数据融合技术在网络安全中的应用

王祖俪, 甘　刚
(成都信息工程学院网络工程系 四川 成都　610225)

【摘　要】: 本文简述了在网络安全管理中引入数据融合技术的原因,介绍了数据融合技术的特点和相关算法。随后根据两个典型的系统框架分析了数据融合技术在网络中的不同应用。最后指出了数据融合技术在网络安全中应用的发展方向和难点。

【关键词】: 数据融合; 多信息源; 网络安全;

0. 引言

数据融合技术，特别是多传感器的数据融合技术是利用多个传感器在时间和空间上的冗余或互补信息依据某种准则进行组合,以获取被观测对象的一致性解释或描述。采用数据融合技术可以增强系统的可靠性,扩展了单个传感器的性能,减少了信息的模糊性。 网络环境中现有的各种安全设备以及对网络流量的监控构成了良好的多信息源， 在信息安全管理中引入数据融合的理念和实践经验,可以更好地维护信息系统的安全。

本文的第二个部分介绍数据融合的概念以及引入到网络安全中的原因,对数据融合中采用的算法进行了说明和比较。在第三部分中结合网络安全的实际背景， 以具体的系统分析了数据融合技术在单一设备和多安全设备中的应用情况。文章最后部分对数据融合在网络安全中的运用需要解决的难点和未来发展方向进行了阐述。

1. 数据融合的特点

1.1 数据融合的概念

数据融合技术最早出现在军事领域， 是将来自多个传感器或多源的信息在一定准则下加以自动分析、综合以完成所需的决策和估计任务而进行的信息处理过程。1973 年美国国防部资助开发了声纳信号理解系统, 数据融合技术在该系统中得到了最早的体现.目前,数据融合在军事方面主要运用在自动目标识别、战场监视、自动飞行器导航、自动威胁识别、遥感等方面; 在非军事领域的应用包括生产过程监控、机器人、医疗诊断、复杂工业过程控制等方面。

数据融合特别是多传感器数据融合技术, 以多个传感器在时间和空间上的冗余或互补信息依据某种准则进行组合, 以获取被观测对象的一致性解释或描述。传感器之间的冗余数据增强了系统的可靠性,互补数据扩展了单个传感器的性能。数据融合技术扩展了时空覆盖范围,改善了系统的可靠性,对目标或事件的确认增加了可信度,减少了信息的模糊性。图 1 为通用的数据融合的过程:

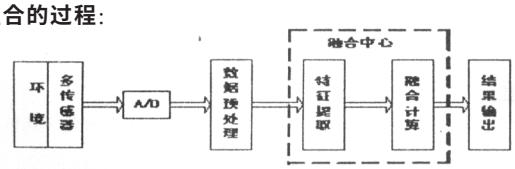

图1　数据融合过程

1.2 数据融合引入到网络安全中的原因

1) 目前的网络攻击越来越多样化和多步骤， 众多攻击手段让一些单一功能的网络安全产品措手无策、无能为力,比如基于病毒码的防病毒软件无法及时识别新的蠕虫攻击， 而孤立地对网络安全设备产生的事件进行分析和处理将无法对整个系统的总体安全状况和态势进行判断,这给网络带来了极大的隐患。

2) 目前的网络的防御手段也随之增多,既包括防病毒软件,也包括防火墙、入侵检测系统等,这些产品产生大量的不同形式的安全信息， 使得整个系统的相互协作和统一管理成为安全管理的难点。

3) 网络攻击手段的融合推动了对抗攻击手段技术的融合。从信息安全管理的角度来说， 网络环境中现有的各种安全设备以及网络流量监控等,构成了良好的多信息源,通过对网络安全事件的融合分析， 可以有效地提高网络安全事件的综合分析处理效率,增强网络安全管理环境中安全事件分析的有效性,并形成对系统环境整体的安全性态势的评价。

1.3 数据融合的常用算法:

目前国内外在对于信息进行融合时通常采用以下算法法进行分析:

1) 综合平均法: 利用加权或者其他考虑因素对收集到的信息采用求平均值的方法,过滤掉错误信息,整合已有的数据信息得到最终的结果。

2) 贝叶斯估计法 : 它通过贝叶斯网络,将先验信息和样本信息合成为后验信息,对检测目标的结果作出推断。但这种算法需要先验知识,且当有多个可解的假设和多个条件相关时,算法显得很复杂,不适合实时的判断。

3) D- S 法(Dempster- shafter) : 它是贝叶斯算法的扩展,利用概率区间和不稳定区间来确定多证据下假设的似然函数,但它不能有效地处理矛盾的信息,且算法复杂性较高。

4) 聚类分析: 根据预先指定的相似标准， 把数据分为若干类。

5) 神经网络方法: 由于具备大规模并行处理能力和固有的容错性以及自学习、自适应功能,目前神经网络正成为数据融合技术的研究热点。

2. 数据融合在网络安全中的具体应用

2.1 数据融合的三个层次:

数据融合技术近几年才逐渐应用到网络安全中， 在网络安全中主要有两大类型的运用,一是运用在单个安全产品上,使得单个安全产品的报警率更准确。二是运用在不同安全产品上,控制整个信息系统的安全状况。数据融合的层次可以分为象素级融合,特征级融合和决策级融合 3 种。

1) 象素级融合是指直接在采集的原始数据基础上进行融合,这是最低层次的融合,这种融合的优点上,可以保持尽可能多的数据,提供其他层次不能提供的细微信息,但局限性也很明显,处理数据过大,代价高,时间长。

2) 特征级融合: 在融合前先对所采集到的原始信息进行特征提取,然后对特征信息进行融合分析和处理,特征级的融合优点是实现了信息的压缩， 有利于实时处理。网络入侵检测系统(NIDS)中所采用的数据融合就是属于特征级的信息融合。

3) 决策级融合:决策级融合是一种高层次的融合,其结果为指挥控制决策提供依据， 决策级融合必须利用特征级融合所提供的多种信息,再次进行融合,得到产生决策的依据。它的优点是具有容错性,抗干扰性好,信息的来源可以来自同一设备,也可以来自不同的设备， 但在融合之前需要对不同格式的信息进行格式的统一转换。

目前在网络安全应用中已使用的数据融合技术大都是基于象素级和特征级的信息融合， 如智能 IDS 的研究， 防毒墙的出现,但在集中式网络安全管理中更强调决策级的信息融合,决策

级所进行融合的数据来自于初级层次上由各种安全设备（防火墙，入侵检测系统等）经过特征级融合后产生的信息。图 2 是在网络安全中通用的数据融合系统模型。

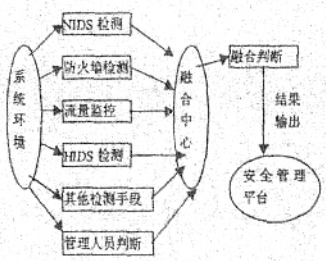

图 2 数据融合通用模型

2.2 在单一设备中运用数据融合的示例：

文献 4 提出了一种基于操作的多层次主机入侵检测的模型，它的工作模型如图 3：

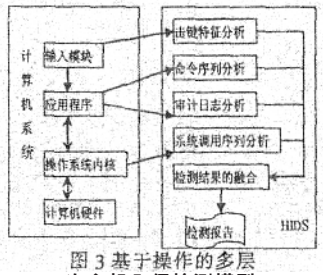

图 3 基于操作的多层次主机入侵检测模型

在该模型中，系统通过对输入设备上进行的输入操作、对某个应用程序下达的操作指令、对系统日志的分析和对系统调用序列的分析这四个层次上监控用户对被保护主机上进行的操作，将所得到的结果进行融合得到一个总的判断。

检测模型中基本决策空间={用户行为正常，用户行为异常}，该检测模型使用了 D-S 公式逐步完成 $m_{keyboard}$、m_{cmd}、m_{audit}、m_{sys_call} 4 个 mass 函数的合成，，最终得到系统对用户行为是否正常的综合判断。比如：若 $m_{keyboard}$（用户行为正常）=65%，m_{cmd}（用户行为异常）=50%，m_{audit}（用户行为异常）=50%，m_{sys_call}（用户行为异常）=30%，根据 D-S 公式所合成的最后概率指定函数：用户正常 25%，用户异常 62%，不能确定 13%得出结论，目前用户异常的可能性最大。可见使用 D-S 合成公式对多个检测结果进行融合可以提高检测系统的检出率，降低误报率，但不足的是使用 D-S 合成公式只能得到一个用户行为异常的判断，且在时效性上不高。

2.3 在多网络安全技术中运用数据融合的示例：

多网络安全技术中运用数据融合着重于安全事件的融合，与单一的安全设备融合数据的层次、目的、效果都不同。比如：入侵检测系统是对检测到的信息的特征级的融合，而入侵检测系统的输出是网络安全事件融合系统的输入。文献 5 中根据 NSAEM（Network-based Security Analysis and Evaluation Model）提出了一个基于多网络安全技术融合的网络安全评估系统。它的功能框架图如图 4

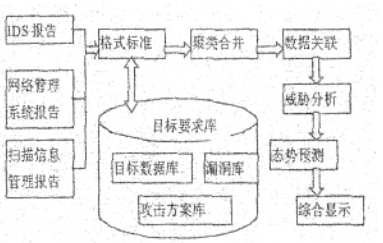

图 4 多网络安全技术融合模型

在该模型中整个评估系统的输入信息为各个安全技术产生的多源信息，这些信息的格式可能是不同的，所以在融合这些信息之前需要对格式进行统一的标准化转换。在格式转换后就要对收集到的多源信息进行确性度的融合。在该系统中融合算法主要采取证据理论的方法，通过聚类合并的方式将各种信息进行归类处理，最后将判断结果存储在数据库中。同时该系统还运用贝叶斯网络和案例推理相结合对整个网络的安全态势进行分析。多网络安全技术的融合有利于发挥各种安全技术的优势，更全面的掌握整个信息管理系统中的安全情况。

3. 结束语

数据融合在其他领域，如军事、海洋监视系统，医疗诊断方面的应用已日趋成熟，近年来数据融合的思想才运用在网络安全管理中，因此还有很多需要急待完善的工作：

1) 在网络安全中所运用的数据融合的算法都是在其他应用领域中已经成熟和有效的融合方法，但有的算法并不完全适应网络安全应用的背景，因此建立符合网络安全特质的融合算法，是目前网络安全信息融合研究的一大热点。

2) 网络安全中的数据融合除了不同安全技术的融合还应该包括安全技术与网络设备的融合，如与路由器等设备的融合。安全技术的融合通过创建新的安全理论，整合各种安全解决方案，构建综合的动态网络来最大化安全防护效果，在技术层面上，整合安全又要通过对诸多安全产品和网络设备的全面整合，为信息网络提供全面动态的安全防护体系。

总而言之，网络安全融合技术要通过创新安全理论、整合各种安全解决方案，整合网络安全资源，构建综合的动态网络来实现安全防护的效果。在理论上通过各种安全理论的整合，构建全新的网络安全理论；在技术上，整合网络安全融合技术和网络安全产品，构建全面动态的安全防护体系，从而满足不同领域网络安全多系统集成问题的需要。

参考文献：

1. 多传感器数据融合技术研究与进展.[J/OL] http://www.chinasensor.cn/news/2006-1/200611091001.htm
2. Zenghuanglin.Theory of rough and neural networks [a].Neual Network Theory and Its Appliacation[M].Chongqing:Publishing Company of Southwest Traffic University,1996.
3. 张文修,梁怡. 《不确定性推理》[M],西安交通大学出版社 2000 年
4. 蔡中闽,彭勤科等.基于操作的多层次主机入侵检测模型与方法[J],计算机工程,2002.7.
5. 刘超,谢宝陵等.基于数据融合模型的网络安全分析评估系统[J],计算机工程,2005.7。
6. 黄声列,陈思国等,网络安全融合技术[J], 《现代情报》,2005.8

（上接第 101 页）

数据以丰富多彩的形式表现出来，不仅增加了数据的直观性，帮助人们更快更准确地做出判断，而且这种丰富的表现形式更有利于 GIS 的进一步普及。

4. 结语

随着 GIS 的应用越来越广泛，GIS 的二次开发也越来越深入，基于 COM 技术的 GIS 二次开发，充分利用 COM 技术的高效性与灵活性，有效降低了 GIS 二次开发的难度，缩短开发周期，是今后相当一段时间内 GIS 二次开发的主流。ATL 作为最新和最为高效的 COM 开发技术，无疑将成为 GIS 二次开发的最好选择。

参考文献：

1. 卢振千，黄杏元. 基于 COM 与 ARC/INFO 8 的系统开发及应用研究[J]. 科技通报,2002,18(1):31-37
2. ESRI. ArcObjects Developer's Guide[M]. RedLands:ESRI,1999
3. 韩鹏，徐占华，褚海峰 等. 地理信息系统开发-ArcObjects 方法[M]. 武汉：武汉大学出版社，2005.9
4. 毛玉龙. ArcGIS 的二次开发[J]. 福建电脑，2006,(2):84-85
5. 贾艳鹏，胡社荣，李凤久，贾青梅. 基于组件的地理信息系统开发技术[J]. 矿山测量，2003,(1):44-45,54
6. 陈方明. 基于 ATL 的 GIS 综合工作平台[D]. 杭州：浙江大学计算机学院，2004