入侵检测系统攻击特征库的设计与实现①

The Design and Implementation of Signatures Database in Intrusion Detection System

费洪晓　肖新华　（中南大学信息科学与工程学院 长沙 410075）
谢文彪　中南大学信息科学与工程学院 长沙 410075）
（长沙理工大学电气与信息工程学院 长沙 410076）
戴宏伟　（中南大学信息科学与工程学院 长沙 410075）

摘要：攻击特征库是基于误用的入侵检测系统的重要组成部分。本文主要介绍了入侵检测系统攻击特征库的设计思想和实现过程，运用 Snort 规则语言描述了规则库设计，规划了规则库的结构，并对规则解析进行了阐述。

关键词：入侵检测 规则库 攻击特征

随着计算机网络的发展，计算机和网络的安全问题也越来越受到人们的重视。入侵检测系统作为防火墙的合理补充，能够实时地检测出计算机和网络中出现的入侵活动。入侵检测系统主要采用误用检测和异常检测两种技术。误用检测是目前 IDS 系统的主要监测手段，它的核心部件就是攻击特征库[1]。因此，入侵检测系统攻击特征库的完备性和准确性非常重要，它直接影响着基于误用检测的入侵检测系统的性能。

1 特征提取

IDS（Intrusion Detection System，入侵检测系统）中的特征就是指用于判别通讯信息种类的样板数据，也就是用于判别入侵行为及种类的特征数据。根据数据来源的不同，特征的含义也随之不同[2]。基于误用检测的 IDS 就是利用入侵行为的特征，采用模式匹配技术来判断入侵发生的。

1.1 特征选取

特征选取问题是入侵检测系统的核心问题之一，准确的特征选取对于降低入侵检测系统的误报率和漏报率，提高入侵检测系统的检测效率都起着重要的作用[2]。特征选项数量要合适，要遵循以下两个准则。

（1）过多，则特殊性强，产生漏报，且计算量大，影响系统效率；

（2）过少，则普遍性强，产生误报。

因此，特征选取就是一种策略，其目的就是降低漏报率和误报率以及提高系统性能方面找到一个最佳切合点。

以下报文头部值和报文数据段的内容常用来作为网络入侵检测系统的特征选项，如表 1 所示。

1.2 特征描述

采用当前使用最普遍的 Snort 规则格式来描述特征，一条 Snort 规则可以分为前后两个部分，规则头和规则选项：

规则 = 规则头 + 规则选项

规则头包含规则的操作（rule's action）、协议（protocol）、源 IP 地址和目标地址及其网络掩码及源端口和目标端口值等信息[3]。规则选项包括报警信息以及用于确定是否触发规则响应动作而需要检查的数据包

① 基金项目：国家自然科学基金面上项目（60673165）；湖南省自然科学基金（05JJ30119）
基金资助：湖南省科技计划项目（2006 JT1040）

区域位置信息[4]。例如，下面就是一条规则：

alert tcp any any - > 192. 168. 1. 0/24 111 (content: " 100 01 86 a51 " ; msg: " mounted access" ;)

这条规则描述了以下信息：任何外部网络主机连接 C 类网络映射端口的数据包中，如果出现二进制数据 00 01 86 a5 便发出警告信息"mounted access"。

表 1　特征选项

特征选项	说明
Protocol ID	协议码，指明该数据报采用的协议(ICMP =0，UDP =1，TCP =2)
Source Address	源地址，数据报源 IP 地址
Destination Address	目的地址，数据报目的的 IP 地址
Source Port	源端口，数据报源地址的端口值
Destination Port	目的端口，数据报目的地址的端口值
Flag	标志位，用于指明 TCP 数据报标志位的置位方式
ICMP Type	ICMP 数据报的类型码，包括应答包、请求包等
ICMP Code	ICMP 数据报的代码，是对类型码的细化
Head Length	报文头部长度
Packet Length	数据报长度，过大过小都可能是恶意的
Data Portion	报文数据段内容，一般取前 30 个字节
Check Sum	数据报头部校验和，恶意的数据报校验和是不正常的

从规则开头到圆括号前的部分是规则头，包含在圆括号中的部分是规则选项。规则选项部分中冒号前面的词组称为选项关键字 (Option Keywords)。这里需要指出规则选项不是对每个规则都必需的部分，它只是用来定义收集特定数据包的特定特征。只有当一条规则中不同部分同时满足时，才能触发对应的规则动作。每个规则的不同部分相当于一种"逻辑与"关系。而同一个规则数据库文件中的所有规则之间相当于一个"或"操作[5]。

(1) 规则头。包含了定义数据包"从哪里来，到什么地方去、干什么"以及发现满足这个规则所有条件的数据包时应该干什么的信息。规则的第一项是规则操作，第二项是协议，第三项是 IP 地址和端口。

(2) 规则选项。规则选项是检测系统的核心，具有易用性和灵活性。不同选项之间使用分号"；"分隔开来，他们之间为"与"的关系。选项有关键字和参数组成，每个关键字和他的参数使用冒号"："分隔。下面是几个关键字的解释，如图 1 所示。

• msg

msg 规则选项告诉日志和警报引擎所需要与数据包一起转储的消息或者需要加入到警报信号中的消息文本信息，该消息是一条简单的文本。

格式：msg：" < message text > " ;

• logto

logto 选项用来通知系统将触发该规则的所有数据包记录到一个指定的输出日志文件中。

格式：logto：" < filename > " ;

• ttl

该规则选项用来测试某一特定 TTL(Time - To - Live) 值。该选项的测试只有在精确匹配时才为真。此选项关键字主要用于对 Traceroute 试探活动的检测。

格式：ttl：" < number > " ;

• id

此选项关键字用于精确匹配 IP 数据包分组的 ID 字段值。某些黑客工具设定该字段为用于各种目的的特殊值，例如，字段值 31337 的设定在某些入侵手法中非常普遍。

格式：id：" < number > " ;

图 1　主要关键字的解释

(3) 规则举例。Land 攻击是针对 CVE - 1999 - 0016 漏洞进行的一种 DOS 攻击，发送一个源地址和目标地址相同，源端口和目的端口也相同的 SYN 包到有漏洞的目标系统，由于 TCP/IP 实现上的问题，目标系统对此种畸形包的处理可能会出问题。规则如下：

alert tcp $ EXTERNAL_NET any - > $ HOME_NET any (msg: " DOS Land attack" ; id: 3868 ;

seq: 3868 ; flags: S ; reference: cve, CVE - 1999 - 0016 ; classtype: attempted - dos ; sid: 269 ; rev: 2 ;)

Teardrop 攻击是针对 CVE - 1999 - 0052 漏洞进行的一种攻击方法，这种攻击由两个相重叠的 IP 碎片组成。第一个是很大的碎片，第二个碎片很小并且会填充到第一个碎片中 (它起始于较后的偏移量但是结束在第一个碎片之前)。这样在易受攻击的系统中的 IP 重组过程将受到冲突的偏移量的混淆并试图采用负数

给新的数据分配内存，从而引起系统崩溃。可以加入如下规则：

alert udp $EXTERNAL_NET any -> $HOME_NET any (msg:"DOS Teardrop attack";id:242;fragbits:M;reference:cve,CAN-1999-0015;reference:url,www.cert.org/advisories/CA-1997-28.html;reference:bugtraq,124;classtype:attempted-dos;sid:270;rev:2;)

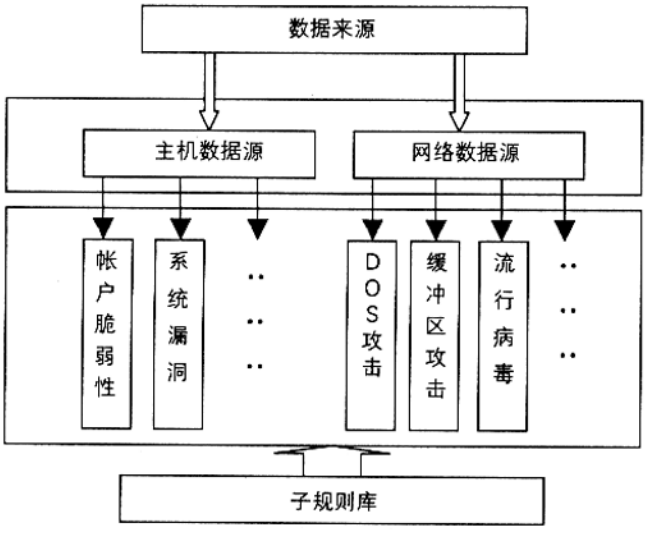

图 2　规则库结构

2　规则库结构设计

规则库采用模块化的结构设计，通过对攻击方法进行系统的分类，将攻击方法进行逐级划分范围，再根据不同的攻击类别建立各个子规则库模块. 例如，DOS 攻击可作为攻击类别中的一大类，其中又包括了 SYN FLOOD、DDOS(Distributed Denial of Service Attack，分布式拒绝服务攻击)等攻击方法。

由于本系统是混合的入侵检测系统，在规则库的上层根据需要检测数据的来源将攻击分为两个方面：来源于网络的和来源于主机的，这样可以使推理机在一个更准确的范围内搜索规则，大大减少了匹配过程的耗时。接下来对于分流后的数据根据攻击方法研究分类，建立子规则库。子规则库将不同类型的攻击特征编写成规则放入各个单独的模块中。同样，以 DOS 攻击为例，在这里可以建立 DOS 攻击子规则库来将它的攻击特征编写成规则放入该模块中。本系统建立了下面这些子规则库：DOS、DDOS 等，该规则库的子规则库是以"rules"为扩展名结尾的，它们用一般的编辑器就可以编辑。各个子规则库描述的是具体各种类型入侵的行为特征。规则库结构如图 2 所示。

3　规则解析

3.1　规则解析

规则解析要先读取规则文件，并读取每一条规则，然后对没条规则进行解析，并用相应的规则语法表示[6]。

系统将检测规则组织成一个二维的链表结构，主要分为两个部分：链表首部和链表选项。链表首部组成主链，然后根据链表选项把规则插入到这个链中，构成规则树。在链表头中包含的多个规则中的共有属性，而不同的检测属性选项则包含在不同的链表选项中。如果在一个规则文件中指定了 45 条检测 CGI-BIN 探测活动的规则，而他们都具有系统的源/目的 IP 地址以及端口号。为了加快检测的速度，这些共同属性就会压缩到一个单独的链表头中，而每一个不同的检测属性将在与表头相连的各个链表选项结构中保存。如图 3 所示为规则链的逻辑结构图。

3.2　规则匹配

规则匹配的过程就是对从网络上捕获的每一条数据报文和规则树进行匹配的过程。如果检测到一条规则匹配这个报文，就表示检测到一个攻击，然后按照规则指定的行为进行处理，如果搜索完所有的规则都没有找到匹配的规则，就表示报文是正常的报文。

规则被组织成规则树，分类存放在规则类列表中。规则检测其实就是对规则树进行匹配扫描，并找到报文所对应的规则。对规则树的匹配过程是先根据报文的 IP 地址和端口号，在规则头链表中找到相应的规则头，找到后再接着匹配此规则头附带的规则链表。

完成规则匹配工作，规则库的实现已经完成。

4　结束语

入侵检测系统是一种主动的防御技术，构建攻击特征库是其核心组成部分。而特征提取是构建特征库最重要的一个步骤[7]，与入侵检测系统的效率和准确性紧密相关。本文阐述了特征库的设计思想和实现过程，提出了提高检测效率的层次结构的规则库。而随

着网络的发展，攻击手段也越来越多，为了减少由于攻击手段的增加所造成的漏报，可以及时添加新的检测规则，更新规则库。同时，也出现了更为先进的协议解析技术，采用模式匹配与协议解析技术相辅助，更能够提高 IDS 的检测效率。

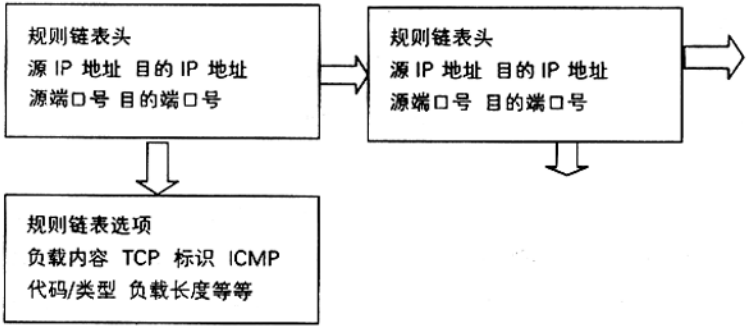

图 3　规则链逻辑图

参考文献

1 温世强、段海新、吴建平，开放式网络攻击特征库的设计与实现[J]，小型微型计算机系统. 2006，27(1)：22～25.

2 唐正军、李建华，入侵检测技术[M]，北京：清华大学出版社，2004. 27～28.

3 齐建东、陶兰、孙总参，入侵检测工具_Snort 剖析[J]，计算机工程与设计，2004，25(1)：36～38.

4 陈海涛，基于多代理的入侵检测系统的实现技术研究[硕士学位论文]，长沙：国防科学技术大学，2002.

5 唐正军，网络入侵检测系统的设计与实现[M]，北京：电子工业出版社，2002，337～347.

6 韩东海、王超、李群，入侵检测系统及实例剖析[M]，北京：清华大学出版社，2002. 107～108.

7 胡威、李建华、陈波，入侵检测建模过程中特征提取最优化评估[J]，计算机工程，2006，32(12)，150～151.