一种用于并行入侵检测系统的数据分流策略

A Data Distribution Strategy for Parallel Intrusion Detection System

伍海波　陶　滔　（南华大学 计算机科学与技术学院 湖南 衡阳 421001）

摘　要：流量分配是影响并行入侵检测系统实时性的重要因素。提出了一种数据分流策略，将捕获的网络数据按照某种分流策略转发到多个探测器进行处理，解决目前基于网络的入侵检测系统跟不上高速网络发展而带来的丢包问题，达到提高整个系统的检测性能。最后通过试验分析表明该策略是有效的。

关键词：高速网络 入侵检测 数据分流

1　引言

入侵检测是一种主动的网络安全防御措施，是指对入侵行为的发觉。它主要是通过对计算机网络或计算机系统中的若干关键点收集信息并进行分析，从中发现是否存在有违反安全策略的行为和被攻击的迹象。它不仅检测来自外部的入侵行为，同时也监督内部用户的未授权活动。实现入侵检测功能的一系列软件和硬件的结合就构成了入侵检测系统（IDS）。现在采用的 IDS 分为基于主机分析和基于网络数据包分析两类，其中使用较多的是基于网络的入侵检测系统。

随着网络带宽越来越宽，网络入侵检测系统的越来越难跟上网络的速度了，利用一台主机进行集中式检测几乎已经无法应付现在高速网络流量的实时分析。因此研究通过多机并行操作实现基于网络的入侵检测系统的相关技术势在必行。并行入侵检测系统的关键问题是如何将流量分配给各探测器。流量分配是影响并行入侵检测系统实时性的一个重要因素。因此，本文提出了一种用于并行入侵检测系统的数据分流策略来实现高速网络下的实时入侵检测。

2　相关研究

2002 年，C. Kruegel 等人首次提出了一种用于高速网络的并行入侵检测系统，虽然提高了网络入侵检测系统适应高速网络的能力，但它采用的流量分配算法是仅仅基于 IP 地址来转发数据包到各探测节点，且该算法没有考虑各探测节点的负载动态平衡。Charitakis 等人提出通过在分流器上使用早期过滤和使用本地缓存两种方法来提高并行入侵检测系统的性能。虽然两种方法都在一定程度上都提高了系统的性能，但早期过滤使的探测器无法检测到所有的数据包，而本地缓存会造成数据包到达探测器的次序和原始次序不一致，从而使得探测器无法进行连接状态的分析和使用一些基于统计的检测方法，另外，该算法只是采用了简单的、基于 Hash 的流量划分算法，并没有考虑划分后流量的动态平衡性。国内的吕志军等人在其描述的高速网络下的分布式实时入侵检测系统中提出了一种流量分配策略，但它的前提条件是所有探测节点的处理能力相同，然而现实中入侵检测系统的所有探测节点的处理能力往往是不同的。

总之，目前这方面的研究都有一定的局限性。本文提出了一种新的数据分流策略，采用两层结构来实现对数据的分流可以提高系统的检测性能。

3　并行入侵检测系统模型

并行入侵检测系统的流量分配结构如（图 1），分流器是整个结构的中心，负责捕获流经该网段的所有数据包并把数据包发送到各分配器。为了保证系统的整体性能，分流器上的计算和操作应该尽可能的简单。各分配器把数据包经交换机转发到各探测器，探测器运行相同配置的 Snort。控制器通过动态反馈机制，定时获取各个探测器当前的负载，负责各探测器的负载平衡。

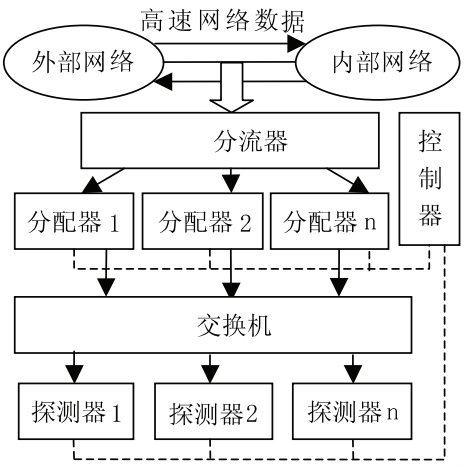

图 1　并行入侵检测系统流量分配结构

4　并行入侵检测系统数据分流策略

目前一般采用的流量分配方法有轮转法(Round Robin)、最少连接法、基于应用层协议的分配算法、Hash 算法等。值得注意的是,在选择流量分配策略时应考虑以下三个问题:第一,每一个探测器的负载应该均衡;第二,属于同一个攻击的数据包应该分配给同一个探测器进行处理;第三,算法的执行效率高。

4.1 流量分配方式

为了准确计算每个探测器的负载情况,下面定义一些参数和函数,设每一个探测器都有一个固定的编号 i(i>0),每个探测器对应的负载函数为 Li(t),连接个数的函数为 Ci(t),缓存中待处理数据包个数的函数为 Pi(t),CPU 占用率函数为 Ei(t),内存占用率函数为 Mi(t),从而可以得到第 i 个探测器在 t 时刻的负载

$$Li(t)=k_1Ci(t)+k_2Pi(t)+k_3Ei(t)+k_4Mi(t) \quad (1)$$

式中 k_1、k_2、k_3、k_4 为动态调整系数,取值应根据实际的流量环境通过试验分析和比较得到,并且 $\sum_{i=1}^{4}k_i=1$。

流量分配分为两部分,第一部分分流器采用的是非常简单的轮转(Round Robin)算法,在进行请求分配的时候不考虑分配器的负载问题,只是简单的快速转发,目的是尽快地将请求分配到后端分配器,避免请求分流器层成为系统的瓶颈。第二部分分配器按照基于连接的流量分配算法把数据包经交换机转发到各探测器,同时各探测器按照式(1)计算出各探测器的当前负载交给控制器。控制器根据各探测器的实际负载情况,通知分配器下一步调整数据包的转发。

4.2 算法描述

绝大多数网络攻击都可以在一个连接中检测出来,一个 TCP 连接可以用五元组表示:〈PN,SIP,SPT,DIP,DPT〉,其中 PN 代表协议号,SIP 代表源 IP 地址,SPT 代表源端口,DIP 代表目的 IP 地址,DPT 代表目的端口。而对于 UDP 和 ICMP 等无连接协议的数据包以及一些不完整 TCP 连接的数据包可以定义相近时间内的具有相同五元标识〈PN,SIP,SPT,DIP,DPT〉的所有数据包为一个连接。

每个分配器动态维持一张分配表,表中内容包括探测器号、连接标识和更新时间。为了提高检索效率,此表可以采用 HASH 算法实现。分配器根据分配表把属于同一个链接的数据包分发给同一个探测器,从而保证攻击证据的完整性。当一个数据包到来时检索分配器中的分配表,若表中存在和该数据包对应的〈PN,SIP,SPT,DIP,DPT〉,则首先更新表中的项“更新时间”,然后将该数据包转发到相应的探测器。若不存在对应的〈PN,SIP,SPT,DIP,DPT〉,则利用(1)式计算的负载选择出一个负载最小的探测器,然后把数据包发送给该探测器,同时在分配表中添加一个新的连接记录。随后同一连接的所有数据包都分发到该探测器。当一个连接结束时,就将该连接记录从分配表中删除。对于 TCP 数据包,FIN 和 RST 标志用来结束一个连接。对于无连接协议(UDP、ICMP 等)的数据包,以及不完整 TCP 连接的数据包,则利用表中的记录项“更新时间”定时删除超时的连接记录。探测器根据已知的规则对所接收到的数据包进行检测。这样每个探测器就不需要检测流经主干网的所有数据包了,因而降低了因网络数据包流量过大而带来的丢包率,实现了高速网络下的入侵检测。

5　性能分析

流量分配采用两层结构,第一层只对流量采用轮转算法进行简单的划分并没有具体分配到各探测器,这样可以更好地与 Gbit/s 的网络速度相匹配,避免出现系统瓶颈;第二层分配是散列法与最小负载的结合,散列可以保证来自相同连接标识的数据包送到同一个探测器进行检测,提高探测器的访问局部性,最小负载法则可以平滑突发性实时任务,改善单纯散列调度算法的负载均衡能力。

为了测试该流量分配结构的有效性,在千兆局域网中搭建了测试环境。通过试验测试结果可以发现,当流量比较小的时候,分流转发和未经分流的性能差不多,但当流量超过200Mbps,未经分流的入侵检测系统的性能急剧下降,而经分流转化的入侵检测系统的性能并没有明显的变化,只有在流量达到400Mbps时才开始下降。

6 结束语

并行入侵检测系统是解决高速网络环境下入侵检测的有效措施,流量分配策略是影响并行入侵检测系统性能的关键。本文在前人的基础上,具体针对高速网络环境下的入侵检测系统的分流结构,在分流时引入动态负载均衡算法,提出了一种并行网络入侵检测系统的数据分流策略,经过理论分析和相关试验研究表明,该分流策略在高速网络环境下的并行入侵检测系统中是有效的,具有重要的理论意义和研究价值。

参考文献

1 卿斯汉,蒋建春,马恒太,文伟平,刘雪飞. 入侵检测技术研究综述. 通信学报,2004,25(7):21 – 24.

2 刘学波,孟丽荣. 高速网络环境下的网络入侵检测系统的研究. 计算机工程与设计,2005,26(5):1236 – 1238.

3 C. Kruegel, F. Valeur, G. Vigna, and R. Kemmerer, " Stateful intrusion detection for high – speed networks," in Proceedings of the IEEE Symposium on Research on Security and Privacy. Oakland, CA: IEEE Press, May 2002:285 – 293.

4 I. Charitakis, K. Anagnostakis, and E. Markatos. An active traffic splitter architecture for intrusion detection. In Proceedings of 11th IEEEWACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2003), Orlando. October 2003: 238 – 241.

5 吕志军,黄皓. 高速网络下的分布式实时入侵检测系统. 计算机研究与发展,2004,41(4):667 – 673.

6 薛华,李祥和,许榕生. 利用分割机制实现高速网下入侵检测的研究. 计算机工程,2004,30(3):33 – 35.