基于多传感器数据融合的入侵检测机制

罗光春，卢显良，张 骏，李 炯

(电子科技大学信息中心 成都 610054)

【摘要】针对特征复杂的入侵方式，设计了一种基于数据融合理论的新型入侵检测机制—DFIDS，结果提高了系统在检测复杂入侵行为时的确定性。DFIDS使用优化的并行分布式检测与决策融合系统模型，可以有效克服传统入侵检测系统因单检测器而在数据采集和分析方面的局限性，从而提高了检测的总体性能。文中建立了DFIDS的理论分析模型，并和传统入侵检测机制进行了对比，结果表明DFIDS在检测准确性方面具有更好的性能。

关 键 词 多传感器； 数据融合； 入侵检测； 性能

中图分类号 TP311 文献标识码 A

A Novel IDS Mechanism Based by Data Fusion with Multiple Sensors

Luo Guangchun，Lu Xianliang，Zhang Jun，Li Jiong

(Information Centre, UEST of China Chengdu 610054)

Abstract This paper presents a novel IDS mechanism based on the theory of data fusion—DFIDS, which is designed, according to intrusion patterns with complex features, to heighten the accuracy of the system while detecting complex intrusion acts. The performance of the whole detecting system is enhanced because DFYDS effectively overcomes the limitations of conventional single sensor detecting system regarding aspects of data collecting and analyzing by using optimized model of the fusion of parallel-distributed detection and decision. In this paper, a theoretical model of DFIDS is established and compared with traditional IDS, which proves that DFIDS is of better performance as to detecting accuracy.

Key words multiple sensors; data fusion; intrusion detection; performance

目前，IDS大都采用中心管理控制平台和检测引擎组成的分布式体系结构。中心管理控制平台基于GUI，用于配置和管理检测引擎，检测引擎分布在需要监控的网段或安装在需要监视的主机上，执行入侵检测[1]。这种分布式结构具有平衡计算工作、灵活配置系统和提高系统并行计算性能的特点，在处理简单攻击时效果较好[2]。但对于字符串匹配弱点攻击、多变shell代码、会话拼接、碎片攻击等特征复杂的攻击方式，传统的IDS在数据分析方面存在不足之处，如单个检测器采集、分析的数据不全面，入侵检测由各个检测引擎独立完成，中心管理控制平台并不具备数据综合分析的功能等[3]，本文引入数据融合技术提出了一种新型入侵检测机制DFIDS。

1 DFIDS的体系结构

DFIDS的体系结构如图1所示，它可作为检测单元布置于系统所需要的环境中。多传感器入侵检测系统包含如下几个部分：1) 传感器，用于原始资料搜集，针对网络系统的不同方面设置，包括系统日志文件、

收稿日期：2002 – 11 – 12
基金项目：国家973项目
作者简介：罗光春(1973 –)，男，在职博士生，讲师，主要从事网络技术和网络安全方面的研究.

SNMP信息、用户资料信息、系统消息、操作命令和网络流量信息传感器等；2) 决策器，对自己的传感器搜集的信息进行决策(图中g_1～g_n表示各决策动作)，并将决策结果(u_1～u_n表示本地决策结果)送到融合中心。同样包括系统日志文件、SNMP信息、用户资料信息、系统消息、操作命令和网络流量信息检测器，不同的检测使用不同的决策策略，常用的传感器包括系统日志文件传感器、网络流量信息传感器、网络数据包传感器等，还应根据需要设置功能、数量不等且与检测器相匹配的决策器；3) 融合中心，对各检测传过来的决策进行融合，做出最终决策评估u。

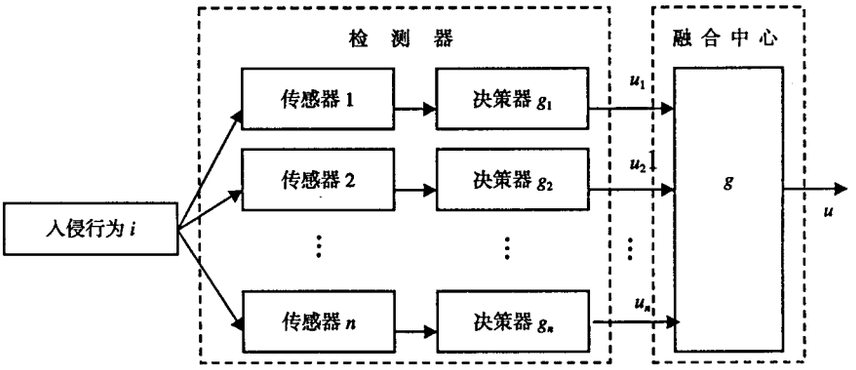

图1 并行分布式检测与决策融合系统

2 DFIDS的理论分析

2.1 决策融合的建模

在入侵检测系统中有m个目标检测类型，每个检测类型分别对应一个检测器。检测器由传感器和决策器组成。第i个检测器对目标的观测矢量为 $\boldsymbol{x}_i = f_i(H), i = 1, 2, \cdots, n$，观测的条件概论密度为 $p(x_i \mid H_j)$，条件联合分布为 $p(x_1, x_2, \cdots, x_n \mid H_j), i = 1, 2, \cdots, n;\ j = 1,2, \cdots, m$。每个检测器$DM_i$根据目标观测矢量 x_i 做出本地决策 $u_i = g_i(x)$。对于二元假设 $U_i \in \{0,1\}$。$U_i = 0$ 表示无目标，支持假设H_0；$U_i = 1$ 表示有目标，支持H_1。n个本地决策$u_1,u_2,\cdots,u_n$传送到融合中心，融合中心根据u_1，u_2，…，u_n，依据某种方法作出全局系统决策 $u = g(u_1, u_2, \cdots, u_n)$[4, 5]。

2.2 二元分布式检测决策融合

二元检测决策融合中心的决策是一个映射|0，1|×|0，1|→|H_0，H_1|，最佳准则函数J：|0，1|×|0，1|×|H_0，H_1|→R，$J(u_1, u_2, H_h)$ 表示H_h为真时检测器DM_1决策为u_1，检测器DM_2决策为u_2的代价函数。融合中心采用最小风险准则，决策规则为

$$\frac{P(u \mid H_1)}{P(u \mid H_0)} = \frac{P(u_1, u_2, \cdots, u_n \mid H_1)}{P(u_1, u_2, \cdots, u_n \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}$$

式中 C_{ij} 表示 H_j 为真但判断为 H_i 的代价。令代价函数 $C_{00} = C_{11} = 0$，$C_{10} = C_{01} = 1$，上式变为

$$\frac{P(u \mid H_1)}{P(u \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}$$

利用Bayes公式[6]，上式等价于

$$\frac{P(u \mid H_1)}{P(u \mid H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

由于 $$P(H_1 \mid u) = \frac{P(H_1, u)}{P(u)} = \frac{P_1}{P(u)} \prod_{i \in S_1} P(u_i = 1 \mid H_1) \prod_{i \in S_0} P(u_i = 0 \mid H_1) = \frac{P_1}{P(u)} \prod_{i \in S_1} (1 - P_{Mi}) \prod_{i \in S_0} P_{Mi}$$

其中，$S_1=\{i\mid u_i=1\}$，$S_0=\{i\mid u_i=0\}$。同理有

$$P(H_0\mid u)=\frac{P(H_0,u)}{P(u)}=\frac{P_0}{P(u)}\prod_{i\in S_1}(1-P_{Fi})\prod_{i\in S_0}P_{Fi}$$

则n个IDS传感器的决策规则为

$$L(u)=\lg\frac{P(H_1\mid u)}{P(H_0\mid u)}=\lg\frac{P_1}{P_0}+\sum_{i\in S_i}\lg\frac{1-P_{Mi}}{P_{Fi}}+\sum_{i\in S_0}\lg\frac{P_{Mi}}{1-P_{Fi}}$$

若令

$$a_0=\lg\frac{P_1}{P_0}$$

$$a_1=\begin{cases}\lg\dfrac{1-P_{Mi}}{P_{Fi}}, & \text{若}u_i=1\\ \lg\dfrac{1-P_{Fi}}{P_{Mi}}, & \text{若}u_i=0\end{cases}$$

则有

$$L(u)=a_0+\sum_{i=1}^{n}(2u_i-1)a_i$$

式中　u_i 为各个检测器的决策结果；$L(u)$ 为融合之后的决策，表示为 u_i 的函数。可以看出，入侵检测系统决策融合的为各个本地决策 u_i 按照其可靠性的加权和。可靠性体现在权值，用常数 a_0、a_1 表示，这个权值实质是各检测器决策的虚警和漏报概率的函数。在数据融合的入侵检测系统中，融合后的决策结果是各个检测器的决策结果和每个检测器的可靠性决定。通过一个标准、统一的功能测试，对各个检测器的可靠性进行评估，根据这个评估结果生成融合的决策公式。

3　分析比较结果

为了避开入侵检测系统的检测，出现了许多IDS逃避技术，主要包括字符串匹配弱点攻击、多变shell代码、会话拼接、碎片攻击等，下面根据碎片重组攻击来比较DFIDS与普通IDS的功能。

碎片重组的问题是在进行字符串匹配以前，入侵检测系统必须在内存中缓存所有的碎片，然后进行重组，还需知道碎片在目的主机如何重组。利用重组的缺陷可以形成攻击，其中包括碎片覆盖、碎片重写、碎片超时和针对网络拓扑的碎片技术等：

1) 碎片覆盖，即发送碎片覆盖先前碎片中的数据，如碎片1“GET x.idd”和碎片2“a?”(缓冲区溢出数据)，第二个碎片的第一个字符覆盖第一个碎片最后一个字符，两个碎片被重组之后就变成了GET x.id**a?**(缓冲区溢出数据)。

2) 碎片数据覆盖，即覆盖全部的碎片数据，如碎片1“GET x.id”和碎片2一些随机的字符，以及碎片“3 a?”(缓冲区溢出数据)。这些碎片在经过目标系统的重组之后，碎片3将完全覆盖碎片2，重组之后的数据变成GET x.ida?(缓冲区溢出数据)，如果入侵检测系统的重组方式和目标系统不同，就无法重组出“GET x.ida?(缓冲区溢出数据)”，因此检测不出这个攻击。

3) 碎片超时，这种攻击依赖于入侵检测系统在丢弃碎片之前会保存的时间，从收到第一个碎片开始计时，大多数系统会在60 s之后将丢弃不完整的碎片流。如果入侵检测系统保存碎片的时间小于60 s，就会漏掉某些攻击，如碎片1(设置了MF位)“GET foo.id”、碎片2(59 s之后发出)“a?”(缓冲区溢出数据)。如果IDS保存起始碎片的时间不到60 s，就会漏过攻击。如果配置没有错误，现在的网络入侵检测系统能够检测此类攻击，这种技术结合其他的网络技术将更有威胁。如果入侵检测系统和被监视的系统不在同一个网段，攻击者可以在TTL上想办法。在某些应用中，由于经费的限制，不能在自己的每个子网都部署IDS节点，只在网络的出入口部署一套IDS，监视所有的网络流量。这种情况下，如果被攻击的主机在其他的子网，攻击数据包到目标系统的跳数就大于到IDS的跳数。攻击者可以伪造碎片的TTL，使某些碎片刚好能够到达，而无法到达目标系统，如碎片序号负载TTL(假设攻击者到目标的跳数是5，到IDS的跳数是3)，1 GET

foo.id 5，2 evasion.html 3，3 a?(缓冲区溢出数据)。从这些碎片中，IDS重组的数据是“ GET foo.idevasion.html a?(缓冲区溢出数据) ”或者“ GET foo.idevasion.html ”(IDS的超时时间小于60 s)。通过这种方式，攻击者成功地在IDS中插入了垃圾数据。

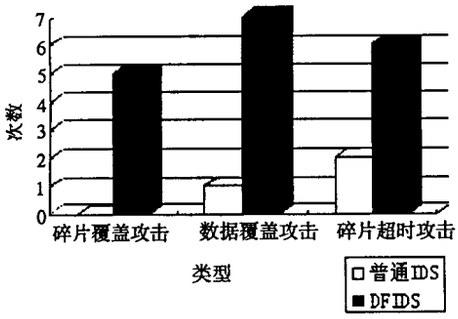

图 2 碎片攻击性能比较

采用碎片攻击的方式来测试DFIDS和普通IDS，在某个网段内安装DFIDS和snort作为对比，DFIDS与snort检测相同的100 M网络段。DFIDS和snort分别安装在PIII800、256 M内存和100 M网卡的相同配置计算机上。在网络段内放置测试的Linux服务器，用网络外的攻击测试机发起攻击。分别进行10次碎片覆盖攻击、碎片数据覆盖攻击、碎片超时攻击。测试数据如图4所示，纵坐标表示检测出的攻击次数，横坐标为攻击类型。从图中可以看出，由于采用了数据融合技术，入侵检测的能力大大提高了。这是因为FDIDS除了搜集普通IDS搜集的网络数据包以外，还通过系统日志、系统负荷的传感器搜集信息，这些信息与网络数据包的信息融合，得到了更强的入侵检测能力。

4 结 束 语

本文将数据融合技术运用到入侵检测系统中，提出了一个使用数据融合技术的入侵检测系统的模型DFIDS。在DFIDS系统中，有多个检测器搜集系统日志文件、网络流量信息、网络数据包等数据，通过决策器进行本地决策，然后传送到融合中心。融合中心决策融合的数学模型，根据这个模型建立了入侵检测系统的数据融合算法。利用碎片重组攻击和字符串匹配弱点攻击，测试了DFIDS系统的入侵检测能力。实践证明，数据融合的入侵检测系统具有更强的入侵检测能力，是入侵检测技术发展的一个重要方向。

本文研究工作同时也得到了校青年基金(YF021501)资助，在此表示感谢。

参 考 文 献

[1] Tenney R R, Sandell N S R. Detection with distributed Sensors[J]. IEEE Transaction. AES., 1981, 17 (4): 501-509

[2] Chair Z, Varshney P K. Optimal data fusion multiple sensor detection system[J]. IEEE Transaction. AES. 1986, 22 (1): 99-101

[3] Baek W, Bommareddy S. Optimal m-ary data fusion with distributed sensors[J]. IEEE Transaction. AES. 1986, 31(1): 1150-1152

[4] Kam M, Zhu Q, Gray W W. Optimal data fusion of correlated local decistions in multiple sensor detection systems[J]. IEEE Transaction. AES, 1988, 18(5): 916-920

[5] Thomopoulos S C A , Viswanathan R, Bougoulias D C. Optimal decision fusion in multiple sensor detection systems[J]. IEEE Transaction. AES,1992, 28(3): 644-653

[6] Chair Z, Varshney P K. Distributed bayesian hypothesis testing with distributed data Fusion[J]. IEEE Transaction. SMC.,1988, 18(5): 695-699

编 辑 徐培红