基于 Agent 的分布式入侵检测系统通信机制设计①

Communication Mechanism Based on XML Designed for Distributed Intrusion Detection System

费洪晓　倪 敏　谢文彪　戴宏伟　裘方敏
（中南大学信息科学与工程学院 湖南长沙 410075）

摘要：分布式环境下的入侵检测系统是入侵检测的研究热点，系统中各 Agent 运行于不同的平台上，具有不同的数据表达格式，这需要有效的通信机制保障 Agent 间的通信与协作。本文在入侵检测标准化组织现有文档的基础上，提出了一个具有通用性的分布式入侵检测系统通信机制框架。Agent 间的消息交换格式参照 IDMEF 标准，并根据入侵检测 Agent 通信需求扩充了警报数据的 XML 描述。框架还给出了 Agent 通信安全机制，使通信机制总体上满足了分布式入侵检测系统警报信息量大、实时通信、安全性高的特点。

关键词：分布式入侵检测　通信机制　IDMEF　XML

分布式入侵检测系统中，由多个检测 Agent 分别检测不同的主机和网络，各 Agent 间需要通过互相协作来完成较复杂的检测任务[1]。然而，各检测 Agent 可能使用不同的检测方法，运行于不同的平台上，具有不同的数据表达格式，这就增加了 Agent 间进行协作的复杂性[2]。因此，在分布式入侵检测系统中，需要一种通用且高效的入侵检测通信机制。

互联网工程任务组（IETF）的入侵检测工作组（IDWG）制定的入侵检测消息交换格式（IDMEF）、入侵检测交换协议（IDXP）、入侵报警（IAP）和美国国防部高级研究计划局（DARPA）制定的公共入侵检测框架（CIDF）等标准提供了一个入侵检测通用的体系结构、入侵消息和入侵对象的表达格式、有效通用的通信协议[3][4]。借鉴以上标准化成果，为了解决分布式入侵检测通信存在的现有问题，在较为深入的研究 XML 文档解析和存取的基础上，提出了一个具有通用性的基于 XML 的分布式入侵检测系统通信机制。文章详细给出了 Agent 之间的通信协议、消息格式以及通信安全机制。该通信机制较好地解决分布式入侵检测系统各 Agent 间的通信协作问题，有利于更有效地发现入侵。

1　基于 Agent 的分布式入侵检测系统结构

为了克服当前入侵检测系统存在的一些不足，本文提出了一种基于 Agent 的分布式入侵检测系统，如图 1 所示。该模型的主要特点是减少网络流量，平衡网络负载，提高容错度，异步交互，增强动态管理和系统配置的灵活性。

（1）感应 Agent 感应 Agent 从所监控的目标环境中广泛地收集各种数据。感应 Agent 获取原始数据后，对数据进行简单的预处理后将数据转换成统一格式，上传给分析 Agent 作进一步分析。

（2）分析 Agent 分析 Agent 在管理 Agent 的控制下向特征库获取信息，对感应 Agent 上传的数据进行分析，检测攻击行为或系统异常。它还负责将向管理 Agent 发送描述入侵的警报信息。

（3）管理 Agent 管理 Agent 负责协调各分析 Agent 之间的关系，它维持有一个注册表，表中记录了当前主机中各分析 Agent 的相关信息。还负责更新行为、特征数据库中的信息，接受控制 Agent 的相关

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）

命令。

(4) 控制 Agent 控制 Agent 运行于服务器上，负责协调控制整个系统，包括向管理 Agent 发送配置和控制指令以及接收管理 Agent 的报告，通过用户界面提供系统动态信息给系统管理员。

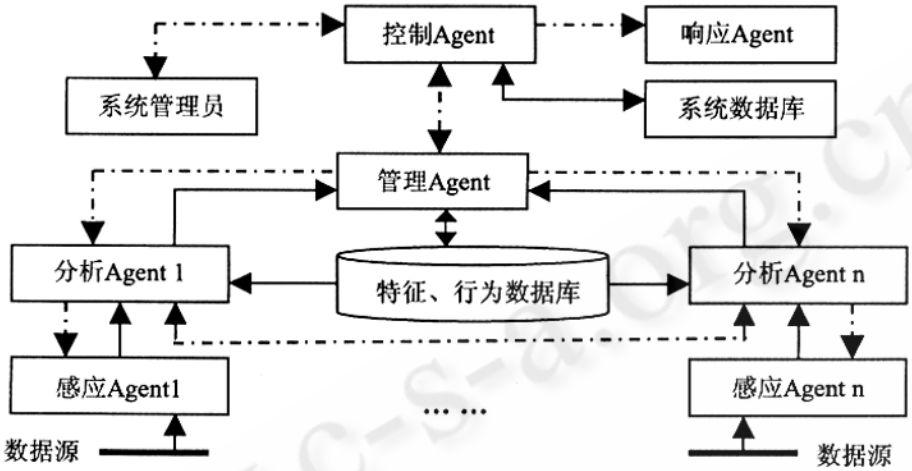

图 1　基于 Agent 的入侵检测系统

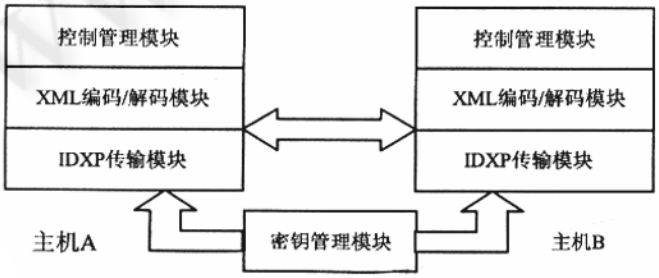

图 2　通信机制的总体框架

2　通信机制的总体设计

通信协议的设计主要是基于 IDWG 所提出的 IDXP 协议，研究并设计适用于基于 Agent 的分布式入侵检测系统的通信协议框架，使之能够支持任意 Agent 之间的加密通信，可信 Agent 双方的相互鉴别和数据的完整性保护。Agent 间的报文交换格式主要参照了 IDMEF 标准，扩充警报数据的 XML 描述，以及查询、应答和响应等控制命令，增强了系统的互操作性。

2.1　通信机制框架结构

通信机制的总体框架主要分为四个模块，分别为控制管理模块、XML 编码/解码模块、IDXP 传输模块和密钥管理模块四个部分。其中控制管理模块的功能包括系统安全配置和日志信息的管理。XML 编码/解码模块主要是实现一个 IDMEF 解释器，对传输数据进行 XML 编码/译码；IDXP 传输模块包括加密、鉴别、完整性保护；密钥管理模块包括会话密钥的协商生成、传递，证书管理等工作。通信机制框架结构如图 2 所示。

2.2　通信协议设计

IDXP 传输层位于传输机制框架的底层，完成 Agent 间警报信息的传输。Agent 之间通过建立会话来传递消息，一个会话建立在一个面向连接的 TCP 协议上。建立连接后，双方要进行身份认证和安全参数协商，用来确保会话的安全性。会话建立后，可以在会话上建立若干的隧道，每一个隧道都可以用来传递如入侵警报、控制信息等特定类型的信息，通信双方可以根据需要来开启或关闭特定隧道，系统通信协议栈如图 3 所示。

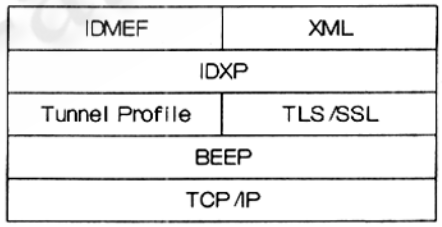

图 3　系统通信协议栈

2.3　系统消息设计

本文提出了一个通用消息模型，按照面向对象思想设计该消息模型的总体框架结构，并采用 XML 设计通用消息格式来封装警报数据，控制命令和配置信息。该模型最高层次类是 IDMEF－Message，分为两个子模型，即报警消息模型和注册消息模型。报警消息模型包括 Alert(警报)、Heartbeat(心跳)两个子类消

息模型。Alert 类消息定义检测器向控制台报告的报警消息的格式;Heartbeat 类消息用于检查探测代理和响应代理网络的状态。注册消息 Register-message 包括 Regmsg(注册)、Regack(注册应答)、Regackconf(注册确认) 三个子类消息。Regmsg 消息定义了代理向控制决策中心申请加入的注册信息;Regack 消息定义了控制决策中心接收到代理的注册消息后所产生的应答消息;Regackconf 消息定义了代理收到控制决策中心的应答消息后所发出的注册确认信息。

3 关键技术的实现

3.1 XML 描述警报信息

XML 文档的内容与数据定义相互分离,数据类型定义便于语义理解和语法检查。IDWG 定义的 IDMEF 以面向对象的方式通过继承和聚集来表示警报数据之间的层次关系。另外,随着入侵检测系统的发展,IDMEF 提供两种方式来扩展表达能力,扩展数据模型和数据类型定义。

其中,XML DTD 用于定义 XML 文档的语法和结构,规定文档允许出现的元标记以及它们出现的顺序、标记可包含的其他标记等。由于 XML DTD 缺乏对 XML 文档的内容及其语义的约束机制,这将限制 XML 处理器进行有效的类型检验。相对于 XML DTD 而言,XML Schema 对 XML 文档给出了更强的语法约束,并增强了语义约束。

XML Schema 提供了对命名空间的支持,具备更丰富的数据类型、数据结构,支持了数据对象间的继承关系。本文基于 XML Schema 实现了对 IDMEF 的改进,使得 IDMEF 支持了数据对象的继承关系。

3.2 XML 消息的解析与存储

(1) XML 消息的解析。当管理 Agent 收到分析 Agent 产生的 XML 消息后,希望将此 XML 消息内容写入数据库,以便进行进一步处理、分析。这里就涉及到 XML 消息的解析。

有两种解析 XML 的方式,分别是 DOM 和 SAX。使用 DOM 的好处是可以引用和操作每一个对象。和 DOM 不同的是,SAX 是基于事件的,这意味着当它在一个 XML 文档中发现特殊的符号的时候,它会产生相关的事件。SAX 不适合于处理包含很多内部交叉引用的文档,不能实现复杂的搜索。在入侵检测系统中,由于要处理的警报消息文件不大,同时希望能够灵活的提取其中的信息,因此选择 DOM 对 XML 文件进行解析。

(2) XML 消息的存储。XML 数据的存储主要有 3 种方式:直接使用文本文件、使用关系数据库和使用对象管理器。文本文件方式最简单,但是每次浏览或查询警报信息时需要重新分析整个 XML 文件;使用关系数据库需要实现 XML-关系映射,利用 DTD 来生成存储 XML 数据的关系模式;对象管理器的方式很适合 XML 的结构,但由于对象的层次一般比较深不便于查询。

在分析了 IDMEF 后发现使用基于栈结构对 IDMEF 数据进行存储和查询是一个比较好的方式。因为 XML 文件本身的结构和堆栈有着很大的类似,XML 栈是指保存 XML 文件结构信息的栈,利用这个堆栈结构可以有效地保存 XML 文件信息,该映射直接由 XML 的语法分析器产生相应的关系数据库中的记录。

3.3 IDXP 传输模块

IDXP 传输模块是一个具有通用性的通信模块,它将经过 XML 编码后的警报信息通过安全传输通道传送给控制 Agent,同时也负责接收来自控制 Agent 的命令信息。该模块以 BEEP 协议核心作为通信框架,BEEP 协议通信双方同时以多个通道(Channel)传输数据,每个通道关联一个轮廓(profile),在轮廓中定义了信息描述、信息安全和身份认证等机制,BEEP 协议通过通道 0 管理其它通道的会话。通信双方在通信开始时通过通道 0 告知各自所支持的轮廓,如果协商成功,就将进行数据传输,否则关闭会话,其过程如下所示。第一步:建立连接。建立连接时,控制 Agent 在固定的端口监听,管理 Agent 发起连接。建立连接后,双方互发送 Greeting 消息来表明身份。对方如果愿意进行下一步协商,就回应一个 OK 消息,否则就发送一个 Error 消息,关闭连接。

第二步:协商安全参数。建立连接后,管理 Agent 发送 TLS-ready 消息。控制 Agent 收到该消息后,

如果愿意协商安全参数，就发送 TLS－proceed 消息，双方通过 TLS 协议来协商安全参数。

第三步：交换 Greeting 消息。协商安全参数完成后，通信双方就可以使用安全的通信信道了。这时双方需要再次发送 Greeting 消息。根据消息中含有发送方的身份信息，接收方可以用来验证是否同建立连接时获得的对方的身份信息相同。

第四步：建立隧道。通信的任何一方都可以发送 Start 消息来申请建立隧道，接收方如果同意建立隧道就发送 OK 消息，否则发送 Error 消息。

第五步：安全的数据传输。通信双方利用已建立的隧道来交换数据。在数据交换过程中，双方仍可以建立新的隧道或关闭旧的隧道。发送方发送 MSG 消息，接收方如果能识别收到的消息，就发送 OK 消息，否则发送 Error 消息。

第六步：结束传输。通信一方在发送的所有数据都得到回应后，可以发送 Close 消息请求关闭其对应的隧道。接收方如同意关闭隧道，就发送 OK 消息，否则发送 Error 消息。如果要关闭整个会话，在 Close 消息中指定的隧道编号应为 0，这样整个会话将被关闭。

3.4 安全通信机制

(1) 会话管理 刚刚建立会话时，双方默认打开隧道 0，利用隧道 0 进行会话管理。通信双方在隧道 0 中交换会话控制消息，完成建立、结束会话以及建立、关闭隧道的协商。

(2) 会话控制消息 通信双方通过交换会话控制消息实现对会话的控制，完成协议中的握手，每种控制消息都有其特定的应答。

(3) 协议消息格式 通信双方通过特定的协议消息来传递会话控制消息和系统消息，协议消息用来封装会话控制消息和系统消息。协议消息的格式如表 1 所示。

表 1　协议消息的格式

消息头	隧道编号	消息编号	消息长度	消息内容

① 消息头：消息头为一个字符串，表示消息的类型。它的值有三个："MSG"、"OK"和"Error"。因此，协议中的消息也就分成三类：MSG 消息，OK 消息和 Error 消息。

② 隧道编号：隧道的编号是一个整数，是通信双方协商好的隧道编号。

③ 消息编号：消息编号为一个整数，表示当前发送的消息的编号。

④ 消息长度：消息内容长度为一个整数，表示消息内容字段的长度。

⑤ 消息内容：消息内容有两大类，一类是会话控制消息，另一类是系统消息。

4　小结

针对分布式环境下的入侵检测系统中各 Agent 运行于不同的平台上，具有不同的数据表达格式，在入侵检测标准化组织现有文档的基础上，提出了一个具有通用性的分布式入侵检测系统通信机制框架。具体设计并实现了分布式入侵检测系统的统一的 Agent 消息交互格式以及 Agent 间的通信协议。框架还给出了 Agent 通信安全机制，使通信机制总体上适用了分布式入侵检测系统警报信息量大、实时通信、安全性高的特点。

参考文献

1　王志敏、覃征、唐文伟等，基于 Agent 的分布式入侵检测系统的研究，小型微型计算机系统，2004，24(4)：703～705.

2　Michael E. Locasto, Janak. Parekh, Salvatore Stolfo. Collaborative Distributive Intrusion Detection. CU Tech Report CUCS－012－04, 2004.

3　H. Debar, D. Curry, B. Feinstein. The Intrusion Detection Message Exchange Format: draft－ietf－idwg－idmef－xml－12.2004.

4　B. Feinstein, G. Matthews, J. White. The Intrusion Detection Exchange Protocol (IDXP): draft－ietf－idwg－beep－idxp－07.2002.