统一入侵检测报警信息格式提案及其实现*

裴晋泽[1,2]，肖枫涛[1]，胡华平[1]，黄辰林[1]
(1.国防科学技术大学 计算机学院，湖南 长沙 410073；2.中国人民解放军 92493 部队，辽宁 葫芦岛 125000)

摘　要：为增强 IDS 之间信息共享和交换的能力，加强 IDS 之间的交流和协作，给出了统一入侵报警信息格式的详细提案，并提出了用 XML Schema 对报警信息建模的方案，最后用 XML 描述语言实现了该提案并通过了 XML Schema 的有效性验证。所提出的统一入侵检测报警信息格式提案给不同 IDS 之间和 IDS 不同组件之间提供了信息共享和信息交换的平台，对于增强 IDS 之间以及 IDS 和其他安全设备的协作能力具有十分重要的意义。

关键词：入侵检测；报警信息；可扩展标记语言；模式
中图法分类号：TP393.08　　文献标识码：A　　文章编号：1001-3695(2006)02-0107-04

Design and Implementation of Unifying Alert Information Format Scheme

PEI Jin-ze[1,2]，XIAO Feng-tao[1]，HU Hua-ping[1]，HUANG Chen-lin[1]
(1. *School of Computer Science, National University of Defense Technology, Changsha Hunan* 410073, *China*; 2. *The* 92493 *Unit of PLA, Huludao Liaoning* 125000, *China*)

Abstract: In order to improve the ability of information share and exchange, and strengthen the communication and collaboration between IDSs, the Unifying Alert Information Format Scheme(UAIFS) is presented, and the alert information modeling method with XML Schema is also presented. At last, the UAIFS is implemented with XML, and its validation is proved by XML Schema. The UAIFS can provide the ability of sharing and exchanging information among IDSs and IDS's components. So it is significant for IDSs and other security equipments to exchange the collaboration ability.

Key words: Intrusion Detection; Alert Information; eXtensible Markup Language; Schema

1　引言

随着网络攻击手段向分布和协同方向发展，对于传统的网络安全防护提出了巨大的挑战，这就使得作为动态安全技术最核心技术之一的入侵检测系统(IDS)之间以及入侵检测组件之间具备共享与攻击相关信息的能力非常重要。这种信息共享是入侵检测合作的基础，同时，真正的网络安全还要求 IDS 能够与访问控制、应急、入侵追踪等系统交换信息、相互协作，形成一个整体有效的安全保障系统。为了提高 IDS 产品、组件及与其他安全产品之间的互操作性，美国国防高级研究计划署(DARPA)和互联网工程任务组(IETF)的入侵检测工作组(IDWG)发起并制定了一系列建议草案。

CIDF[1]是理论界标准，目的是解决不同入侵检测系统互操作性和共存问题。但是 CIDF 的事件描述语言 CISL 并不适合于描述复杂事件的详细情况，特别是 IDS 产品的开发商也未对 CIDF 表现出多少兴趣，CIDF 的后续影响将很有限。IDWG[2]是商业界标准，目的是制定 IDS 之间、IDS 和网管系统之间共享的数据格式和统一的通信规程。比较而言，IDWG 的解决方案相比 CIDF 要适用得多。但是，IDWG 只集中在过于简单的报警数据的相互交换上，不完全如其章程所说的“入侵检测信息交换”；而且其描述入侵报警的数据模型更倾向适用于网络型入侵检测系统，对主机型入侵检测系统的支持相对要弱一些，因而可以用其交换的数据量比较有限。

目前 CIDF 和 IDWG 都还不成熟，需要不断地改进和完善，但可以预见，标准化是未来入侵检测系统发展的必然方向。

收稿日期：2004-11-20；修返日期：2005-03-30
基金项目：国家“863”计划资助项目(2003AA142010)

2　IDS 报警信息格式现状

入侵描述所关注的是互联网的安全，它应该具有被广泛接受的可能和基础，即报警信息的统一化描述应该是在分析各种目前被广泛采用的 IDS 产品报警信息格式基础上得出的。本文的统一报警信息格式提案正是针对这一关键点展开。

通过收集并从分析几种有代表性的 IDS 报警信息格式，如免费 IDS(TCPdump, Snort 和 Portsentry)报警信息格式、GIAC(Global Incidents Analysis Center，全球事件分析中心)的信息格式、SecurityFocus(www.securityfocus.com)的事件格式以及目前使用比较广泛的商业入侵检测系统(Dragon, RealSecure, NFR)报警信息格式，可以得出以下结论[3]：①报警消息内容各异，对于报警事件的描述过于简单或过于复杂，不具备可操作性；②报警格式各异，不便于识别和理解，通用性差；③各厂家 IDS 产品出于各自的商业利益，可扩展性不强，可移植性差。

3　报警信息建模

3.1　采用 Schema 对报警信息建模的可行性分析

在信息建模中，虽然 UML 建模方法在技术精确度方面有

着明显的优势,但是针对所要建模对象的实际情况——报警信息本身内容简单、报警信息数据之间的关系不复杂,只有包含和被包含的关系等特点,其建模的目的是为了便于生成 XML 文档,是一种较小规模的建模。本文采用 XMLSPY 工具提供的可视化的 Schema 树型结构来进行建模。

3.2 报警信息模式模型(Alert Information Schema Modeing, AISM)

Alert 是 XML 文档中的根元素,是 AISM 的父类,它包含三个属性:消息类型标志 MessageClassID(实数型)、报警唯一标志 AlertID(字符串型)和报警发送者估计此事件可能对应的响应优先级别 ResponsePriority(二进制枚举型,000 ~110 表示攻击后响应,111 表示攻击进行中响应)。

为了保证在某入侵检测环境中所有交换的报警消息都能够被唯一地标志,AISM 为每个报警消息打上一个标记 AlertID,这个标记由两部分组成:①在所属的入侵检测环境中,赋予每个报警消息的发送者一个唯一的标志,如 AgentID;②每个发送者对其发送的所有消息赋予唯一的编号。合成后的标记就可以唯一地表达报警标志。

另外,在网络安全防护中,及时发现问题和及时解决问题是一对不可分割的整体,对于必须并且可以作出及时反应的报警,决不能延误时机,因此本文在报警信息中特别添加了事件响应优先级 ResponsePriority 作为联系 IDS 和实时响应系统的纽带。消息类型标志 MessageClassID 则是为了区分非报警消息而设。

依据标准化报警信息格式提案要求和一般建模原则,将 Alert 类细分为五个子类,它们之间的关系如图 1 所示。

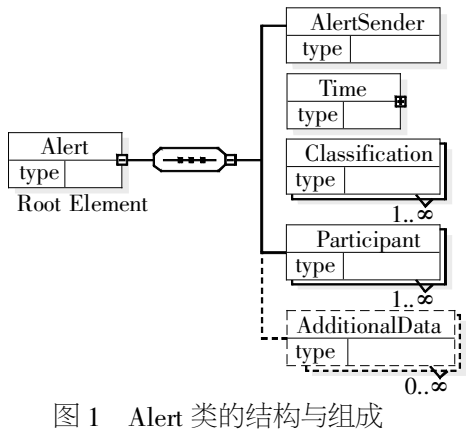

图 1 Alert 类的结构与组成

(1) AlertSender

它用来描述报警消息的发送者,如分析器,且有且只有一个。它包含一个属性 AgentID(字符型),也可以为 AnalyserID,其意义是相同的。

(2) Time

它用来描述与报警相关的日期和时间信息。Time 下包含三个子类:CreateTime, TimeStamp 和 EndTime。CreateTime 是入侵检测系统的分析器检测到可疑事件的时间。TimeStamp 时间戳的作用是更加准确地记录事件,因为在某一秒内很可能发生许多事件,为每个报警信息设置一个唯一的时间戳(按时间先后增量分配)就不会造成混淆。其中 CreateTime, TimeStamp(整型)为必须有的时间,而且只能有一个。EndTime 是考虑可能会对报警信息进行关联处理的需要而设,故为可选。分别为 CreateTime 和 EndTime 定义一个字符串类型的属性 DateTime,采用 ISO8601: 2000 表示日期和时间的格式,即将日期和时间采用一个字符"T"来连接形如 YYYY-MM-DD 的日期表达和 HH: MM: SS 的时间表达。

(3) Classification

它用来描述报警信息对应的攻击类型。Classification 包括两个可选的属性:CVE(Common Vulnerabilities and Exposures)和 EventType,缺省值为 0。目前 CVE 是被广泛认可的标志和分类,CVE[6] 使得漏洞表示方法标准化,这有助于数据共享,有助于 IDS 的互操作性(Interoperability)、报告的一致性,便于 IDS 间的比较。当然,用户可以根据需要选择某种特定的攻击分类法,为使 AISM 更具一般性,我们并不具体指定,而是由用户视特定的需要而定。

(4) Participant

在报警信息中必然涉及到入侵或滥用活动中实施者和实施对象这两种角色。考虑到如果对攻击源 Source 和攻击目标 Target 分别定义一个类,两者会有很多重叠的部分,因此本文定义一个统一的类 Participant 来描述这两种不同的角色,而用一个属性 Role 来区分这两种不同参与者的角色。这样不仅涵盖了更多的内容,而且简化了数据模型。属性 Role 类型为 Enum(枚举),枚举值为 Source 和 Target,由于存在源和目标都不能确定的可能,Source, Target 可以有零到多个。

源和目标包含的内容因不同 IDS 对入侵源的表示而各不相同,有网络地址、主机名、用户名甚至是进程。例如,基于网络的 IDS 提供攻击源的 IP 地址,而基于主机的 IDS 则提供主机名或用户名。可能用于描述报警实体的还有网络节点、用户账号、进程 ID 以及服务,而对于基于主机的入侵检测系统还需要涉及文件对象。这样 Participant 类必须包含描述这些不同类型数据的子类,因此需要定义 Node, User, Process, Service, File 等类来描述和定位报警源及目标实体。

①Node 用于描述网络上的主机或其他网络设备,如交换机、路由器等。Node 需要表达的信息应该包括网络节点当前所用的地址类型以及放置的位置等。鉴于网络地址的种类繁多,这里为 Node 定义一个子类 Address 来具体描述节点的地址,它可以描述网络地址、硬件地址或应用地址。Address 的属性 Category 有三种枚举类型:IPv4, IPv6 和 MAC, Address 的另一个属性 address 分别表示对应这三种类型的地址。

②User 说明了标志用户的不同方式,它包含 UserID 和 UserName 两个属性,表达两种标志用户的方式,即分别用 ID 号和名字来表示用户或用户组。

③Process 类描述正运行于 Participant 或 AlertSender 上的有关进程的信息,这些信息本身都比较简单,可以直接采用属性 ProcessName 和 ProcessPath(可选)来表达。

④Service 类用于描述源和目标上的有关网络服务的信息。通常来说,服务就是网络上可用的资源,它包含服务编号 ServiceID 和对应的端口号 Port 两个属性。当 Service 作为 Source 角色的 Participant 类的聚集类时,应该理解其为发起服务请求的一方,而且是 Node, User 和 Process 的补充信息。同样,当 Service 作为 Target 角色的 Participant 聚集类时,自然应理解其为提供服务的一方。Service 类是数据模型中唯一包含有关端口信息的类。为便于描述网络中经常遇到的端口扫描攻击,本文定义端口表达方式如下:any 表示任何一个端口;1:1024表示 1 到 1024 端口;:6000 表示小于 6000 端口;500: 表

示小于 1024 到大于或等于 500 的端口;!6000:6010 端口范围可以使用非的功能,表示除 X Windows 端口外的任何端口。

⑤File 类主要用于描述当 Participant 类的 Role 属性为 Target 时可能需要包含的有关文件的信息,这类信息在基于主机的入侵检测系统中常常需要用到。File 类包含的信息本身也比较简单,可以直接采用属性来表达,包括文件名 FileName,文件存储路径 FilePath,文件所有者 FileOwnerID 等,三者类型均为 String。

(5) AdditionalData

为满足对数据模型扩展性的要求,定义一个单独的类 AdditionalData(零到多个)来供 IDS 开发方或使用者需要时描述附加信息。这就使得用户在无须对数据模型本身进行扩展的情况下就可以表达附加的信息,它包含一个 Byte 型的属性 Data,以描述诸如 IP 数据包的头部信息、不同系统的审计日志等附件信息。

综上所述,可以得到 AISM 模型图,如图 2 所示。

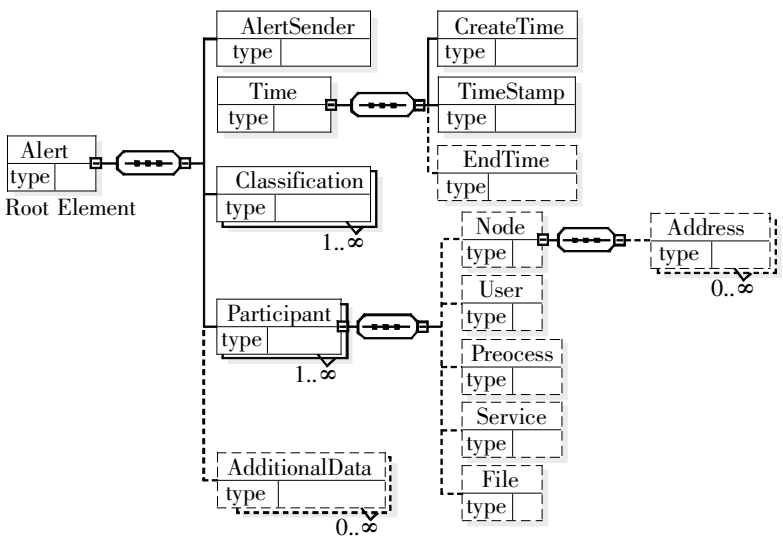

图 2　AISM 模型图

3.3　AISM 的扩展

以上 AISM 包含的报警信息内容是一个比较完整的报警消息的普遍内容,然而对于有不同需求的 IDS 用户,应该允许 IDS 厂商针对特殊客户的特殊需求提供不同程度涵盖以上内容的报警信息,或在必要时进行相应的扩展。

本文提供两种方式来扩展 AISM 的表达能力,即扩充数据类型定义和扩展数据模型。扩充数据类型定义主要是扩充消息的附加数据;扩展数据模型主要是扩充继承关系和聚集关系。AISM 的两种扩展方式为:①定义一个新类添加到数据模型中已存在的类,它类似于 UML 中的聚集关系,如将 Node 和 User 定义为 Participant 的下属类就是这种方式的一个例子。这种方式允许对所有的类扩展提供更多的信息。②细化模型中已定义的类,它类似于 UML 中的继承关系,如可以细化 Service 类得出 WebService, HostService 等下属类。这种方式是扩展报警信息表达方式的首选,因为继承扩展方式不仅能维持原有报警信息的数据结构,而且还可以维持原有的对数据结构的操作。

4　基于 XML 的报警信息格式提案的实现

1998 年 W3C(World Wide Web Consortium)正式公布了 XML 1.0 规范[4]。XML 是一种元语言,即用来描述其他语言的语言,它可以让用户定义自己的标记语言,从而在 XML 文件中描述并封装数据。此外,XML 还具有严格的规范、清晰的语义、数据的共享与重用、良好的扩展性、平台无关性等特点(具体可参见文献[5]),基本符合标准化报警信息格式的要求。因此,与 XML 成为越来越多 Internet 应用首选的理由相同,XML 的灵活性或者说是可扩展性以及易用性使其成为实现报警信息交换格式的首选方式。

4.1　AISM 到 XML 的映射

由于 XML Schema 本身就是符合 XML 语法规范的 XML 文档,使得建立从 AISM 到 XML 的映射非常简单。

(1) 表示对象类型

AISM 中的对象类型就是相应 XML Schema 中元素类型,因此可平滑映射到 XML 文档中的元素类型。

(2) 表示关系

在 AISM 中,聚合关系的聚合端用　　表示,对应到相应的 XML Schema 中就是元素和子元素的嵌套关系,也可以平滑映射到 XML 文档中包含的元素和重复的子元素。

(3) 表示属性

AISM 中类的属性和 XML 文档中元素的属性表示是完全对应的,因此无须转换。

4.2　报警信息的 XML 代码实现

建立从 AISM 到 XML 的映射比较简单。下面给出端口扫描的报警格式的信息编码(对应文件为 Alert_message.xml):

```
<?xml version="1.0" encoding="UTF-8"?>
<Alert xmlns="http://www.Alert_message.com/ids/alert"
Alert AlertID="rs0156785" ResponsePriority="100" MessageClassID="0">
    <AlertSender AgentID="rs01">
    </AlertSender>
    <CreateTime DateTime="2004-03-09T15:31:07Z">
    </CreateTime>
    <TimeStamp timestamp="66712">
    </TimeStamp>
    <Classification EventType="dns"> </Classification>
    <Participant Role="source">
      <Node>
        <Address Category="ipv4" address="192.168.100.1">
        </Address>
      </Node>
    </Participant>
    <Participant Role="target">
      <Node>
        <Address Category="ipv4" address="192.168.102.1">
        </Address>
      </Node>
      <Service ServiceID="9" Port="5:25,37,42,43,53">
      </Service>
    </Participant>
    <AdditionalData>
      <IPhdr saddr="192.168.100.1" taddr="192.168.102.1" proto="TCP">
      <TCPhdr> sport="10033" dport="5" flags="1" </TCPhdr>
      </IPhdr>
    </AdditionalData>
</Alert>
```

5 报警信息 XML 文档的有效性验证

5.1 XML 文档有效性验证的必要性

完全遵循 XML 规范的正确格式的文档并不总能满足需要。许多情况下还需要保证文档的有效性,否则在应用程序中对 XML 文档进行操作时,XML 代码中的错误会引起其他程序的中断,或者导致错误的数据进入系统。

5.2 报警信息 XML 文档的有效性验证

由于在建模阶段已经分析清楚文档元素的名称、结构和属性,这使得编写 XML Schema 非常容易。并且基于 AISM 模型,使用工具 XMLSPY 还可以自动生成 XML Schema 文档 Alert_message.xsd。采用 XMLSPY 验证工具执行的验证操作步骤如下:

(1) 将要检验的 Alert_message.xml 文件加载到 XMLSPY 中;

(2) 从“XML”菜单选择“Validate”;

(3) 执行验证操作。如果发现有效性错误,状态栏将予以指示并提供详细说明,否则,系统提示“This file is valid”而通过验证。

凭借 AISM 的优势,XML 示例文档 Alert_message.xml 一次性通过了其对应的 Alert_message.xsd 文档的有效性验证,其界面如图 3 所示,(图 3 左下角“√”标记表示通过有效性验证)。

图 3 XML SPY 效验工具及效验结果图

6 结束语

针对黑客分布式攻击的新特点,为增强 IDS 之间信息共享和交换的能力,加强 IDS 之间的交流和协作,本文给出了统一报警信息格式的详细提案,并提出了用 XML Schema 对报警信息建模的方案,最后用 XML 描述语言实现了该提案并通过了 XML Schema 的有效性验证。本文的研究成果已经在国家“十五”“863”项目分布式网络监控与预警系统中得到了应用,并取得了较好的效果。

参考文献:

[1] taniford-Chen S, *et al*. The Common Intrusion Detection Framework (CIDF)[C]. Orlando, Florida: ISW'98 Workshop, 1998.

[2] D Curry, H Debar. Intrusion Detection Message Exchange Format Data Model and Extensible Markup Language (XML) Document Type Definition[EB/OL]. http://xml.coverpages.org/draft-ietf-idwg-idmef-xml-10.txt, 2002-02.

[3] 裴晋泽. 基于 IDS 的网络安全预警系统关键技术研究与实现[D]. 长沙: 国防科技大学, 2004.

[4] World Wide Web Consortium (W3C). Extensible Markup Language (XML)[EB/OL]. http://www.w3.org/TR/1998/REC-xml-19980210, 1998-02.

[5] Fabio Arciniegas. XML 开发指南[M]. 天宏工作室. 北京: 清华大学出版社, 2003. 230-235.

[6] CVE Editorial Board. The Common Vulnerabilities and Exposures [EB/OL]. http://www.cve.mitre.org, 2003-09.

作者简介:

裴晋泽,硕士研究生,研究方向为信息安全;肖枫涛,博士研究生,研究方向为信息安全;胡华平,教授,博导,研究方向为信息安全。

(上接第 106 页) 企业同样要关注容灾的流程、规范及其具体措施。作为金融行业,管理者更应该明确本单位的关键职能,更要评估灾难发生时所造成的风险以及潜在的影响,因此要制定合理计划来减少这些风险和影响,也要制定规章制度来保证整个灾难恢复的顺利实施,这正是绝大多数企业所缺乏的[4]。

4 未来容灾技术的发展展望

灾难备份已由原来的磁带备份技术慢慢转向了磁盘镜像技术,由单机备份转向网络备份,备份数据中心也由冷备份逐渐向热备份转换,灾难恢复级别的要求也越来越高。目前,建设连续可用系统、保证业务系统可持续性操作、更好地为用户提供服务是许多企业所追求的目标,通过容灾尤其是远程应用级容灾,建立无数据丢失、灾难发生时能够自动切换的数据中心来保证业务系统的连续可用是未来容灾的发展方向。但远程应用级容灾的研究目前刚处于起步阶段,相关的技术与文档比较少,实现起来非常困难。但它的重要性是不言而喻的,是建设业务连续可用系统的基础,是未来容灾的发展方向,目前需要一套完整的技术理论来支持。

参考文献:

[1] ttp://www.longterm.cn/tech/article_content.asp?id=337, 2005-01-03.

[2] 周世超,郭利江. 远程应用级容灾系统(RALDRS)模型研究[EB/OL]. http://www.chinabyte.com/20030321/1658579.shtml, 2003.

[3] http://media.ccidnet.com/media/ciw/1267/c1401.htm, 2005-01-15[EB/OL].

[4] Jon William Toigo. Disaster Recovery Planning [M]. 北京: 电子工业出版社, 2004. 7-10, 104-106.

[5] Andries J, M de Jong. Extended Remote Copy[R]. 2004. 82-104.

作者简介:

岳友宝,男,硕士研究生,主要研究方向为灾难备份与恢复技术;张艳,女,博士研究生,研究方向为计算机网络与信息安全、灾难备份和恢复技术;李舟军,男,教授,博士生导师,博士,主要研究方向为进程代数理论、安全协议的形式化验证、灾难备份与恢复技术。