基于概率的入侵容忍系统表决机制设计

费洪晓,李钦秀,李文兴,覃思明
(中南大学 信息科学与工程学院,湖南 长沙 410075)

摘 要:普通的大数表决算法不能区分一致正确的响应和一致错误的响应,导致错误输出和没有输出的问题。针对入侵容忍系统中输出结果准确性不高的问题,提出了一种应用概率理论作为额外信息的基于概率的入侵容忍表决机制。该机制根据没有出错的服务器产生的响应结果和出错服务器产生响应结果的概率计算出每个响应结果出现的概率,并根据最大可能性原则找出出现概率最大的那个响应结果,这个响应结果就是正确的输出。分析结果表明,该机制比普通的大数表决输出的结果更为准确,提高了系统的输出准确性。采用了冗余技术和多样性技术,系统即使在遭受到恶意攻击的情况下仍能安全、可靠地运行。

关键词:入侵容忍;冗余;表决;概率理论;最大可能性原则
中图分类号:TP302.1 **文献标识码**:A **文章编号**:1673-629X(2010)03-0136-04

Design of Intrusion Tolerant Voting Mechanism Based on Probability Strategy

FEI Hong-xiao, LI Qin-xiu, LI Wen-xing, QIN Si-ming
(School of Information Science and Engineering, Central South University, Changsha 410075, China)

Abstract: Traditional majority voting algorithm can not distinguish the coherent right response and coherent wrong response, and there will be wrong and no output situation. Current intrusion tolerant systems have some limits on accuracy. To solve this problem, a novel intrusion tolerant architecture with voting mechanism based on probability strategy is proposed. The probability of voting mechanism is based on the application of probability theory as an additional information. The mechanism is based on the probability of no-error servers responsing to results and the error servers responsing to results to calculate the probability of every results. It is based on maximum-likehood principle to find the greatest probability of results, which is correct output. Analysis shows that the mechanism based on probability strategy can vote more accurate output than traditional majority voting algorithm, so a more accurate system output will be improved. Both diversity, redundancy intrusion tolerance technology are used to build Web server system with intrusion tolerance properties. This system can provide correct and reliable services for the system continually despite the existence of fault or intrusion.

Key words: intrusion-tolerance; redundancy; voter; probability theory; maximum-likehood principle

0 引 言

网络的发展使越来越多的人使用网络服务,但是随着互联网应用的普及不可避免地伴随着攻击,越来越多的漏洞被利用,几乎每天都有新安全警告发布。大多数的网络信息系统采用入侵防御机制,例如使用加密和数字签名技术保护信息的完整性和可用性,然而许多攻击却在这些系统中找到了新的漏洞。入侵检测系统只能较有效地抵挡已知的和定义好的攻击,而对于新出现的攻击,它们抵御和检测的措施或策略往往是滞后的,这些安全技术的主要目的是防御攻击或入侵。但是事实上,这些防御措施对于一些恶意攻击有时是无效的[1]。因此在存在入侵的情况下,系统如何为合法用户提供正常的服务就成为了一个非常重要的问题[2]。

入侵容忍技术[3]的出现,为解决上述问题提供了很好的思路,它使系统具有弹性,能承受一定限度的攻击,在系统受攻击后仍然可以为合法用户提供不间断的服务。入侵容忍系统利用容错理论、门限密码、冗余、多样性技术和拜占庭等技术[4]为系统提供入侵容忍能力,它对入侵的检测主要运用表决技术[5],而表决技术在入侵容忍系统中起到了非常关键的作用。

目前广泛应用的是大数表决策略,其基本思想是:在有 n 个输入请求的大数表决器中,若至少有「(n +

收稿日期:2009-07-01;修回日期:2009-10-21
基金项目:湖南省科技计划基金资助项目(2006JT1040)
作者简介:费洪晓(1967-),男,副教授,CCF会员,研究方向为信息过滤和网络安全。

1)/2⌉个复制品的输入满足一致协商则可产生一个输出,此时会在满足一致协商的输入值中任意地选择一个作为表决器的输出,若少于⌈(n + 1)/2⌉个复制品的输入满足一致协商则不能产生输出[6]。在以下的条件下普通大数表决能够正常工作:(1) 不超过 $n - \lceil (n+1)/2 \rceil$ 个复制品产生错误的结果;(2) 所有的正确结果是一致的。但是在系统执行过程中往往会出现差强人意的结果,普通的大数表决算法可能会选择不正确的响应作为输出,如果有至少⌈(n + 1)/2⌉个错误的响应结果它也认为此结果是正确的,因此这种算法不能够区别一致正确的响应和一致错误的响应,会出现错误输出的情况。基于概率的表决应用概率理论作为额外的信息可以很好地解决错误输出和没有输出的问题,使系统产生更加准确的输出结果。

1　系统结构

应用多样性和冗余技术构建容忍入侵系统,由冗余多样性的 Web 服务器、代理服务器、自适应重配置模块和表决模块等组成,系统结构如图 1 所示。

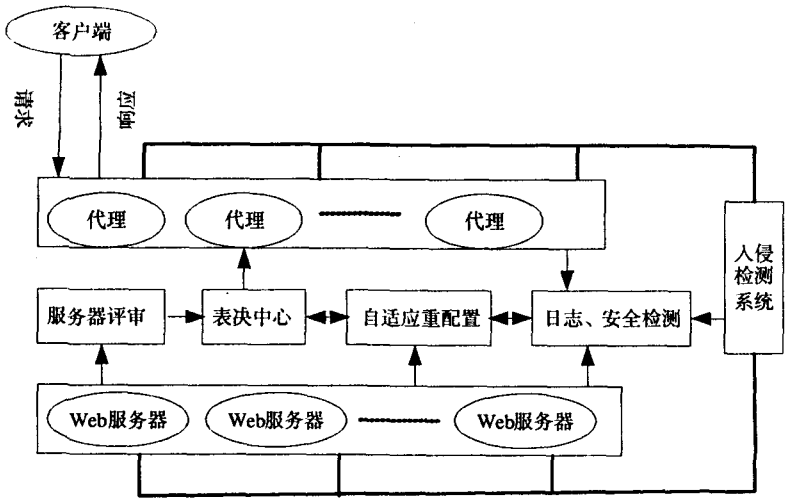

图 1　系统结构图

代理服务器和 Web 服务器在功能上都是冗余的,但是在实现上各不相同,这样被同一种攻击入侵的可能性就非常小。

代理服务器接收客户端发送出的请求,对请求进行过滤验证,如果请求通过验证,请求将被发送到 Web 服务器,Web 服务器对请求进行响应,并把响应结果送到表决模块,表决模块通过表决把最终的正确响应结果通过代理服务器送到客户端。日志、安全检测模型能够对系统的日志进行收集,并且能够检查 Web 服务器的状态和服务的内容。为了使代理服务器和 Web 服务器能够正确的执行,系统中安装了入侵检测系统 IDS,它能够检测出攻击并且能够监视服务器和代理服务器的性能。自适应重配置模块从表决模块和日志安全检测模块、IDS 中得到信息,发现系统异常情况时,将系统恢复到原始状态,或者通过相关策略重新构造 Web 服务器系统。

2　基于概率的入侵容忍表决机制设计

表决器从冗余的 Web 服务器中得到输入,选举出正确的结果,如果选举出的响应结果是不可能的,表决器会产生异常。冗余的服务器对请求进行处理期间会产生多种响应结果,每个值是否是正确的都有一定的概率,因此可以通过概率的方法衡量响应结果的正确性,从而计算出响应结果是否是正确的概率。为了能有效的计算出响应结果是正确的概率,主要应用了(1)每个响应结果的先验概率;(2)出错服务器产生每个响应结果的概率;(3)每个服务器出错的概率。

2.1　表决模型的建立

基于概率的表决方案的逻辑结构主要由三部分组成:n 个冗余的服务器,服务器评审模块和表决器模块,逻辑结构如图 2 所示。

大数表决算法没有考虑到服务器的可靠性和响应结果的可信度,在实践中,软件的可靠性是不同的。入侵容忍系统中采用冗余技术设计 Web 服务器,每个服务器的可靠性是不同的,服务器的可靠性能通过统计的方法进行测量[7~9]。

服务器评审模块应用验收测试技术评估各个 Web 服务器的概况,从而计算出 Web 服务器出错的概率,以及响应结果出现的概率。它对服务器的评估用于每轮表决的概率计算,为了得知 Web 服务器中出现哪些错误类型和可能有哪些安全隐患,服务器评审模块主要从以下几个方面对服务器进行评估:

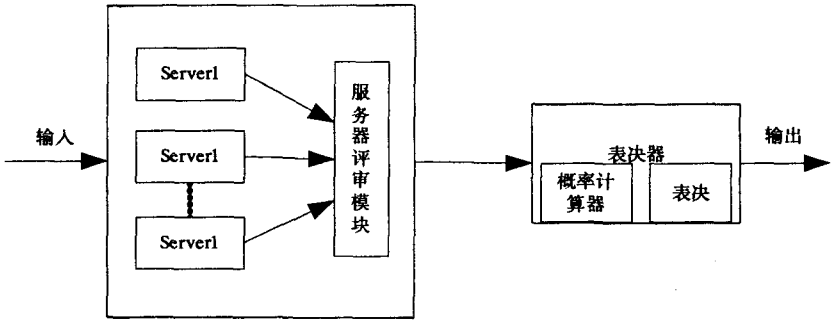

图 2　基于概率的表决方案的逻辑结构图

请求测试:在有些情况下,一些条件强加到容忍入侵结构中去完成任务。这些条件是预计顺序的事件或给定事件的一个子集,请求测试时要测试这些强加的条件是合理的。

合理性评审:合理性检查用于通过预先计算的范围检测软件故障和预计序列程序状态,合理性检查是根据物理因素,而需求测试主要是基于逻辑或数学关系。

定时测试:在容错中,时间是用来测试系统随着时间的推移,以确定敏感元件执行时间是否达到了限制。系统中,使用时间测试,以检测拒绝服务在时间敏感和非时间敏感的服务中的状况。在时间敏感的服务中,限制时间参数以确定服务器的响应到达时间;在非时间敏感的服务中,合理的时间参数模块也给出了。

表决器的输入就是对象组中复制品所发出的消息序列即为$\{Rp_1, Rp_2, \cdots m\}$,假设系统中有两类基本消息:请求(request)和响应(response)。概率计算器从服务器评审模块得到各个服务器以及响应结果的信息,并利用公式进行概率计算,通过表决可以表决出概率最大的响应结果。

2.2 基于概率的表决方案设计

2.2.1 方案中的符号含义

首先,介绍方案中相关符号的含义。① n:系统中冗余服务器的数目;② p_k:第 k 个服务器出错的概率;③ Rp_i:服务器可能出现的响应;④ r_i:出错服务器的响应为 Rp_i 的概率;⑤ q_i:Rp_i是正确值的概率;⑥ R:所有冗余服务器响应结果的集合;⑦ $C_i(R)$:在集合 R 中产生 Rp_i 为响应结果的服务器的数目。

2.2.2 基于概率的表决方案的建立

只考虑获得正确的响应,出错的服务器的响应结果可以为任意的值,没有出错的服务器的响应结果可以认为是正确的。方案用到了每个正确响应结果和出错服务器产生响应结果的先验概率,先验概率可以通过系统中的服务器评审模块中的验收检测来获得。该方案需要选择出现概率最大响应结果,概率公式:

$$Pr(R \text{ occurs} \mid Rp_i \text{ is correct}) = \frac{Pr\{R \text{ occurs} \mid Rp_i \text{ is correct})\} \cdot Pr\{Rp_i \text{ is correct}\}}{Pr\{R \text{ occurs}\}} =$$

$$q_i \cdot \sum_{h=0}^{G(R)} Pr\{R \text{ occurs} \mid (Rp_i \text{ is correct}) \cap (h \text{ servers are non-faulty})\} \cdot Pr\{h \text{ servers are non-faulty} \mid Rp_i \text{ is correct}\} = q_i \cdot \Big(\prod_{Rpi \in RR:\, Rpj \neq Rpi} r_j^{G_j(R)}\Big) \cdot \sum_{h=0}^{G(R)} \binom{G_i(R)}{h} \cdot (1-p)^h \cdot p^{n-h} \cdot r_i^{G_i(R)-h} \quad (1)$$

此方案用到了最大可能性原则[10],并且当一个响应结果概率是最大的并且大于某个门限值 α 的时候将此响应输出,否则产生异常信息。此方案的流程图如图 3 所示。

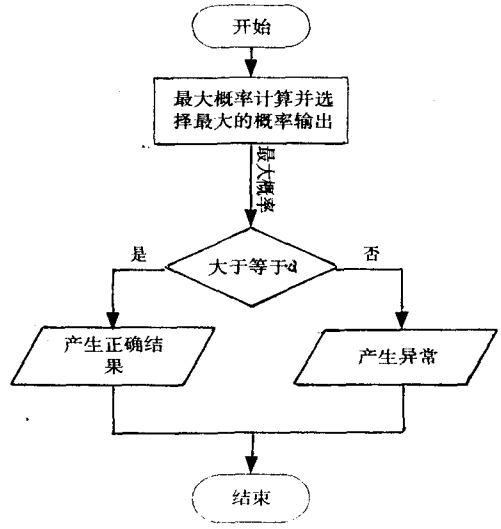

图 3 基于概率表决方案的流程图

2.2.3 策略举例

下面以一个具体实例阐述该策略的运行过程:

对于一个具有 7 个冗余服务器的系统即 $n = 7$,服务器可产生5个响应结果即 $Rp_1, Rp_2, \cdots, Rp_5$。R 是响应结果的集合$\{Rp_1, Rp_2, \cdots, Rp_5\}$,每个服务器出错的概率 $p = 0.4$,并且响应为 Rp_3 的服务器的数目为 4 个,响应结果为 Rp_5 的服务器的数目为 3 个。

已知条件如表1所示:编号1所在的行代表所有的响应有着相同的发生概率,出错服务器产生此种响应的概率也相同;编号 2 所在的行代表所有的响应有着不同发生的概率,出错服务器产生此种响应的概率相同;编号 3 所在的行代表所有的响应有着相同的发生的概率,出错服务器产生此种响应的概率不相同。

表 1 已知条件

编号	q_1	q_2	q_3	q_4	q_5	r_1	r_2	r_3	r_4	r_5
1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
2	0.15	0.1	0.05	0.1	0.6	0.2	0.2	0.2	0.2	0.2
3	0.2	0.2	0.2	0.2	0.2	0.2	0.15	0.4	0.15	0.15

2.2.4 结果分析

计算结果如表 2 所示。

表 2 计算结果

编号	Pr_1	Pr_2	Pr_3	Pr_4	Pr_5	Max
1	0.00017132	0.00017132	0.89427700	0.00017132	0.10520906	$Rp3$
2	0.00023816	0.00015877	0.41440373	0.00015877	0.58504056	$Rp5$
3	0.00054257	0.00054257	0.27620622	0.00054257	0.72216606	$Rp5$

第一种情况,当所有的响应有着相同的发生的概率并且出错服务器产生此种响应的概率也相同即都为 0.2 的时候,Rp_3 出现的概率比较大,有 89% 正确的可

能性,并且大于门限值 α,所以将 Rp_3 作为响应结果输出。第二种情况,即所有的响应有着不同发生的概率,出错服务器产生此种响应的概率相同出现的时候,由计算结果可以看出出现的最大可能性不再是 Rp_3 而是 Rp_5,并且大于门限值 α,因此把 Rp_5 作为最终的输出结果输出。第三种情况,即所有的响应有着相同的发生的概率,出错服务器产生此种响应的概率不相同,由计算结果可以看出出现的最大可能性不再是 Rp_3 而是 Rp_5,因此把 Rp_5 作为最终的输出结果进行输出。由此可见对 r_i 进行细微的改变,Rp_5 是正确的可能性由不到 0.11 上升到 0.72。应用普通的大数表决算法产生的输出是 Rp_3,普通的大数表决算法没有考虑到每个错误服务器产生响应的概率是改变的,它们产生错误的结果作为输出的概率达到了 3/4,出现了低的准确性。

3　结束语

基于概率的入侵容忍表决机制实现了系统的安全性、可靠性,为系统提供了容忍入侵能力,与普通大数表决方法相比,基于概率的表决能输出更准确的响应结果。显然,文中提出的这种容忍入侵结构与模型具有一定的理论价值和实验应用价值,特别适合应用于对系统准确性、可靠性、可生存性的分布式环境中。

参考文献:

[1] 朱建明,马建峰.基于容忍入侵的数据库安全体系结构[J].西安电子科技大学学报,2003,30(1):85-89.

[2] 彭文灵,王丽娜,张焕国,等.基于角色访问控制的入侵容忍机制研究[J].电子学报,2005,33(1):91-95.

[3] 张险峰,张峰秦.入侵容忍技术现状与发展[J].计算机科学,2004,31(10):20-22.

[4] 柴争义.入侵容忍技术及实现[J].计算机技术与发展,2007,17(2):229-231.

[5] 殷丽华,何　松.一种入侵容忍系统的研究与实现[J].通信学报,2006,27(2):137-142.

[6] Reynolds J, Just J, Lawson E. The design and implementation of an intrustion tolerant system[C]//Proceedings of Int'l Conference on Dependable Systems and Networks. Washington D.C.:[s.n.], 2002:258-290.

[7] Musa J D. Tools for measuring software reliability[J]. IEEE Spectrum, 1989,26(2):9-42.

[8] Brocklehurst S, Littlewood B. New ways to get accurate reliability measures[J]. IEEE Software, 1992,9(4):34-42.

[9] Sheldon F T, Kavi K M, Tausworthe R C. Reliability measurement: From theory to practice[J]. IEEE Software, 1992, 9(4):13-20.

[10] Leung Y W. Maximum likelihood voting for fault-tolerant software with finite output-space[J]. IEEE Trans. Rel, 1995,44(3):419-427.

(上接第 135 页)

送组播数据,在小规模的范围内采用 IPv6 PIM-DM 的部署方式,配置和管理较为容易。

(5)接入主干网设计。

在石牌校区采用两条 1Gbps 独立光纤分别接入教科网现有的 CERNET 主干网华南核心节点华南理工大学和 CNGI-CERNET2 主干网华南核心节点华南理工大学,目前线路已经正常使用。在纯 IPv6 子网建设中,我校将把大学城校区到大学城汇接中心的出口也升级为双栈协议,并通过动态路由协议实现多路由出口。

4　结束语

目前,我校已经完成了校园网 IPv6 的升级部署,而我校 CNGI 驻地网建设子项目也已经顺利通过验收。通过 CNGI 驻地网子项目建设,我校校园网出口完成与下一代互联网的高速对接,各项基础设施已逐步升级更新,IPv6 试验和应用范围逐步扩大,各项 IPv6 的应用开始逐步实施。接下来将进一步完善 IPv6 网络服务与网络管理,进一步研发 IPv6 应用系统,在新一代 IPv6 网络上实施高性能流媒体传输应用等,为我校启动下一代互联网的全面建设和应用及进一步参与国家下一代互联网的研究积极做好准备工作。

参考文献:

[1] 王相林.IPv6 技术新一代网络技术[M].北京:机械工业出版社,2008:12-56.

[2] Li Qing. IPV6 Advanced Protocols Implementation[M].北京:人民邮电出版社,2009:23-49.

[3] Childress B, Cathey B, Dixon S. The adoption of IPv6[J]. Journal of Computing Sciences in Colleges, 2003, 18(4):39-51.

[4] 张　乐,夏　昕,陈　萌.IPv4 向 IPv6 过渡策略研究综述[J].科技广场,2008(12):94-95.

[5] 张五红,王　宇.高校 IPv6 校园网的部署与配置[J].计算机工程与设计,2007,28(13):3106-3110.

[6] 张天云.IPv6 技术及其在校园网的部署[J].信息技术,2007,35(1):25-26.

[7] 郭东恩,沈　燕.ORACLE 透明网关技术实现异构数据库互连[J].电脑开发与应用,2008,21:68-126.

[8] 张宏科,苏　伟.IPv6 路由协议栈协议栈原理与技术[M].北京:北京邮电大学出版社,2006:50-125.