg S. Revisitation patterns in world wide web navigation. In: Proc. of CHI97, 1997. 399～406
15 Ester M, Kriegal H P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis E, Han J, Fayyad UM, eds. Proc. of thesecond intl. conf. on knowledge discovery and data mining. AAAI Press, Menlo Park, CA, 1996. 226～231
16 Ester M, Kriegal H P, Sander J, Wimmer M, Xu X. Incremental clustering for mining in a data warehousing environment. In: Gupta A, Shmueli O, Widom J, eds. Proc. of 24rd intl. conf. on very large data bases, New York. Morgan Kaufmann, San Mateo, CA, 1998. 323～333
17 Savasere A, Omiecinski E, Navathe S. An efficient algorithm for mining association rules in large databases. In: Dayal U, Gray PMD, Nishio S, eds. Proc. of 21st intl. conf. on very large data bases, Zurich, Switzerland. Morgan Kaufmann, San Mateo, CA, 1995. 432～444
18 Bonchi F, Giannotti F. Web log data warehousing and mining for intelligent web caching. Data & Knowledge Engineering, 2001, 39: 165～189

(上接第68页)

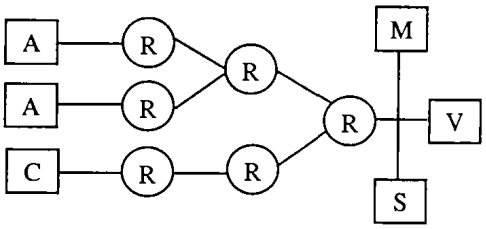

图4 实验网络模型

结论 PPL 对网络和路由器产生的开销都比较低，而且支持逐增配置。例如：在现实环境中，引入这些组件到一个管理域，如一个企业内联网，首先在域中实现回溯功能。如果相邻域也配置了回溯系统，域之间通过其 manager 交换回溯信息来实现跨网络的回溯。

通过我们的仿真研究，可以很明显地看到，路由器上存储量的增长率降低了，同时又没有降低逐跳(hop by hop)回溯的性能。

PPL 对当今协议没有改变，克服了 PPM 方案的一些问题。在 PPM 方案中，路由器必须放一些信息到分组中，这样会产生大量计算并增加了算法的复杂性，而且需对 IP 协议进行一些修改。

回溯所需要的时间与分组在缓冲器中的驻留时间有直接关系。当概率 P 值较小时，分组的驻留时间就会变长，被回溯分组的丢失率就会降低。

PPL 对骨干网更有意义，因为骨干网上的通信量非常高，更有必要采取有效的方法来降低对存储空间的要求。

但是，我们的方法也有一些弱点限制了它在一些情况下的可用性和有效性。用这种方法回溯由单包引发的攻击[6]很困难甚至不可能。该方法是否可有效地用于大规模 DDOS 攻击，还有待于验证。如何决定概率 P 的值也是一个有待于在实践中解决的问题。

参 考 文 献

1 Savage S, et al. Practical Network Support for IP Traceback. In: Proc. 2000 ACM SIGCOMM, ACM Press, New York, Aug. 2000. 295-306. Aavailable online at: http:// www. cs. washington-edu/ homes/ savage/ traceback. html.
2 Baba T, Matsuda S. Tracing Network Attacks to Their Sources. IEEE Internet Computing, 2002. 20～26
3 Song D X, Perrig A. Advanced and Authenticated Marking Schemes for IP Traceback. IEEE INFOCOM 2001
4 Bellovin S M. Security Problems in the TCP/IP Protocol Suite. Computer Communications Review, 1989, 9(2): 32～48
5 Microsoft Corporation. Stop 0A in tcpip. sys when receiving out of band (OOB) data. Available Online at: http://support. microsoft. com/sup-port/kb/articles/Q143/4/78. asp
6 Snoeren A C, et al. Single-Packet IP Traceback. IEEE/ACM Transactions on Networking, 2002, 10(6): 721～734