广西科学院学报
Journal of Guangxi Academy of Sciences

2007,23(3):184~187
Vol. 23,No. 3 August 2007

一种基于状态树和拓扑分析的网络入侵预警系统
An Intrusion Detection Warning System Based on Status Tree and Topology Analysis

徐邕江
XU Yong-jiang

(广西经济信息中心,广西南宁 530022)
(Guangxi Economic Information Center,Nanning,Guangxi,530022,China)

摘要:在LINUX平台下,采用C/C++语言实现了基于状态树和基于拓扑分析的预警系统。该系统能够快速准确地对即将发生的入侵行为进行预警,为大规模网络系统的安全性提供一定保障。该系统框架具有良好的可靠性和扩展性,能够适应于分布式入侵检测预警的需要。

关键词:入侵检测 入侵状态 拓扑分析 预警
中图法分类号:TP393.08 **文献标识码**:A **文章编号**:1002-7378(2007)03-0184-04

Abstract:In this paper,a warning system based on state tree and topology analysis is implemented by C/C++ on linux platform. An intrusion can be warned rapidly and precisely and a large scale network security can be archieved in some degree. The framework of this system is reliable and expansible,and especially proper for distributed intrusion detection network warning system.

Key words:intrusion detection,intrusion status,topology analysis,warning

随着计算机网络技术的发展,网络安全问题日益突出,如何有效防御非法的入侵和攻击是维护网络安全最关键的问题之一。入侵检测技术[1]是近几年发展起来的一种能够动态监控和防御计算机和网络系统入侵行为的安全机制。它从系统内部和各种网络资源中主动采集信息,从中分析可能的网络入侵或攻击并及时通知管理员,以便及时采取相应的防范措施,把危害减少到最低程度。入侵检测系统分析网络数据包和系统的审计数据,实现对网络和系统的智能监控、实时探测、动态响应,可以防范来自系统内部和外部的入侵行为。因此,入侵检测技术通常被看作是继防火墙之后的又一道安全闸门。

随着网络规模扩大,攻击手段日益多样,网络环境日益复杂,单纯依靠基于误用和基于异常的检测机制已不能及时地起到防范作用,如何有效地检测高速网络入侵行为并及时准确地进行预警成为当前研究的重点。

收稿日期:2007-06-19
作者简介:徐邕江(1956-),男,工程师,主要从事网络信息安全研究工作。

国外早已开展了早期预警系统及入侵检测技术的研究,在一些重要的政治、军事和经济网络上对非法入侵实施监控。这些系统在保障信息网络安全,尽早发现入侵攻击迹象,分析入侵攻击的技术手段等方面发挥着重要的作用。目前我国在局域网的入侵检测系统与预警系统方面的研究已取得初步成果,但是典型的入侵检测产品,如Snort、Warsher等,在应用中会产生大量的报警信息,而在分布式应用环境下这种数据量更是以几何级增长,许多有价值的信息往往被淹没在这些海量的警报数据中[2]。目前我国还没有研究出适用于大规模网络的入侵检测与预警系统。为了保障我国信息系统支持和适应信息战争的要求,开展网络入侵检测与预警系统的研究是十分必要的,它对于提高网络系统的应急响应能力、缓解网络攻击所造成的危害、提高系统的反击能力等方面均具有十分重要的意义。在现代信息战争与维护国家信息安全中,它是有效抵御敌方的信息攻击,维护自身网络体系的正常运作,并进行追踪和反击所必不可少的[3]。

本文采用基于状态树和基于拓扑分析的预警系统,该策略能够快速准确地对即将发生的入侵行为

进行预警，从而为大规模网络系统的安全性提供一定保障。

1 预警原理及框架结构

我们所构建的预警系统就是基于各个检测代理所产生的警报数据。针对报警信息的有序性，通过对入侵检测代理所产生的报警数据进行相关性分析，构建最小状态树，结合预警规则库，对状态树最有可能的生长方向进行预测，从而实现针对主机的入侵的预警。针对攻击的广泛性，利用已有的规则库中的拓扑地图对拓扑的最有可能发展方向进行预测，从而实现对入侵的攻击范围的预警。

1.1 基于状态树的预警

通过分析大量的攻击实例和相关文献，我们发现入侵者在攻击一个目标的过程中，通常是要采用一些基本的模式，进行一些比较固定的步骤，从而会产生一系列的警报状态。一般来说，入侵者的攻击过程主要包括以下几个基本阶段。

(1)信息搜集：入侵者做的第一件事情就是收集情报，他们需要先判断系统是否存在安全漏洞然后才能确定攻击目标，发动攻击。他们通常会先对一个网段内的主机系统进行端口扫描，检查系统开放的端口，操作系统的类型、版本号，是否存在弱口令等等，进而探测出存在的漏洞。入侵者经常使用的工具有：NSS，Strobe，Netscan，SATAN，Jakal，IdentTCPscan，FTPScan 等，以及各种 Sniffer。从广义上说，特洛依木马程序也是收集信息情报的重要手段。

(2)考虑攻击方法：根据搜集的情报对系统的脆弱性进行评估，从而确定攻击方法，选取合适的攻击工具。

(3)远程攻击：利用漏洞(缓冲区溢出、Unicode 等等)使系统堆栈崩溃，进而取得一定控制权，或者通过一定手段猜测口令。

(4)入侵系统：登录到被攻击的主机。对 Linux 系统，争取获得一个 Shell 以便直接运行命令。寻找系统漏洞，获取用户帐号和口令。利用系统的配置错误和漏洞，获取 Root 或 Administrator 权限。

(5)收尾工作：为方便下次入侵，需要留些后门，例如安装木马程序，为隐藏身份，还要消除系统日志，以及其它可能的痕迹。对应这些阶段，入侵者通常采用不同的工具和命令达到目的，有的直接编写脚本来完成。

由于入侵检测系统对上述攻击步骤进行检测，记录每一步产生的警报，会逐渐形成一条入侵状态序列。我们把这种状态序列与预警规则库中状态树进行匹配，就会找到一棵最有可能的生长子树，再通过概率分析从而确定该子树最有可能的生长方向，从而在该子树还未生长到状态树的叶端之前进行状态预测。

具体方法是，将规则库中存放长成的状态树林，实时接收来自检测部件的警报数据，根据关联分析，形成警报数据中最可能的入侵状态序列，并在此基础上构建最小入侵状态树。再通过和规则库中状态树林的匹配，找到一棵最优匹配树，并对入侵过程中即将到来的下一个攻击状态进行推断。基于状态树的预警如图 1 所示。

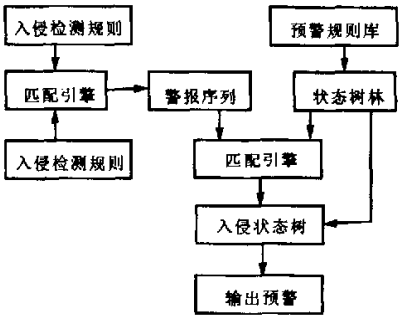

图 1 基于状态树的预警原理

这种方法是利用警报数据的相关处理技术上的一种延伸，但存在一个问题，由于警报是依次到来，所以状态树是由树根开始逐渐生长，最初匹配的状态树的分枝会比较多，对入侵的预测也就有多种可能，我们应该选取可能性最大的预测结果。随着状态树的不断生长，分枝会逐渐减少，我们对攻击的步骤，状态树的生长方向也会有比较明析的判断。但若等完全匹配规则库中的攻击状态树，可能攻击过程已经完成，就没有必要预测了。因此，必须在状态树长成之前进行合理的预测，要在完全匹配之前，也就是在匹配部分规则状态树时就进行预测。但这时可能会产生很多预测结果。

1.2 基于拓扑分析的预警

基于拓扑分析的预警，主要是对在一定物理范围内发生的连续的针对不同目标所进行的同一种攻击的识别，并对其相邻可能的受牵连主机给予预警。这种攻击形式上类似于蠕虫病毒，逐步扩散，由点及面。为及早发现并遏制这种大规模攻击行为，需要对正遭受攻击和已被判定为被攻陷的主机进行记录，并从网络整体布局上进行分析以预测发展趋势。

由于被攻击主机的选取具有随机性(绝大多数

取决于攻击者的主观意图），仅仅基于攻击代码特征很难对下一个被攻击结点进行正确判断。比较现实的方法是对下一个可能遭到攻击的范围以及可能会受到威胁的主机结点都进行预测。为此，首先要对网络中的主机结点按安全等级进行分类，这需要利用网络漏洞扫描一类的安全工具对网络中的主机进行安全评估并标定等级。对属于同一等级的主机划为一类，将其同等对待。当确认某一类别的某一主机结点受到攻击时，则同安全等级主机都有可能成为接下来的受害者，应适当给予警告。预警过程可行式化描述为：

设 $A=\{a_i, i\in 1\sim m\}, m>0$，$A$ 是目前接到的入侵警报集合。设 $N_i=\{n_i|a_i, i\in 1\sim m\}, m>0$，$N_i$ 是受 a_i 类型攻击的主机结点集合。

(1) 当接到某一新报警 a_j，如果 $a_j\in A$，则将警报 a_j 所指结点加入 N_j。

(2) 如果 $a_j\notin A$，则在 A 中添加 a_j，并将警报 a_j 所指结点加入 N_j。

(3) 当 $\mathrm{Count}(N_i)>p$（p 为某一预定值，根据经验获得），则可认为遭受 a_i 攻击的主机已超过一定规模，应该向相关主机结点发出预警。

2 预警系统的关键技术

考虑系统的稳定性和可靠性，软件开发在LINUX平台下完成。系统主要采用LINUX C/C++编写主程序，GTK/GDK完成用户界面，数据库为MYSQL。程序中使用多线程技术、SOCKET网络编程技术、非阻塞技术、知识表示等。

2.1 规则文件

规则文件由多段规则组成，其中每一段代表一棵规则状态树。规则状态树是由攻击状态序列组成的若干分支，叶子节点包括攻击状态信息及系统的报警建议组成，用"!"进行分割。前一部分是攻击状态信息，由攻击名称组成，多个攻击名称之间用"->"分割；后一部分是针对这类攻击序列的报警建议。如下式所示：

Rule={AttackSequence! ResponseSuggestion}

AttackSequence={AttackName1, AttackName2, …, AttackName$I$, …, AttackName$N$}

其中Rule为一条规则；AttackSequence为攻击序列；AttackName$I$为攻击序列中的第I个攻击名称。

程序中规则文件的结构如图2所示。

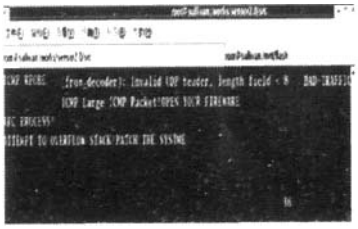

图2 规则文件的建立

2.2 创建规则链表算法

系统初始化时，从规则库文件中加载规则，建立规则树链表。建立规则树林链表的算法用C语言描述。

2.2.1 规则链表的节点结构

对应于规则表示，规则链表节点包括三个域，如图3所示。

Rule	Suggestion	NextRule

图3 规则链表的节点结构

2.2.2 规则链表的建链算法

系统从规则文件中逐条读取规则，建立规则树链表。用C语言描述该链表的建立过程的程序片断如下。

```
……
FILE * fdd;
fdd=fopen("./Alertrules","a+");
  while(!feof(fdd)){
  if(fgets(rulepool,sizeof(rulepool),fdd)
  ==NULL)break;
  /*将缓存中数据进行分割，分别赋值给p的rule和suggestion域*/
  mSplit(rulepool,"!",2,&num—tock5,0);
  sprintf(p->rule,"%s",tock5[0]);
    sprintf(p->suggest,"%s",tock5[1]);
  temp->next=p;temp=p;
  ……
  }
```

2.3 基于攻击状态的预警

本算法中，对同一主机的攻击状态序列是系统预警的依据。预警算法分为四步：

(1) 动态建立被攻击主机状态树链表MachineList

MachineList中的主机以其IP地址唯一标识。因此，IP地址能够唯一表示链表中一个节点，可将其作为链表节点的主键。链表节点结构如图4所示。其中，IpAddr是被攻击主机的IP地址；Matchstring是正在形成的攻击状态名称，格式与规则文件中的

Rule 字段相同；NextMachine 指向下一个节点。

IpAddr	Matchstring	NextMachine

图4 被攻击主机链表节点结构

我们把系统检测到第一台被攻击主机受到的第一个攻击步骤之前的 MachineList 的状态称为链表的初始状态。这种初始状态为空链表。当系统收到第一个报警数据包以后，为该 IP 地址建立第一个表结点，并添加到被攻击主机链表中。对于随后收到的报警数据包，首先提取其 IP 地址，然后，遍历 MachineList，若该 IP 地址对应的结点不存在，则为其创建新结点，并将报警信息写入 Matchstring 字段；若该 IP 地址对应的节点存在，则转向(2)。

(2)为被攻击主机生成状态树 Matchstring

若 IP 地址对应的结点存在，表示该主机已经发现过入侵者的攻击，并已经记录了一个未能与规则匹配的攻击序列。这时，应将报警信息按照既定的格式合并到该结点的 matchstring 字段，扩充其被攻击状态序列树。

(3)Matchstring 与规则状态树链表匹配

如果节点的 Matchstring 字段发生更新，系统立即将更新后的 Matchstring 字段，并匹配程序初始化时建立的规则状态树链表每一结点的 rule 字段，若未能匹配，转向(1)；若正确匹配，转(4)。

(4)预警

若节点 Matchstring 字段匹配到规则状态树链表中的某一子树，表示发现了预先定义的入侵，可以对其生长方向进行预测并采用系统规定的方式向管理员发出预警。预警内容以结点的 Suggest 字段为基础，比如可以将 Suggest 字段的内容输出到屏幕，或者发送 E-mail 等。算法主体的 C 语言描述如下。

```
……
while(p1->next!=NULL)
{if(strcmp(p1->next->ipaddr,addrbuff)==0)
  {strcat(p1->next->matchstring,"\n");
   strcat(p1->next->matchstring,alertmsg);
   break;}
    p1=p1->next;
}
 if(p1->next==NULL){
p2=(struct ipnode *)malloc(sizeof(struct ipnode));
strcpy(p2->matchstring,alertmsg);
strcpy(p2->ipaddr,addrbuff);
p2->next=NULL;p1->next=p2;
p2=p1->next;}
tmp=head;
while(tmp)
{if(strstr(p2->matchstring,tmp->rule))
  sprinft(intrude,"\n Host:%s being intruded! \n
Attack order:\n%s\n",p2->ipaddr,tmp->rule);
      tmp=tmp->next;}
……
```

3 结束语

从警报数据处理分析的角度研究入侵检测技术中的预警技术，分析并初步构建了基于状态树和基于拓扑分析的预警系统。本系统为今后的网络安全建设提供一个很好的参考模型和设计思路。

参考文献：

[1] 蒋建春，马恒太，任党恩，等. 网络安全入侵检测：研究综述[J]. 软件学报，2000，11(11)：1460-1467.

[2] 孙宏伟，田新广，李学春. 一种改进的 IDS 异常检测模型[J]. 计算机学报，2003，26(11)：1450-1455.

[3] CROSBIE M，SPAFFORD E H. Defending a computer system using autonomous agents[C]. Baltimore：Proc of the 18th National Information Systems Security Conference，1995：549-558.

(责任编辑：邓大玉 尹 闯)

科学家在植物中找到天然防冻物质

北海道大学教授藤川清三的研究小组已在植物色素中找到一种防止水在低温下结冰的物质。只要将这种物质以0.01%的比例混入水中，水在零下10摄氏度的环境中也不会结冰。这种物质或许可以用于移植用脏器的低温保存等领域，脏器的保存时间将可能得到大幅延长。

(据科学网)