正态分布与分布式拒绝服务攻击的主动预防

赵 英,倪 铮
(北京化工大学 网络中心,北京 100029)

摘 要:随着信息技术的发展和应用的普及,网络安全问题已经成为人们关注的焦点问题。目前分布式拒绝服务(DDoS, Distributed Denial of Service)攻击已经成为影响 Internet 正常运行的一个比较严重的问题,并影响合法用户获得正常的服务。文中首先阐述了 DDoS 形成的原理,然后分析了预防 DDoS 攻击的措施和机制。随后借助于 SSFNet(Scalable Simulation Framework Net)仿真软件构建相应的网络环境,模拟了一种分布式拒绝服务攻击。针对在实验中发现的攻击特征,即攻击发生时通过路由器的新 IP 数量呈现正态分布的变化趋势,结合统计学中正态分布的概率理论知识,提出了一种通过正态分布模型结合网络中新 IP 数量变化趋势应对分布式拒绝服务攻击的主动防御方案。然后利用实验中采集的数据,对所提出的应对分布式拒绝服务攻击的防御方案进行了验证。

关键词:DDoS;正态分布;主动;安全
中图分类号:TP393.08 **文献标识码**:A **文章编号**:1673-629X(2006)12-0237-03

Normal Distribution and Active Denial-of-Service Defense Mechanisms

ZHAO Ying, NI Zheng
(Network Center, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: As the development and prevalence of information technology, network security is currently a hot issue in the Internet. The DDoS is becoming a serious problem to affect the running of the Internet by preventing legitimate users of a service from using the desired resource. The theory of DDoS is introduced at first in this paper. Then discuss the countermeasures and mechanism of DDoS. Relying on SSFNet simulation software, sets up corresponding network environment, and simulates a DDoS attack. Whereafter, regarding remarkable characteristics of the attack found in the experiment, namely the normal distribution trend which is presented by the number of new IP addresses passing the router, and combining with normal distribution probability theory in statistics, it comes up with an active defense scheme against DDoS attack. Subsequently, by taking advantage of collected data and integrating with abnormity detecting method raised from the scheme, it tests the validity of DDoS defense scheme.

Key words: DDoS; normal distribution; active; security

0 概 述

随着 Internet 的普及,联网计算机的数量迅速增加,网络入侵问题也随之突出[1]。目前分布式拒绝服务攻击(DDoS)已经成为威胁网络正常运行的黑客软件之一。是目前黑客经常采用而难以防范的攻击手段。尽管 DDoS 的攻击方式有很多种,但最基本的 DDoS 攻击就是利用合理的服务请求来占用过多的服务资源,从而使合法用户无法得到正常的服务和资源。这些问题的出现不仅影响了网络的正常运行,而且也阻碍了信息技术的应用和普及。

1 DDoS 工作原理

由于 Internet 本身的开放性和任意性,使其在近 30 多年里得到了迅猛发展。然而这些特性所带来的安全问题同时也给网络的攻击者提供了可乘之机。这里首先介绍一下 DDoS 的工作原理。

尽管 DDoS 的攻击方式有很多种,但其基本思想就是想通过合理的服务请求,恶意占用被攻击对象的过多资源。一个最简单的 DDoS 攻击过程如图 1 所示。

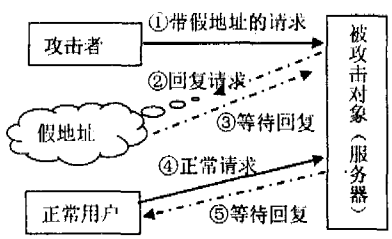

图 1 简单的 DDoS 结构

首先攻击者向服务器发送大量的带假地址的请求①,然后服务器对这些请求进行回复②,并等待请求发送者的回复③。由于请求的地址是一系列的假地址,这就造成了服务器永远不会等到回复③。由于服务器在等待过程中

收稿日期:2006-03-09
基金项目:国家"十五"资助项目("十五"211CERS-08)
作者简介:赵 英(1966-),男,天津人,教授,研究方向为计算机网络、分布式系统。

没有释放资源,如果这种请求过多的话,就会造成服务器资源被耗尽。如果此时有一个正常的用户向服务器发送正常请求④的话,由于此时服务器无可用的资源,则不会对用户的请求发送回复⑤。这也就是说正常的用户从服务器上不能够得到正常的服务。这也正是拒绝服务攻击的由来。

2 主动 DDoS 的防范

一般认为,传统的 DDoS 防范策略采用"保护－监测－应对"这个循环过程。通常情况下,DDoS 攻击开始后,目标网络和主机已经受到了相当程度的伤害。如网络的拥塞、服务器不能工作等。通常这种防范策略被叫做被动的防范策略。鉴于被动 DDoS 防范策略存在着许多问题,人们开始思考如何更早地发现 DDoS 的出现,并采取相应的应对措施,即主动 DDoS 防御系统[2]。主动防御系统的目标是尽早地控制住攻击,并减少所带来的破坏[3]。具体的内容包括:

* 在分布式拒绝服务攻击的部署阶段就试图找到恶意的行为。

* 在同级的防御系统间通讯,以共享攻击的信息。

* 记录恶意行为的线索以备日后分析、学习和讨论。

* 构建一个可以升级的网络资源保护体系,以备日后部署分布式的安全系统。

一般来说,主动 DDoS 防范系统是一个分布式系统。系统中的节点主要分布在要监控网络的关键路由器、防火墙或系统网关上。这样当网络出现异常时,这些设备能够及时捕捉到这些异常信息,并可以采取相应的应对方案[4]。

3 基于正态分布概率统计分析的防御措施

为了探测由于 DDoS 攻击引发的网络异常的方法,人们提出了众多的理论方案。这其中基于统计的分析方法[5]在文中被采用,主要是因为它的获取和处理过程相对迅速,随着统计数据量的增加,所获得的预测结果也越来越准确。该方法其理论依据是,网络异常遵循一种正态分布,通过统计学的概率模型来探测网络异常现象。正态分布的实验频率曲线有以下特征:曲线的纵坐标值为非负值;观测值在平均值附近出现的机会最多,所以曲线存在一个高峰;大小相等、符号相反的偏差发生的频率大致相等,所以曲线有一中心对称轴;曲线两端向左、右延伸逐渐趋近于零,这表明特大正偏差和特大负偏差发生的概率极小,一般很少出现;在对称轴两边曲线上,各有一个拐点,具有这 5 个特征的曲线,并且要求该曲线下的总面积等于 1,即符合理论频率曲线的要求。

如果正常情况下的网络用 $x(t)$ 表示,异常情况下为 $y(t)$,那么 $y(t)=x(t)+n(t)$,这里的 $n(t)$ 就表示网络的异常,如果网络异常超过设定的合理范围,那么就表示网络中存在某些异常现象。根据正态分布的特性:

$$\zeta \sim N(\mu_\zeta,\sigma_\zeta^2)=\frac{1}{\sqrt{2\pi}\sigma_\zeta}e^{-\frac{(\zeta-\mu_\zeta)^2}{2\sigma_\zeta^2}}$$

可以得到如下的可能性分析结果:

$$P_d=\int_{\frac{V-\mu_\zeta}{\sigma}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt \quad P_f=\int_{\frac{V}{\sigma}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$P_m=\int_{-\infty}^{\frac{V-\mu_\zeta}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

P_d 表示成功探测可能性,P_f 表示把正常网络误认为存在异常的可能性,P_m 表示没有探测到异常的可能性。ξ 表示网络异常,V 表示设定的阈值,μ_ξ 表示 ξ 均值,σ 表示 ξ 的方差。然后令 $\phi(t)=\int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$,可以得到如下结论:

$$P_d=1-\varphi[(V-\mu_\xi)/\sigma] \quad P_f=1-\varphi[V/\sigma]$$

$$P_m=\varphi[(V-\mu_\xi)/\sigma]$$

假设 f 为错误探测可能性,令 $P_f<f$,则有

$$V_f\geqslant-\sigma\varphi^{-1}(f)$$

假设 d 为成功探测可能性,令 $P_d>d$,则有

$$V_d\leqslant\mu_\xi-\sigma\varphi^{-1}(d)$$

为了获取最好的效果,令 $f=0$ 且 $d=1$,则有

$$V\in[4\sigma,\mu_\xi-4\sigma],\mu_\xi-4\sigma>0$$

所以通过设定合适的阈值 V,就有可能获取最大的异常探测可能性和最低的错误探测可能性。

4 实验方案及实施

由于组建一个真实的分布式测试环境需要很高的代价,所以通过有效的网络模拟软件 SSFNet 来组建这样一个模拟实验环境是很实际的。

具体的网络模拟环境如图 2。它由 3 个子网组成:其中 Net0 表示一般的客户端,Net1 表示被攻击服务器(Server)所在的子网,Net2 被定义成发起 DDoS 的客户端。在这个实验环境中,存在着一个中心路由器,其负责连接这 3 个子网,而每个子网都有一台边界路由器。

实验中用到 SSFNet 中的两类操作:DDoS 和 TcpDump。DDoS 主要用于发起 DDoS 攻击,具体采用的是基于 TCP 的 SYN 溢出攻击;TcpDump 主要用于捕捉混合模式下所经过的数据包,以用于分析攻击时的网络特征。

在实验中,攻击发起后(攻击时间为 time = 100s 至 time=400s),客户机不断地向目标服务器发送具有 IP 欺骗的 SYN 包,使其无法确认发来的连接请求,随着这些无法成功建立连接的假请求包的增多,不断消耗自身的资源,当服务器的资源消耗殆尽时,便会无法响应其他用户的连接请求,产生拒绝服务的现象。

通过对每个 Router 的 IP 包进行检查,发现在攻击发生的这个时间段内,通过每个 Router 的 IP 包中的某些包的数量突然增加,而这些包的源地址是在之前从未出现过的、实际上又不存在的,正是这些地址根本不存在的包,利用 TCP 协议需要进行地址确认才能建立有效的连接的特

点,消耗服务器资源。图 3 列出了通过每个 Router 的之前未出现过的 IP 的数量。

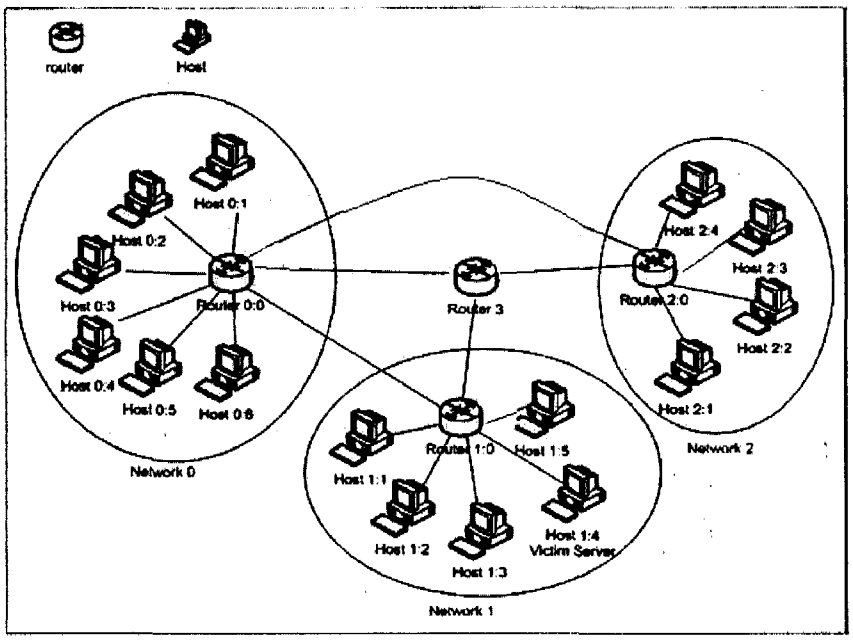

图 2　网络拓扑结构

通过观察不难发现,在攻击时段通过 Router1 的新 IP 数量呈明显的增加趋势,而其他 3 个 Router 的情况没有那么明显,这是因为 Router1 所在的子网是服务器所在的位置,另外 2 个客户端子网的请求最终要汇聚到服务器子网中,所以 Router1 的趋势十分明显。既然如此,如果能判别出网络中存在这一趋势的新 IP 的数量,那么就可以表明网络中存在被 DDoS 攻击的可能性。

值得注意的是,这一趋势是指的新 IP 数量,而不应该是通过每个 Router 的包的数量,原因有两点:第一,暂时的网络拥塞或者由于某些原因带来的网络流量在某段时间的提升很有可能也会符合这种趋势,而这种情况的产生并不是 DDoS 引起的;第二,当发生 DDoS 攻击时网络流量也不一定会出现如此的趋势,DDoS 攻击目的在于消耗服务器资源,使其不能正常提供服务,而不是要破坏网络。

由于 Router 中新 IP 数量的分布趋势与正态分布十分类似,所以文中提出了将基于正态分布的可能性分析方法与通过路由器中新 IP 数量相结合,探测 DDoS 攻击的思想。通过对正常网络中新 IP 数量趋势进行建模,获取合适的探测阈值,然后其应用在各个散步的路由器上,部署一种分布式的防御体系,对通过路由器的新 IP 数量进行统计,与标准模型对比,如果实际结果超过设定的阈值,则表示网络中存在 DDoS 攻击的可能性,然后及时地采取措施,屏蔽掉这些新 IP 地址发送来得连接请求,就可以有效地缓解服务器的压力,去除那些恶意的请求,维护合法用户享受服务。

参照之前得到的公式 $V \in [4\sigma, \mu_\xi - 4\sigma], \mu_\xi - 4\sigma > 0$,分别计算出了适用于每个 Router 的 V 值:$V\text{router}_0 = 4$; $V\text{router}_1 = 7$; $V\text{router}_2 = 2$; $V\text{router}_3 = 3$;通过对比图 3 中的图例,可见求得的阈值对此拓扑结构的网络是适用的,也从实际上验证了基于正态分布概率理论的 DDoS 探测方法。

5　结　论

DDoS 攻击直接威胁着网络和计算机的安全。文中不仅详细分析了 DDoS 产生原理,而且论述了 DDoS 防范策略。特别强调了主动式 DDoS 防范方法——基于正态分布的统计学方法来分析和检测 DDoS 发生。最后通过一个仿真模拟实验验证了这种方法的可行性。随着信息技术和网路技术的发展,分布式主动 DDoS 攻击防范策路将成为保护信息和网络安全的重要手段。

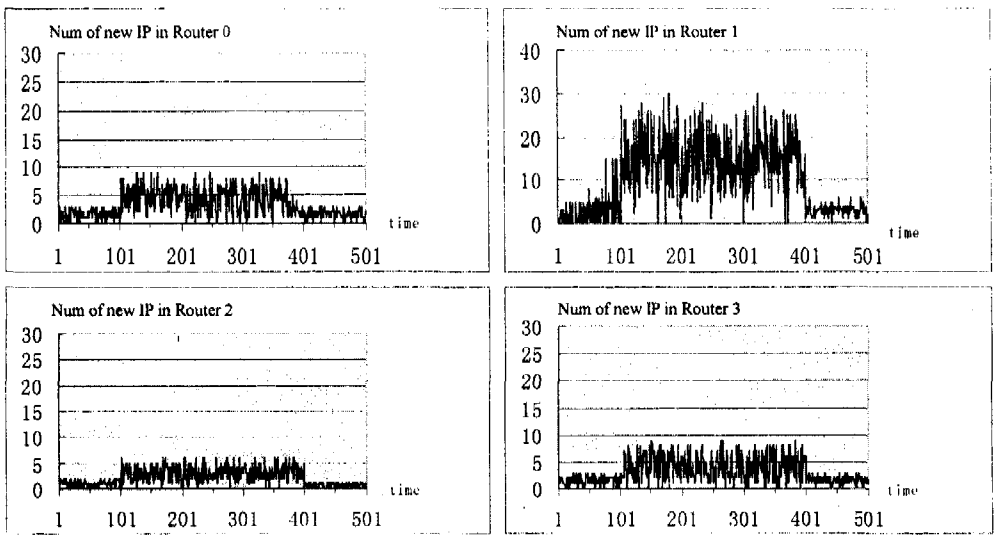

图 3　通过每个 Router 的新 IP 地址的数量

(下转第 243 页)

表单定义、表单业务数据和当前运行流程实例等数据。

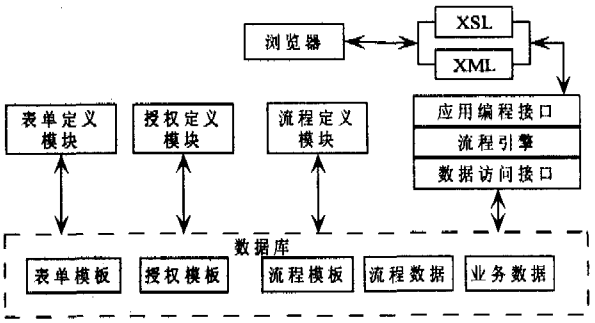

图 2　细粒度授权系统架构图

为降低授权操作难度,系统提供简单、易用的定制工具,将表单模型、授权模型、业务规则、流程模型定义四种功能集成。定制工具的程序流程如图 3 所示。通过定制工具,可以实现 XML 格式的流程文档的可视化编辑,极大减小了流程定义和维护的工作量。

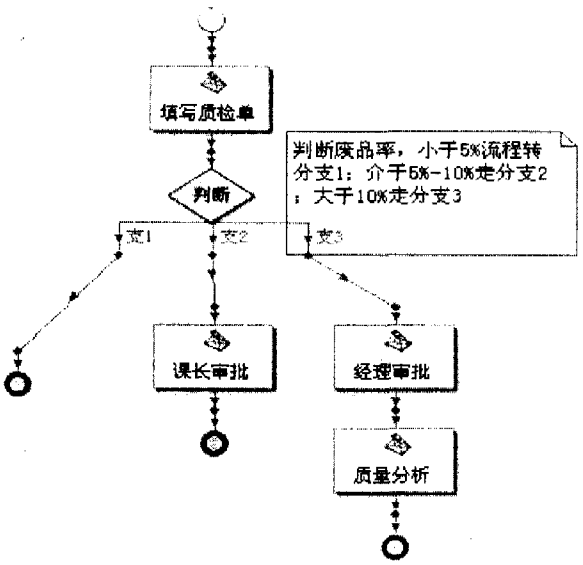

图 3　业务规则定制工具

2.2　性能优化

系统在实现时,考虑到 XML 解析时资源消耗大的缺点,进行了一系列的性能优化工作,从而大大降低了程序对计算机资源的依赖性,减少了冗余计算,适时释放了内存空间,其具体优化策略如下:

(1) 流程引擎的解析算法优化。

根据流程模型、授权模型、表单模型中的自定义标签的规律进行解析优化,保证基本数据解析操作的高效。

(2)XML 表单预处理。

针对系统中动态表单、静态表单并存的情况,在表单定义阶段将无动态业务授权逻辑的静态表单预先生成并存入数据库的特定表。在流程执行时无需动态创建即可获得 XML 表单,极大降低系统运行时的计算量。

(3)流程文档缓存。

当同一流程有多个实例被激活时,重复解析同一流程文档将导致效率低下。系统采用缓存技术,将已解析流程文档用关键字标识,缓存在系统内存中;系统激活同流程的实例时,则从缓存中获取已解析的流程文档,从而大大减少系统的内存消耗。

3　结束语

针对基于 Web 的工作流对授权模型的细粒度、灵活性的要求,提出一种工作流细粒度授权框架。框架通过表单模型、授权模型和流程模型的协同工作,较完善地解决了工作流授权的细粒度问题,具有一定的可扩展性、重用性和灵活性。在下一步的研究工作中,准备进一步完善流程模型的描述能力,同时优化系统执行性能,以适应复杂的企业流程需求。

参考文献:

[1] 范玉顺 . 工作流管理技术基础[M].北京:清华大学出版社,2001.

[2] Sandhu. Role-based Access Control models[J]. IEEE Computer, 1996,29(2):38-47.

[3] Bertino E, Bonatti P A, Ferrari E. TRBAC: A temporal role-based access control model[J]. ACM Transactions on Information and System Security,2001,4(3):191-223.

[4] 黄　建,卿斯汉, 温红子.带时间特性的角色访问控制[J].软件学报,2003,14(11) : 1944-1954.

[5] 王小明,赵宗涛,郝克刚.工作流系统带权角色与周期时间访问控制模型[J].软件学报,2003,14(11):1841-1848.

[6] Thomas R K, Sandhu R. Task-based authentication controls (TABC):A family of models for active and enterprise-oriented authentication management [C]//In: Proc of the IFIP WG1113 Workshop on Database Security. London : Chapman & Hall,1997:166-181.

(上接第 239 页)

参考文献:

[1] 程　光,龚　俭,丁　伟.基于抽样测量的高速网络实时异常检测模型[J]. 软件学报,2002,13(4):594-599.

[2] Yang Xiang,Zhou Wanlei,Chowdhury M. A Survey of Active and Passive Defense Mechanisms against DDoS Attacks[R]. Australia:Deakin University,2004.

[3] Bellovin I J. Pushback:Router-Based Defense Against DDoS Attacks[C]//In Proceeding of Network and Distributed System Security Symposium. San Diego, USA:[s.n.],2002.

[4] Mohiuddin S, Hershkop S, Bhan R, et al. Defending against a Large Scale Denial of Service Attack[C]// In Proceedings of the 3rd Annual IEEE Information Assurance Workshop. New York:United States Military Academy West Point,2002.

[5] Li M, Chi C. Decision Analysis of Statistically Detecting Distributed Denial-of-Service Flooding Attacks[J]. International Journal of Information Technology and Decision Making,2003,2(3):397-405.