ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 37,No. 3,May 2022,pp. 488－501
DOI:10. 16337/j. 1004-9037. 2022. 03. 002

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

内部威胁发现检测方法研究综述

郭世泽[1]，张　磊[1]，潘　雨[1]，陶　蔚[2]，白　玮[1]，郑奇斌[3]，刘　艺[4]，潘志松[1]

(1. 陆军工程大学指挥控制工程学院，南京 210007;2. 军事科学院战略评估咨询中心，北京 100091;3. 北京大数据先进技术研究院，北京 100091;4. 军事科学院国防科技创新研究院，北京 100010)

摘　要：组织内部网络不仅面临着外部攻击者的威胁，同时也面临以破坏组织网络结构、内部信息资料窃取以及各种诈骗手段为主的内部威胁。内部威胁因为其多元化、伪装性强等特点，对组织机构内部造成了严重影响，因此对于内部威胁发现检测方法的研究变得非常有必要。本文首先对内部威胁进行了描述，重点针对内部威胁发现检测方法的现实意义进行了论述。同时将现有的内部威胁发现检测方法分为3类：基于异常行为的检测方法、基于审计日志异常的检测方法和其他检测方法，分别介绍了现有3类方法的研究现状，并对它们的研究进展进行了总结、归纳和分析。最后对内部威胁发现检测方法的未来研究方向进行了展望。

关键词：内部威胁；发现检测；异常行为；审计日志异常

中图分类号：TP181　　**文献标志码**：A

Survey on Insider Threat Detection Method

GUO Shize[1], ZHANG Lei[1], PAN Yu[1], TAO Wei[2], BAI Wei[1], ZHENG Qibin[3], LIU Yi[4], PAN Zhisong[1]

(1. Command and Control Engineering College, Army Engineering University of PLA, Nanjing 210007, China; 2. Evaluation Center, Academy of Military Science, Beijing 100091, China;3. Advanced Institute of Big Data, Beijing 100091, China;4. Defense Innovation Institute, Academy of Military Science, Beijing 100010, China)

Abstract: The internal network of the organization is not only faced with the threat of external attackers, but also faced with the insider threat including destruction of the organization network structure, internal information theft and various means of fraud. Because of the characteristics of concealment, destructiveness and diversification of attack methods, the insider threat poses a serious threat to the internal network. Therefore, it is very necessary to study the detection methods of insider threat. This paper analyzes the characteristics of insider threat and expounds the significance of studying the detection methods of insider threat. The existing insider threat detection methods are divided into three categories, namely, detection methods based on abnormal behavior, detection methods based on abnormal audit diary, and other detection methods. The current research status of each aspect is introduced respectively, and the progress of the research status of each aspect is summarized and analyzed. At last, the future research direction of insider threat detection methods is prospected.

Key words: insider threat; detection; abnormal behavior; abnormal audit diary

---

基金项目：国家自然科学基金面上项目(62076251)；国家重点研发计划(2017YFB0802801)。
收稿日期：2021-08-12；**修订日期**：2022-04-24

引 言

随着社会信息化程度越来越高,企事业单位等组织的内部网络不仅面临着外部攻击者攻击的风险,而且面临着来自内部攻击者的安全威胁。一般来说,内部攻击者一般是指组织机构内部人员,包括在职或离职员工、承包商以及商业合作伙伴等[1-2]。与外部攻击者不同,内部攻击者可以利用正常业务流程获得相关内部网络的信息,或者利用社交关系获取网络管理员权限,从而对内部网络安全造成负面影响[3],这些潜在的威胁统称为"内部威胁"。内部威胁从幕后到台前的关键事件是斯诺登的"棱镜门"事件,该事件表明内部威胁已经发展成为网络安全领域所面临的重要问题[4]。文献[5]的调研报告显示,只要价钱合理,20% 的人愿意将自己的工作账号和密码卖给无关人员;美国计算机安全协会有报告指出,同样的攻击成本,内部威胁所造成的损失要远远高于外部攻击[6];普华永道的报告显示,中国公司的网络安全事件中,由内部人员造成的网络安全事件占50% 以上[7]。同时各大机构更频频爆出内部网络安全的事件,例如:巴林银行职工安全事件导致14亿美元的损失,最终导致银行破产;法国兴业银行内部网络安全事件造成的72亿美元的损失[8]。这些事件都证明了内部威胁会对公司发展产生重大影响,因而发现检测内部威胁问题变得刻不容缓。

内部威胁问题也引起了学术领域的广泛关注。中国计算机学会推荐A类会议CCS(Conference on Computer and Communications Security)针对内部威胁的发现检测技术召开研讨会进行讨论。卡耐基梅隆大学的CERT(Computer Emergency Response Teams)研究中心早在2000年就开始了对内部威胁检测方法的研究,取得了一系列的成果。根据CERT和文献[9-14]对内部威胁的定义,内部威胁和外部攻击的主要区别在于其攻击者主要来自内部,因此内部威胁一般具有以下特征:

(1) 比外部威胁损失更大:外部威胁由于是黑盒攻击,不一定会产生严重影响。但内部威胁由于攻击者是组织内部人员,相当于整个网络的配置都暴露在攻击者面前,比如网络的配置弱点、核心服务和核心资产等,因此内部威胁造成的危害相比于外部威胁更大,会对组织的经济资产、业务运行及组织信誉造成更大的破坏。在2014年CERT发布的网络安全调查显示,不到30% 的内部威胁可以造成将近50% 的经济资产损失[15],说明了内部威胁较普通的外部攻击危害更大。

(2) 具有极强的伪装性:由于攻击者来自内部,熟知内部网络的安全防护配置等情况,可以有效逃避已有的网络安全检测方法,所以内部威胁具有极强的隐蔽性。例如:内部人员发动内部攻击一般不会经过防火墙、安全防护系统等设备,导致网络安全设备无法检测到内部攻击行为,因此安全设备对内部威胁防御能力较低;内部攻击可以发生在工作时间,导致攻击行为和正常工作行为较难区分,增大了攻击数据挖掘和攻击行为分析的难度;内部攻击者具有组织安全防御机制的相关知识,并且与网络管理员和工作人员具有相应的社交关系,发现难度较大,具有极强的伪装性。

(3) 攻击者和攻击方法的多样化:在大数据时代,组织内部核心资产与业务的信息化导致内部攻击难度降低,攻击元素日趋多样化。首先是攻击者的多元化,网络管理员、组织内职工、第三方职工、合作方和服务甲方等,都有可能成为内部威胁的攻击者;其次是攻击方法的多样化,内部攻击者可以在内部网络中植入病毒,可以使用逻辑炸弹,可以利用自己的权限获得相关组织内部信息,也可以删除组织内部数据库或基础软件代码,还可以篡改相关信息进行诈骗等。攻击方法的多元化增加了内部威胁发现检测的复杂性,使得内部威胁检测问题面临更为严峻的挑战。

对内部威胁的检测发现研究有助于识别内部威胁的发生条件和原因,以及了解和应对组织内部面临内部威胁的状况,从而提高整个网络的安全防御能力。因此,无论对于网络安全管理员还是其他人员,内部威胁研究都具有非常高的理论研究价值和实践价值,主要体现在以下两个方面:(1) 对内部威胁进行研究可以有效防范其带来的危害。通过科学的方法分析系统中面临的内部威胁、研究内部威胁

的构成因素、量化内部威胁风险,可以有效对内部威胁进行发现和防御。(2) 内部威胁的发现检测防御方法是网络安全建设的重要保障。传统的网络安全防御比较依赖于外部网络防御,对内部威胁的关注不多,但内部威胁是网络安全建设中的重要环节,只有有效识别内部威胁,才能在防御和控制内部威胁的基础上保障网络的整体安全。

1 内部威胁发现检测方法

对内部威胁的发现检测是学术界多年来的一个热点问题。内部威胁发现检测方法主要通过发现内部用户的行为痕迹[16],建立相应的内部用户行为模型以检测其是否为内部威胁。因此可以将当前内部威胁发现检测方法分为基于异常行为的发现检测方法与基于形式化建模的发现检测方法两类。

1.1 基于异常行为的发现检测方法

基于异常行为的发现检测方法是基于大数据和人工智能技术的内部威胁发现检测方法,也是目前的主流方法[17-19]。该方法主要基于数据驱动,通过大数据分析,对用户的行为进行特征提取并进行分类,判断该用户行为是否正常,进而从中找出可能的异常行为,判断其是否为内部威胁。这种分析的对象既可以基于端设备的文件访问、函数调用等用户行为,也可以基于流量信息、服务访问信息和文件访问信息等用户网络行为[20]。在用户的异常行为检测中,提取用户行为数据的传统方法主要包括朴素贝叶斯、主成分分析和隐马尔可夫模型等。近年来,随着深度学习和强化学习技术的不断发展,深度学习和强化学习的联合检测方法也逐渐受到更多的关注。其中文献[21]提出一种通用的时序图分析框架来发现内部用户的异常行为,它将内部用户与系统的交互过程抽象为一系列的图,并通过图分类的不同来发现内部用户行为的异常,以此来判断是否存在内部威胁。文献[22]通过隐马尔可夫模型和单类支持向量机的联合模型来发现内部用户的异常行为和攻击行为。文献[23]提出了基于混合高斯模型的内部威胁发现检测方法,并在其中融合了相关专家的判断信息,进一步提高了发现准确率。文献[24]通过建立员工情绪档案,采用深度学习方法,提出了一种发现内部威胁早期潜在风险的方法,即通过员工情绪的异常变化来发现内部用户可能会发生的异常行为和攻击行为。文献[25]使用深度递归神经网络模型构建在线无监督学习方法,用于实时监测和检测内部系统日志中的异常活动,可随时发现内部用户的异常行为和攻击行为。文献[26]提出一种基于行为序列的预测方法,通过该方法可以为用户的相关行为进行画像,从而判断内部用户是否存在内部威胁。文献[27]提出了基于多模型、多数据融合的内部威胁发现检测方案,从而发现内部用户的异常行为和攻击行为。根据模型中使用数据源的不同,异常行为发现方法分为基于审计日志异常的发现检测方法和基于用户命令异常的发现检测方法。

1.1.1 基于审计日志异常的内部威胁发现检测方法

基于审计日志的异常行为发现检测方法主要针对用户操作留下的审计日志进行审查,从而发现其中的异常行为和攻击行为,如登录网站、访问网络以及邮件收发等记录,该方法可以对用户的行为进行较为全面的刻画。文献[28]提出一个不同类型数据融合的典型方法。他们从用户的工作组属性出发定义了不同类型数据之间的一致性,之后检验用户不同类型数据的一致性,并使用词频-逆文档模型融合用户对不同类型数据上一致性的评分。但该方法依赖于一个假设,即用户组属性必须一致,但一般实际情况与该假设恰恰相反。文献[29]针对现有内部威胁发现检测模型,将内部威胁的各个审计日志要素均纳入其中进行考虑。但是该模型概念性太强,没有实验依据,不具有指导意义。文献[30]提出了一个发现检测系统,该系统主要从3个层次来选择特征集,提出了基于场景分析的内部威胁检测系统,但是该系统存在较多漏洞,仅针对内部信息窃取进行了实验,其他内部威胁的发现方法并没有进行实验验证。文献[31]针对不同用户的角色行为特征进行提取分析,提出一种基于角色异常行为挖掘的

发现检测方法。根据序列模式挖掘原理挖掘角色正常行为,使用Knuth Morris Pratt字符串匹配算法进行发现,判断角色行为是否存在异常。结果表明,该方法可有效对用户的异常行为进行检测,同时减少了挖掘时间,并且在异常行为检测精确度上有所提高。

现有的基于审计日志异常的内部威胁发现检测方法的难点在于不同类型审计日志数据之间的结合方式。如果只是使用单一类型的审计日志数据,则对用户的网络行为反映不够全面,不具有代表性。而如果只对不同数据进行简单拼接,会造成部分特征失效、模型训练复杂度过高以及模型过拟合等问题,因此相关研究成果不多,现有的研究成果较多集中于基于用户命令异常的内部威胁发现检测方法。

1.1.2 基于用户命令异常的内部威胁发现检测方法

文献[32]是基于用户命令异常检测方法的较早研究成果,它将用户的命令序列作为分析对象,分别计算相邻命令模式出现的概率,新命令与历史命令的匹配程度来判断是否属于异常。在此之后,机器学习算法开始广泛应用起来[33],如利用朴素贝叶斯方法、期望最大化算法和支持向量机等。其中文献[34]提出基于隐马尔可夫模型的内部威胁检测方法,利用本地应用程序接口建立用户的正常行为轮廓库,当有内部用户不符合该库的行为即可认为存在内部威胁。文献[35]将朴素贝叶斯方法运用到内部威胁发现检测中,基于网络分层的方法,提出以同时出现的非邻接命令来补充用户命令模型。文献[36]提出了评价检测系统,并且指出应当在不同会话层区分内部攻击者与普通用户的能力。文献[37]提出了一种监控内部用户系统命令,从用户的文件和用户的进程中分析出文件访问与进程调用的联系,但该方法检出率不能达到100%。文献[38]主要通过分析系统窗口主题信息,即打开一个窗口自动记录主题信息以及窗口进程等信息,从而刻画出用户窗口行为特征,从这些窗口行为特征来判断内部用户是否为攻击用户。文献[39]基于系统用户的统计数据建立异常行为检测方法,提取出约1 500个系统属性特征,从提取的属性来分析用户的攻击行为,从而能够准确地刻画出用户行为。基于用户命令异常的内部威胁发现检测方法主要从用户命令序列或系统命令调用序列作为数据集,运用机器学习建立分类器,但是由于数据源过于简单以及分类器过于简单,导致多数方法检测成功率并不高,检测效果并不是太好。总体来说,基于异常行为的内部威胁发现检测方法十分依赖于数据获取的准确性和全面性,需要大量现有的内部用户行为数据,因此常常受制于实际数据采集的困难,在小样本或无样本数据的情况下,通过对目标系统进行形式化建模的方法发现检测内部威胁同样受到业界的关注。

1.2 基于形式化建模的内部威胁发现检测方法

基于形式化建模的内部威胁发现检测方法是指通过建立用户的正常行为模型,通过对比用户的现有行为,检测出偏移该模型的异常行为[40]。基于该思想,学术界很早就提出了攻击图、攻击树、Petri网和信息获取图等多种形式化的建模工具,这些工具能够有效地对内部网络信息进行建模,但是在表达能力、适用范围等方面存在不足[41-43]。近年来,学术界不断扩展形式化工具的表达能力。文献[44]考虑了物理信息系统中物理域和信息域的交互关系,以构建相应的内部威胁发现模型。文献[45]将组织架构的形式化建模分析与社会学解释相结合,并基于高阶逻辑构建了内部威胁发现检测分析框架。文献[46]建立了基于本体的组织物理安全体系弱点发现方法,用于发现潜在的内部威胁。文献[47]通过最小化网络空间安全风险,提出了基于博弈论的内部网络安全防护模型。文献[48]将物理空间信息与社会网络信息相关联进行建模来发现可能的内部威胁,实现了物理域信息和社会域信息的结合。文献[49]构建了组织内部的动机和机会模型,通过多域联合的方式来发现内部威胁。文献[50-51]则构建了层次化的本体集,用于描述内部威胁的社会因素和组织因素,从顶层设计层面给出了组织行为和个人因素对内部威胁的影响。

为提高模型的泛化性以及对用户行为正常行为变化的适应性,文献[52]提出了一种解决方法,在模型中同时部署了K个分类器,采用"K-投票"的形式对用户行为进行判断,同时K个分类器实时更新。

文献[53]提出了基于文件使用的内部威胁检测系统,该系统用于检测攻击者的攻击,从用户查看文件系统以及访问文件的角度建立用户的行为模型,一旦有偏离该模型的异常行为即报警。文献[54]将文件目录作为用户任务的抽象,对内部用户的行为进行画像,通过与朴素贝叶斯方法与马尔可夫模型对比实验证明了其内部威胁发现检测系统要优于其他两种模型。文献[55]针对用户遍历文件系统时的文件时间序列关系,建立了文件目录图与用户访问图的行为模型,并且使用朴素贝叶斯分类器检测文件访问行为的突然变化。除了上述检测方法外,文献[56]基于一个枚举攻击方式全集的假设,即用户在做任何行为前需要说明其使用意图,任何偏离其使用意图的行为都被看作可能会发生内部威胁。但现实中不可能枚举所有的攻击方式,因此该方法可操作性不强,并且需要存储用户意图集,增加了计算难度,降低了检测效率。文献[57]针对内部威胁,提出了基于文件内容的异常检测模型,该模型使用文本分割与朴素贝叶斯方法对文件内容分类,然后根据内部用户行为以及内部用户组群间的行为偏移来检测文件访问中的异常行为,实验证明了该模型在保护系统内部文件访问上具有一定的有效性,但是效果完全取决于所用数据库的丰富程度,也需要大量数据集的支持。

基于形式化建模检测方法的另一个重要工作是基于图方法的内部威胁发现检测方法[58-59]。文献[60]在攻击树的基础上提出了关键挑战图,顶点代表着主机或服务器,边代表着实体间的通信,每个顶点上标注了相关的资源信息,如用户名密码等[61]。用户访问行程可建模为一个关键挑战序列,一旦符合关键挑战序列即可能存在内部威胁,并且可以计算图中单独分支的内部攻击成本。文献[62]提出了基于图的检测发现算法,该算法的核心是刻画图的输入输出等变化状态。具体可以分4步:(1)获取相关的数据集,检测是否存在相关异常;(2)基于异常与相关内部用户创建图;(3)建立相关数据移动图;(4)学习用户频繁子图模式,以该模式作为正常模式,其余的即为异常模式,依此来发现内部威胁。但该方法检测效率不高,图计算在1 000多个顶点就需耗费72 h时间,实用性不强。文献[63]通过设置“设备-操作-属性”三元组对用户及对应的角色行为进行树结构抽象,目的是为了更加全面地刻画用户行为;同时还设计了一个3层内部威胁评估系统,每层检测到异常都会分类为异常行为,并且通过反馈实时更新模型参数。文献[64]综合了攻击树与行为树,提出了活动树模型,它通过记录用户的工作模式,从分支长度对应节点相似性方面判断新行为与已有工作模式的相似性,相似性高即为正常内部用户,否则可能存在内部威胁。文献[65]基于内部用户构建了贝叶斯网络攻击图,通过计算不同攻击路径的概率从而检测该路径下内部用户行为的异常程度。文献[66]在攻击图中加入了用户意图信息,基于意图信息构建用户的合法单元操作集合,然后生成用户最小攻击树,通过该模型实时监控用户行为在最小攻击树中的进度判断用户是否存在内部威胁。文献[67]提出了一种可以改变组织内部角色集合的防御思路,即角色动态调整算法。该算法首先定义了带有参数的目标函数,通过启发式搜索策略和子集结对操作得到一组候选角色,使用启发式函数计算角色分值,按照角色分值的高低对候选角色集进行选择,得到一个调整角色集;进一步,以降低角色冗余度为目标,使用调整角色集继续为用户重新分配角色,从而得到新的系统角色配置。该模型主要目标是调节公司内部系统的角色集,从而降低内部威胁的发生概率。

高级持续威胁(Advanced persistent threat, APT)作为内部威胁的一种恶意、故意、有明确目标以及安全威胁持续时间长等特点,近年来对组织、企业和政府内部网络的攻击次数和危险程度都在不断增长,而基于形成化建模的发现检测方法可以有效发现此类新型内部威胁。文献[68]综述了现有APT攻击防御框架研究的现状,并分析网络流量异常检测、恶意代码异常检测等基于网络安全大数据分析的APT攻击检测技术的研究内容与最新进展,并指出现有技术所面临的挑战和下一步发展方向。文献[69]对APT攻击检测进行了研究,完善了APT攻击检测过程,同时考虑用户行为特征对APT攻击检测的影响,增强网络的安全性。文献[70]为解决攻击数据样本少、攻击持续时间长、准确率低的问题,

提出了基于生成式对抗网络(Generative adversarial networks, GAN)和长短期记忆网络的APT攻击检测方法。该方法基于GAN模拟生成攻击数据,为判别模型生成大量攻击样本,从而提升模型的鲁棒性和准确率。文献[71]提出了一种基于深度学习和网络流量的APT攻击检测方法,将网络流量分析为基于互联网协议地址的网络流,然后从网络流中重构出地址信息,最后利用深度学习模型提取特征,从其他地址中检测出APT攻击地址。文献[72]提出了一种新的APT攻击报警系统。该系统参考美国国家标准与技术研究所的网络安全框架,是一种实时检测并响应反恶意软件的主机入侵检测系统,可以监控异常改变主机配置的行为。实验证明,该系统能提供一个良好的网络环境防止APT攻击。文献[73]提出了一种频率粒子群算法对APT攻击进行分类。该分类方法共有4个阶段,在功能提取阶段使用频率粒子群算法提取最佳特征集,在预测阶段采用最佳加权特征方法。同时该方法使用K-means聚类算法和K近邻分类器,最后利用混淆矩阵评估该方法的性能并通过实验给出了结果。

1.3 其他方法

除了上述有关研究,还有其他诸多内部威胁发现检测方法[74-76]。首先是外设设备使用检测。外设设备检测是研究用户使用外设设备行为的方法,主要以鼠标与键盘为代表。文献[77]从用户使用鼠标的角度,通过用户使用浏览器时的鼠标操作,构建鼠标行为数据库。该数据库涉及鼠标使用时的光标移动坐标、移动距离等不同特征,并且定义了一个方形光标矩阵,重点记录了鼠标在移动、点击以及推拽时的坐标、速度等特征。但该方法数据集过少,代表性不强,不能反映真实的内部威胁状况。键盘使用检测有两种方法:一种是静态文本监测,即研究用户输入同一段文本的行为;另一种是动态文本监测,即收集用户随意输入文本的方式。文献[78]是一种静态文本监测方法,从用户输入口令中分析用户输入方式的变化,并从包含用户的数据集中提取不同的键盘输入特征。但该方法同样数据量过少,代表性不强,不能反映真实的情况。文献[79]分析了用户的邮件信息,主要记录击键行为与时间戳。此类方法的不足在于监控文本输入通常缺乏良好的用户交互性,但模型却较为确定。

此外,基于传统的蜜罐研究也颇有成果。文献[80-81]针对密标的使用方式提出了诸多改进,主要有:(1)针对内部网络嗅探行为检测,设计了在网络内大量散布密标的方法;(2)在邮箱中加入有密标的伪邮件,当伪邮件中密标被内部攻击者使用时就会触发报警。文献[82]提出了哈希运算消息认证码、陷阱主机以及灯塔诱饵3种密标文件设置方法,能够在内部攻击者使用密标时获取攻击者信息,具有很大的实用性。文献[83]结合心理学知识预测内部威胁,一方面从设置的诱饵主机中监测用户的行为轨迹,分析其异常程度,同时借助心理学相关知识,计算、检测每个用户的攻击行为倾向以及所处压力水平。但该方法只使用了心理学知识,并且心理计量测验过于简单,不能完整反映用户的真实心理状态。文献[84]提出了基于攻击面转移的内部威胁检测方法,具体阐释了攻击面以及攻击面转移的形式化定义,然后分析了攻击面4个层次的动态转移技术,并且针对不同的动态转移技术进行分析和比较。文献[85]设计了一个博弈理论框架,该框架是一种由1个生成器、1个激励调节器和1个信任操纵者组成的欺骗机制。该机制以动态蜜罐配置为例,研究蜜罐配置对蜜罐内部威胁的影响。实验验证了最优的欺骗机制能够诱导内部威胁人员采取相关行动,从而改善内部网络的安全态势;实验也表明防御者总是能从伪造蜜罐中获益。

传统的内部威胁检测方法严重依赖特征工程,由于底层数据的高维性、复杂性、异构性、稀疏性、缺乏标记以及内部威胁的微妙性和适应性,很难准确捕捉内部攻击用户和内部普通用户的行为差异。文献[86]综述了利用先进的深度学习技术从复杂数据中学习特征的方法,为内部威胁检测提供了一种新的思路。与传统算法相比,深度学习模型可以提高内部威胁检测的性能。文献[87]提出了一种基于用户中心机器学习的内部威胁检测的新系统。该系统结合了无监督异常检测和有监督机器学习方法,可以从无标记数据和非常少量的标记数据中学习,具有较高的检测率。文献[88]提出了一种集成深度神

经网络技术学习自适应合成技术采样方法,对内部人员进行了整合威胁检测,解决了数据不平衡问题。实验使用CERT数据集,结果表明所提出的集成模型提高了模型的可靠性。文献[89]提出并评估一个贝叶斯网络架构,它可以考虑行为方面与网络数据的串联。同时利用机器学习来理解数据的结构,根据内部威胁的理论基础输入专门制作的特征,并对上述行为特征进行分析。文献[90]提出一种基于无监督机器学习的内部威胁异常检测方法,该方法采用了自动编码器和隔离森林两种不同的方法,并探索了带有时间信息的数据的各种表示方式。

文献[91]分析比较现有的内部威胁项目,提出系统中应具有管理员、分析员、工程师和执法员4种实际的系统角色,不同的系统角色在内部威胁中发挥着不同的作用。文献[92]提出检测系统应当基于一种智能化的管理机制,通过安全数据收集与安全管理的分工合作形成树形动态的管理机制,从而对内部威胁进行有效预防。文献[93]提出了在内部威胁监测系统中使用引导算法,该引导算法用于从大量非标记数据集中自动标记出数据的类别。文献[94]主要提出了一种量化方法,该方法根据内部威胁目标的资产重要性权重与受威胁程度,使用相关函数计算威胁指数,从而评估内部网络系统的内部威胁态势。文献[95]针对内部威胁中攻击图不确定性导致攻击图检测误报率过高的问题引入了攻击图步骤间的转移概率,通过计算每步变化的概率,进而计算整体攻击的不确定性。该方法可以有效降低误报率,但是该方法在概率攻击图的构建上完全依赖于知识库,智能化程度不高。

现阶段云计算发展迅速,因此基于云计算的内部威胁发现检测方法研究也取得了相应的成果。文献[96]从云计算领域内分析了内部威胁的不同应对方案。由于在云计算环境下,多种资源高度融合,因此内部威胁可窃取大量用户信息,将导致更多用户受到影响和威胁。针对上述问题,提出了云计算服务商与用户双方均应采取安全措施应对内部威胁,用户方面实施强健的密码策略,云计算服务提供商安装入侵检测系统和入侵防御系统,并从数据冗余备份、认证访问控制等方面加强安全监管[97]。文献[98]指出,随着物联网技术的发展,现有的内部威胁检测框架面临新的安全挑战。由于攻击面显著增大,可能会导致公司内部威胁管理方面的风险增大。根据物联网的特点和结构对考虑物联网环境的数据源进行研究,结果表明,对于物联网环境中的内部威胁问题,使用网络和应用层的数据源比使用感知层更合适。文献[99]将云模型应用在内部威胁感知中,在基于系统访问控制关系建立的分层内部威胁模型上应用云模型感知算法,对内部威胁特征同时进行定性、定量的分析,有效提高了内部威胁检测系统的准确性。文献[100]针对网络攻击持续高发的现状,构建了基于随机博弈的内部威胁态势检测发现模型。该模型将外部威胁情报与系统内部威胁事件之间的相似度进行比较,对目标系统进行威胁察觉,在此过程中利用博弈论的思想对系统当前的网络安全态势进行量化,最终实现对内部威胁的预测。

1.4 总结

综上所述,现有的内部威胁发现检测方法总体来说有以下几个共性问题:(1)如何判断异常行为是内部威胁。在基于异常行为的内部威胁发现检测方法中,检测出的异常行为有可能就是普通用户的异常行为,并不存在内部威胁,这导致内部威胁的发现检测难度增大,因此首先要区分该异常行为是内部威胁还是普通存在的异常行为。(2)对内部威胁的检测是否准确全面。现有研究对内部威胁的发现检测仅从单一方面考虑,例如仅从用户的异常行为考虑。实际上影响内部威胁的方面有很多,仅从单一方面去判断是否存在内部威胁不能全面地了解内部威胁,同时说服力也不够。(3)真正应用于实际的研究较少。前文所述的内部威胁发现检测方法,普遍存在效率不高、发现检测时间较长等问题,同时还有部分研究仅停留在理论层面,无法在组织机构内部普遍应用。

2 下一步研究方向

未来内部威胁研究需从内部异常发现、更加全面的内部威胁检测以及实际应用3个方面进行。本

文分析未来关键技术如下:

(1)更准确的发现检测技术。在内部威胁的发现检测技术中,较为突出的问题就是如何区分是正常用户的异常行为还是攻击用户的内部攻击行为。而在基于异常行为的内部威胁发现检测中,上述方法不能对这两种行为进行有效区分。因此在下一步研究中,应进一步提高内部威胁的判别率,区分哪些是正常的异常行为,哪些是真正的内部威胁。针对该问题,可以将深度学习引入到内部威胁的研究中,一方面研究内部威胁检测集成学习方法,如深度学习加多步分类器、深度学习加横向K投票分类器等;另一方面可以研究内部威胁基准模型选择方法,例如基于用户自身行为的纵向比较,或基于用户角色行为的横向比较等,从而提高内部威胁发现的准确率。未来内部威胁检测还应考虑不同公司机构内部特点,基于不同领域的业务特点与组织特点,分析攻击的薄弱处,提取威胁特征建立多域数据的关联,实现内部威胁检测的融合分析,有效提高内部威胁的检测率和降低误报率。

(2)如何应对内部威胁与外部威胁的联合攻击行为。现有的内部威胁发现检测技术大部分都是基于内部人员的攻击行为进行研究,对象仅为内部人员。随着攻击手段越来越多样化,内部人员与外部人员联合发动的内外部协同攻击也时有发生,在此情况下将更难发现内部的攻击用户,攻击产生的后果更严重,导致网络安全防御的压力也更大。而现有的研究仅针对内部威胁或外部威胁进行发现检测,对于两者联合攻击行为发现检测的研究成果较少。由于内外部联合攻击可以从内部人员的社交关系中去发现提取,因此可以对内部人员采集数据,而且数据不仅要包括用户主观要素,还应包括基社交媒体状态,以及历史档案、性格分析和心理状态等多方面个体特征数据,从而有效对内部人员的用户系统调用行为和社交行为进行刻画,建立多层监视器。同时利用社会工程学对APT攻击进行预判,有效解决攻防对抗中信息不对称的问题,保证系统效率。此外,还可以基于数据关联方式进行检测发现,通过主客观特征综合检测方式,如用户主客观行为关联图中异常子图检测,或分析主客观数据的使用顺序,如先对客观数据异常检测、然后关联可疑用户、之后通过主观数据行为进行发现检测,则可以有效降低误报率。

(3)更为高效简单实用的发现检测方法。现有的大部分内部威胁发现检测技术都存在应用难的问题,即方法实用性不强,效率较低。相反,现有的攻击技术较为成熟,可供防御者的反应时间较短。如何提高发现效率,提升防御者反应速度,做到攻击即被发现,将是未来内部威胁的研究方向之一。现有方法实时检测能力差,主要原因是特征分析技术效率较低。因此可以先从内部威胁检测系统发现威胁信息,提取特征后上传数据库;然后建立内部威胁信息与特征映射方法,比如多层次行为特征文件、哈希行为模式层次文件等,通过建立特征数据库,将威胁特征输入检测系统。此时由于有大量内部威胁特征的数据被清洗、优化以及分类,可再针对特征进行二次挖掘,从而发现内部威胁的本质特征,提高内部威胁的发现效率。

3 结束语

随着科技的不断发展,网络技术的普及使得人们的工作生活变得越来越方便,但随之而来的网络安全问题也越来越严峻。内部威胁是指公司组织内部人员发动的网络攻击,具有隐蔽性强、破坏性大的特点,能够造成比外部威胁更大的损失,因此内部威胁也越来越受到研究人员的重视。本文首先分析了内部威胁的特征,然后综述了国内外关于内部威胁的研究现状,并指出现有研究虽取得了一定的成果,但仍存在不足:首先异常行为与内部威胁方面,现有研究还不能有效区分普通用户的异常行为和攻击用户的攻击行为;其次针对内部威胁和外部威胁联合攻击的研究不多;最后现有研究应用性较弱,还不能对内部威胁进行有效而快速的发现检测和防御,确保内部网络信息的安全。通过本文综述可以得出以下结论:(1)内部威胁危害大、防御难,主要体现在内部人员掌握一定的组织知识,可对组织内最

薄弱的环节入手实施内部攻击,同时防御者重心在外部防御上,导致内部威胁不易被发现;(2)通过发现内部用户的异常行为,可对内部威胁实施有效的发现检测,但存在异常行为不一定就存在内部威胁;(3)通过形式化建模可对用户行为进行刻画,当存在偏离模型的可疑行动时,即可触发内部威胁报警,但如何进一步提高报警的精度是下一步研究的重点方向;(4)随着内外部威胁的联合攻击越来越多,如何应对内外部威胁的联合攻击将是下一步研究的主要方向;(5)内部威胁的发现检测效率问题有待进一步提高,为下一步在组织内部大规模应用打下良好基础。信息化时代的内部威胁越来越严重,尤其攻击手段和攻击人员的多元化导致内部威胁的发现检测难度不断加大,因此要及时加大互联网时代内部威胁的研究,对更多的组织普及内部威胁的危害,建立起既能防御内部威胁、又能防御外部威胁的联合防御机制,积极应对内部威胁带来的一系列挑战。

参考文献:

[1] HOMOLIAK I, TOFFALINI F, GUARNIZO J, et al. Insight into insiders and IT: A survey of insider threat taxonomies, analysis, modeling, and countermeasures [J]. ACM Computing Surveys, 2019, 52 (2): 30.

[2] SCHULZE H. 2019 Insider threat report [EB/OL]. (2019-12-19) [2021-06-15]. https://enterprise. verizon. com/resources/reports/ insider-threat-report.

[3] 张焕国, 韩文报, 来学嘉, 等. 网络空间安全综述[J]. 中国科学: 信息科学, 2016, 46(2): 125-164.
ZHANG Huanguo, HAN Wenbao, LAI Xuejia, et al. Survey on cyberspace security[J] Scientia Sinica (Informationis), 2016, 46(2): 125-164.

[4] CLAYCOMB W R, NICOLL A. Insiderthreats to cloud computing: Directions for new research challenges[C]//Proceedings of the IEEE 36th Annual Computer Software and Applications Conference (COMPSAC). Lzmir, Turkey: IEEE, 2012: 387-394.

[5] 孟子航. 信息安全与密码学[J]. 中文信息, 2018, 11: 1-3.
MENG Zihang. Information security and cryptography[J]. Chinese Information, 2018,11: 1-3.

[6] 巩琦. 计算机网络安全大事件分析及防范[J]. 信息周刊, 2019(16): 194-195.
GONG Qi. Analysis and prevention of computer network security incidents[J]. Information Weekly, 2019(16): 194-195.

[7] 普华永道. 企业内鬼:网络安全的最大威胁[J]. 首席财务官, 2016(1): 82-83.
PWC. Enterprise ghost: The biggest threat to network security[J]. Top CFO, 2016(1): 82-83.

[8] YANG Y. The revelation that insider threats cost billions of dollars[J]. Information Security, 2010, 24(6): 53.

[9] CAPPELLI D M, MOORE A P, TRZECIAK R F. The CERT guide to insider threats: How to prevent, detect, and respond to information technology crimes[M]. New Jersey, USA: Addison-Wesley Professional, 2012.

[10] SCHULZE H. Insider threat 2018 report [EB/OL].(2018-12-23) [2021-06-15]. https://www.ca.com/content/dam/ca/us/files/ebook/insider-threat- report .pdf. 2018.

[11] BERTINO E. Data protection from insider threats[J]. Synthesis Lectures on Data Management, 2012, 4(4): 1-91.

[12] GONZALEZ G G, DUBUSB S, MOTZEK A, et al. Dynamic risk management response system to handle cyber threats[J]. Future Generation Computer Systems, 2018, 83: 535-552.

[13] GEORGIADOU A, MOUZAKITIS S, ASKOUNIS D. Detecting insider threat via a cyber-security culture framework[J]. Journal of Computer Information Systems, 2021, 19: 336-346.

[14] LE D C, HEYWOOD N Z, HEYWOOD M I. Analyzing data granularity levels for Insider threat detection using machine learning[J]. IEEE Transactions on Network and Service Management, 2020,17(1): 30-44.

[15] CERT. 2014 US state of cybercrime survey[R]. Pittsburgh,PA,USA: Carnegie Mellon University, 2014.

[16] XIONG W, LAGERSTROM R. Threat modeling—A systematic literature review [J]. Computers and Security, 2019, 84: 53-69.

[17] 杨光, 马建刚, 于爱民, 等. 内部威胁检测研究[J]. 信息安全学报, 2016, 1(3): 21-36.
YANG Guang, MA Jiangang, YU Aimin, et al. Survey of insider threat detection[J]. Journal of Cyber Security, 2016, 1(3):

21-36.

[18] 王伟. 基于深度学习的网络流量分类及异常检测方法研究[D]. 合肥: 中国科学技术大学, 2018.
WANG Wei. Research on network traffic classification and anomaly detection method based on deep learning[D]. Hefei: University of Science and Technology of China, 2018.

[19] 邬江兴. 网络空间拟态防御研究[J]. 信息安全学报, 2016, 1(4): 1-10.
WU Jiangxing. Research on cyber mimic defense[J]. Journal of Cyber Security, 2016, 1(4): 1-10.

[20] AGRAFIOTIS I, EROLA A, GOLDSMITH M, et al. A tripwire grammar for insider threat detection [C] // Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats. New York: ACM Press, 2016: 105-108.

[21] MORIANO P, PENDLETION J, RICH S, et al. Insider threat event detection in user-system interactions [C]// Proceedings of the 2017 International Workshop on Managing Insider Security Threats. Dallas, Texas, USA: ACM, 2017: 1-12.

[22] RASHID T, AGRAFIOTIS I, NURSE J R. A new take on detecting insider threats: Exploring the use of hidden Markov models[C]// Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats. New York: ACM Press, 2016: 47-56.

[23] TABASH K A, HAPPA J. Insider-threat detection using gaussian mixture models and sensitivity profiles[J]. Computers & Security, 2018, 77: 838-859.

[24] SOH C, YU S, NARAYANAN A, et al. Employee profiling via aspect-based sentiment and network for insider threats detection[J]. Expert Systems with Applications, 2019, 135: 351-361.

[25] ZHANG D, ZHENG Y, WEN Y, et al. Role-based log analysis applying deep learning for insider threat detection[C]// Proceedings of the 1st Workshop on Security-Oriented Designs of Computer Architectures and Processors. New York: ACM Press, 2018: 18-20.

[26] 郭渊博, 刘春辉, 孔菁, 等. 内部威胁检测中用户行为模式画像方法研究[J]. 通信学报, 2018, 39(12): 141-150.
GUO Yuanbo, LIU Chunhui, KONG Jing, et al. Study on user behavior profiling in insider threat detection[J]. Journal of Communications, 2018, 39(12): 141-150.

[27] BROWN D P, BUEDE D, VERMILLION S D. Improving insider threat detection through multi-modelling data fusion[J]. Procedia Computer Science, 2019, 153: 100-107.

[28] HODA E, EVGENIY B, LIU J, et al. Multi-domain information fusion for insider threat detection[J]. IEEE Security & Privacy Workshops, 2013, 42(6): 45-51.

[29] NURSE J R, BUCKLEY O, LEGG P A, et al. Understanding insider threat: A framework for characterizing attacks[C]// Proceedings of the IEEE Symposium on Research for Insider Threat. Piscataway: IEEE Press, 2014: 215-228.

[30] TED E S, HENRY G G, ALEX M, et al. Detecting insider threats in a real corporate database of computer usage activity [C]// Proceedings of the ACM International Conference on Knowledge Discovery & Data Mining (SIGKDD). New York: ACM Press, 2013: 1393-1401.

[31] 顾兆军, 郭靖轩. 基于角色异常行为挖掘的内部威胁检测方法[J]. 计算机工程与设计, 2020, 41(10): 2740-2746.
GU Zhaojun, GUO Jingxuan. Internal threat detection method based on role abnormal behavior mining[J]. Computer Engineering and Design, 2020, 41(10): 2740-2746.

[32] DAVISON B D, HIRSH H. Predicting sequences of user actions[C]// Proceedings of the AAAI/ICML 1998 Workshop on Predicting the Future AI Approaches to Timeseries Analysis.[S.l.]: AAAI, 1998.

[33] PARVEEN P. Evolving Insider Threat detection using stream analytics and big data[M].[S.l.]: University of Texas at Dallas, 2013.

[34] 黄铁, 张奋. 基于隐马尔可夫模型的内部威胁检测方法[J]. 计算机工程与设计, 2010, 31(5): 965-968.
HUANG Tie, ZHANG Fen. Method of insider threat detection based on hidden Markov model[J]. Computer Engineering and Design, 2010, 31(5): 965-968.

[35] MAXION R A, TOWNSEND T N. Masquerade detection using truncated command lines[C]// Proceedings of the 2002 International Conference on Dependable Systems and Networks. Piscataway: IEEE Press, 2002: 219-228.

[36] OKA M, OYAMA Y, KATO K. Eigen co-occurrence matrix method for masquerade detection[C]// Proceedings of the Publications of the Japan Society for Software Science and Technology. Berlin: Springer, 2004.

[37] NAM N, PETER R, GEOFFREY H. Detecting insider threats by monitoring system call activity[C]// Proceedings of the 2003 IEEE Workshop on Information Assurance. Piscataway: IEEE Press, 2003: 18-20.

[38] TOM G. Authenticating users by profiling behavior[C]// Proceedings of the ICDM Workshop on Data Mining for Computer Security. Piscataway: IEEE Press, 2003.

[39] SHAVLIK J, SHAVLIK M. Selection, combination, and evaluation of effective software sensors for detecting abnormal computer usage[C]// Proceedings of the 10th International Conference of Knowledge Discovery and Data Mining (SIGKDD). New York: ACM Press, 2004: 276-285.

[40] CAMINA J B, HERNANDEZ G C, MONROY R, et al. The windows-users and intruder simulations logs dataset (WUIL): An experimental framework for masquerade detection mechanisms [J]. Expert Systems with Applications, 2014, 41(3): 919-930.

[41] KAYNAR K. A taxonomy for attack graph generation and usage in network security[J]. Journal of Information Security and Applications, 2016, 29: 27-56.

[42] 胡鹤, 胡昌振, 姚淑萍. 基于攻击图的主动响应策略选择[J]. 北京工业大学学报, 2012, 38(11): 1659-1664.
HU He, HU Changzhen, YAO Shuping. Decision on optimal active response based on intrusion graph[J]. Journal of Beijing Polytechnic University, 2012, 38(11): 1659-1664.

[43] GARG U, SIKKA G, AWASTHI L K. Empirical analysis of attack graphs for mitigating critical paths and vulnerabilities[J]. Computers & Security, 2018, 77: 349-359.

[44] MENG W, LI W, WANG Y, et al. Detecting insider attacks in medical cyber-physical networks based on behavioral profiling [J]. Future Generation Computer Systems, 2018,108: 1258-1266.

[45] KAMMULLER F, PROBST C W. Modeling and verification of insider threats using logical analysis[J]. IEEE Systems Journal, 2017, 11(2): 534-545.

[46] MAVROEIDIS V, VISHI K, JOSANG A. A framework for data-driven physical security and insider threat detection[C]// Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. Piscataway: IEEE Press, 2018: 1108-1115.

[47] MUSMAN S, TURNER A J. A game oriented approach to minimizing cybersecurity risk[J]. International Journal of Safety and Security Engineering, 2018, 8(2): 212-222.

[48] BARACALDO N, PALANISAMY B, JOSHI J. G-SIR: An insider attack resilient geo-social access control framework[J]. IEEE Transactions on Dependable and Secure Computing, 2019, 16(1): 84-98.

[49] SAFA N S, MAPLE C, WATSON T, et al. Motivation and opportunity based model to reduce information security insider threats in organizations[J]. Journal of Information Security and Applications, 2018, 40: 247-257.

[50] GREITZER F L, PURL J, LEONG Y M, et al. SOFIT: Sociotechnical and organizational factors for insider threat[C]// Proceedings of the 2018 IEEE Security and Privacy Workshops. Piscataway: IEEE Press, 2018: 197-206.

[51] GREITZER F L, LEE J D, PURL J, et al. Design and implementation of a comprehensive insider threat ontology[J]. Procedia Computer Science, 2019, 153: 361-369.

[52] PATCHA A, PARK J M. An overview of anomaly detection techniques: Existing solutions and latest technological trends[J]. Computer Networks, 2007, 51(1251): 3448-3470.

[53] ZHANG R, CHEN X J, SHI J Q, et al. Detecting insider threat based on document access behavior analysis[C]// Proceedings of the Asia Pacific Web Conference Workshops. Berlin: Springer, 2014: 376-387.

[54] CAMINA J B, RODRIGUEZ J, MONROY R. Towards a masquerade detection system based on user's task[C]// Proceedings of the International Symposium on Recent Advances in Intrusion Detection. New York: ACM Press, 2014: 447-465.

[55] CAMINA J B, MONROY R, TREJO L A, et al. Towards building a masquerade detection method based on user file system navigation[C]// Proceedings of the Mexican International Conference on Artificial. Berlin: Springer, 2011: 174-186.

[56] RAY I, POOLSAPASSIT N. Using attack trees to identify malicious attacks from authorized insiders[C]// Proceedings of the European Symposium on Computer Security (ESORICS). Berlin: Springer, 2005: 231-246.

[57] LIU A, MARTIN C, HETHERINGTON T, et al. A comparison of system call feature representations for insider threat

detection[C]// Proceedings of the IEEE SMC Information Assurance Workshop. Piscataway: IEEE Press, 2005: 340-347.

[58] SHARIF M, URAKAWA J, CHRISTIN N, et al. Predicting impending exposure to malicious content from user behavior [C]// Proceedings of the ACM Conference on Computer and Communications Security (CCS). New York: ACM Press, 2018: 1487-1501.

[59] WUCHNE T, CISLAK A, OCHOA M, et al. Leveraging compression-based graph mining for behavior-based malware detection[J]. IEEE Transactions on Dependable and Secure Computing, 2019, 16(1): 99-112.

[60] LYER A, NGO H Q. Towards a theory of insider threat assessment[C]// Proceedings of the International Conference on Dependable Systems & Networks (DSN). Piscataway: IEEE Press, 2005: 108-117.

[61] NASR M, SHOKRI R, HOUMANSADR A. Machine learning with membership privacy using adversarial regularization[C]// Proceedings of the ACM Conference on Computer and Communications Security (CCS). New York: ACM Press, 2018: 634-646.

[62] EBERLE W, GRAVES J, HOLDER L. Insider threat detection using a graph-based approach[J]. Journal of Applied Security Research, 2010, 6(1): 32-81.

[63] LEGG P A, BUCKLEY O, GOLDSMITH M, et al. Automated insider threat detection system using user and role-based profile assessment[J]. IEEE System Journal, 2015 11(2): 1-10.

[64] AGRAFIOTIS I, LEGG P, GOLDSMITH M, et al. Towards a user and role-based sequential behavioral analysis tool for insider threat detection[J]. Journal of Technology Transfer, 2014, 4:127-137.

[65] 王辉, 杨光灿, 韩冬梅. 基于贝叶斯网络的内部威胁预测研究[J]. 计算机应用研究, 2013(9): 2767-2771.
WANG Hui, YANG Guangcan, HAN Dongmei. Research of predicting insider threat based on Bayesian network[J]. Application Research of Computers, 2013(9): 2767-2771.

[66] 王辉, 刘淑芬. 一种可扩展的内部威胁预测模型[J]. 计算机学报, 2006, 29(8): 1346-1355.
WANG Hui, LIU Shufen. A scalable predicting model for insider threat[J]. Chinese Journal of Computers, 2006, 29(8): 136-1355.

[67] 潘恒, 李景峰, 马君虎. 可抵御内部威胁的角色动态调整算法[J]. 计算机科学, 2020, 47(5):6.
PAN Heng, LI Jingfeng, MA Junhu. Role dynamic adjustment algorithm for resisting insider threat[J]. Computer Science, 2020, 47(5): 313-318.

[68] 付钰, 李洪成, 吴晓平, 等. 基于大数据分析的APT攻击检测研究综述[J]. 通信学报, 2015, 36(11): 1-14.
FU Yu, LI Hongcheng, WU Xiaoping, et al. Detecting APT attacks: A survey from the perspective of big data analysis[J].Journal on Communications, 2015, 36(11): 1-14.

[69] 胡伟, 洪熠, 王静雅. 基于活动行为特征的APT攻击检测方法研究[J]. 中国高新科技, 2021, 14: 2.
HU Wei, HONG Yi, WANG Jingya. Research on apt attack detection method based on activity behavior characteristics[J].Chinese High and New Technology, 2021, 14: 2.

[70] 刘海波, 武天博, 沈晶, 等. 基于GAN-LSTM的APT攻击检测[J]. 计算机科学, 2020, 47(1): 281-286.
LIU Haibo, WU Tianbo, SHEN Jing, et al. Advanced persistent threat detection based on generative adversarial networks and long short-term memory[J]. Computer Science, 2020, 47(1): 281-286.

[71] CHAO D X, HOANGA D M, DINHA N H. APT attack detection based on flow network analysis techniques using deep learning[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(3): 4785-4801.

[72] HONG S P, LIM C H, LEE H J. APT attack response system through AM-HIDS[C]// Proceedings of 2021 23rd International Conference on Advanced Communication Technology (ICACT). Berlin: Springer, 2021.

[73] AHMAD K, HWAITAT A, MANASEER S, et al. An investigator digital forensics frequencies particle swarm optimization for detection and classification of APT attack in fog computing environment[J]. Journal of Theoretical and Applied Information Technology, 2020, 98(7): 937-952.

[74] 李立勋, 张斌, 董书琴, 等. 基于脆弱性变换的网络动态防御有效性分析方法[J]. 电子学报, 2018, 46(12): 3014-3020.
LI Lixun, ZHANG Bin, DONG Shuqin, et al. Effectiveness analysis approach based on vulnerability mutation for network dynamic defense[J]. Acta Electronica Sinica, 2018, 46(12): 3014-3020.

[75] JI Y J, ZHANG X Y, JI S L, et al. Model-reuse attacks on deep learning systems[C]// Proceedings of the ACM Conference

on Computer and Communications Security (CCS). New York: ACM Press, 2018: 349-363.

[76] MITROPOULOS D, LOURIDAS P, POLYCHRONAKIS M, et al. Defending against web application attacks: Approaches, challenges and implications[J]. IEEE Transactions on Dependable and Secure Computing, 2019, 16(2): 188-203.

[77] PUSARA M, BRODLEY C. User Re-authentication via mouse move-motets[C]// Proceedings of ACM Workshop on Visualization and Data Mining for Computer Security. New York: ACM Press, 2004: 1-8.

[78] KILLOURHY K, MAXION R. Why did my detector do that?! —Predicting keystroke-dynamics error rates[J]. Recent Advance in Intrusion Detection, 2010, 13: 256-276.

[79] MESSERMAN A, MUSTAFIC T, CAMTEPE S, et al. Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics[C]// Proceedings of the International Joint Conference on Biometrics. Piscataway: IEEE Press, 2011: 1-8.

[80] SPITZNER L. Honeypots: Catching the insider threat[C]// Proceedings of the Computer Security Applications Conference. Piscataway: IEEE Press, 2003: 170-179.

[81] SPITZNER L. Honeypots: Tracking hackers[M]. [S.l.]: Addison Wesley Press, 2003.

[82] BOWEN B M, SALEM M, HERSHKOP S, et al. Designing host and network sensors to mitigate the insider threat[J]. IEEE Security & Privacy, 2009, 7(6): 22-29.

[83] KANDIAS M, MYLONAS A, VIRVILIS N, et al. An insider threat prediction model[C]// Proceedings of the Trust, Privacy and Security in Digital Business. Berlin:Springer, 2010: 26-37.

[84] 周余阳, 程光, 郭春生, 等. 移动目标防御的攻击面动态转移技术研究综述[J]. 软件学报, 2018, 29(9): 2799-2820.
ZHOU Yuyang, CHENG Guang, GUO Chunsheng, et al. A survey on attack surface dynamic transfer technology based on moving target defense[J]. Journal of Software, 2018, 29(9): 2799-2820.

[85] HUANG L, ZHU Q Y. Duplicity games for deception design with an application to Insider threat mitigation[J]. IEEE Transactions on Information Forensics and Security, 2021,16: 4843-4856.

[86] YUAN S H, WU X T. Deep learning for insider threat detection: Review, challenges and opportunities[J]. Computers & Security, 2021, 104: 234-244.

[87] LE D C, HEYWOOD N Z. Exploring anomalous behavior detection and classification for insider threat identification[J]. International Journal of Network Management, 2020, 31(4): 434-456.

[88] MOHAMMED N A, AHMED R, ABIDIN Z, et al. An integrated imbalanced learning and deep neural network model for insider threat detection[J]. International Journal of Advanced Computer Science and Applications, 2021,12(1): 573-577.

[89] WALL A, AGRAFIOTIS I. A Bayesian approach to insider threat detection[J]. Journal of Wireless Mobile Networks, 2021, 12(2): 48-84.

[90] LE D C, HEYWOOD N Z. Exploring adversarial properties of insider threat detection[C]// Proceedings of 2020 IEEE Conference on Communications and Network Security (CNS). Basel,Switzerland: MDPI, 2020.

[91] MARK D G, MARC W B. Insider threat program best practices[C]// Proceedings of the 46th Hawaii International Conference on System Science. Piscataway: IEEE Press, 2013: 1831-1839.

[92] 陆军, 刘大昕, 战扬. 基于Agent的内部威胁监视系统的动态管理[J]. 计算机研究与发展, 2006, 43(Z1): 341-346.
LU Jun, LIU Daxin, ZHAN Yang. Dynamic management of internal threat watching system based on agent[J]. Journal of Computer Research and Development, 2006, 43(Z1): 341-346.

[93] AZARIA A, RICHARDOSN A, KRAUS S, et al. Behavioral analysis of insider threat: A survey and bootstrapped prediction in imbalanced data[J]. Computational Social Systems IEEE Transactions, 2014, 1(2): 135-155.

[94] 陈亚辉. 层次化内部威胁态势量化评估模型的研究和分析[D]. 长沙: 国防科学技术大学, 2008.
CHEN Yahui. Layering of quantitative evaluation model for the situation insider threat[D]. Changsha: National University of Defense Technology, 2008.

[95] 陈小军. 意图驱动的内部威胁检测技术研究[D]. 北京: 中国科学院大学, 2014.
CHEN Xiaojun. Research on intentional-driven insider threat detection[D]. Beijing: University of Chinese Academy of Sciences, 2014.

[96] KANDIAS M, VIRVILIS N, GRITZALIS D. The insider threat in cloud computing[C]// Proceedings of the International

Workshop on Critical Information Infrastructures Security. Berlin, Heidelberg: Springer Press, 2013: 93-103.

[97] 王国峰, 刘川意, 潘鹤中, 等. 云计算模式内部威胁综述[J]. 计算机学报, 2017, 40(2): 296-316.
WANG Guofeng, LIU Chuanyi, PAN Hezhong, et al. Survey on insider threats to cloud computing[J]. Chinese Journal of-Computers, 2017, 40(2): 296-316.

[98] KIM A, OH J, RYU J, et al. A review of insider threat detection approaches with IoT perspective[J]. IEEE Access, 2021, 8: 78847-78867.

[99] 张红斌, 裴庆琪, 马建峰. 内部威胁云模型感知算法[J]. 计算机学报, 2009, 32(4): 784-792.
ZHANG Hongbin, PEI Qingqi, MA Jianfeng. An algorithm for sensing insider threat based on cloud model[J]. Chinese Journal of Computers, 2009, 32(4): 784-792.

[100] 张红斌, 尹彦, 赵冬梅, 等. 基于威胁情报的网络安全态势感知模型[J]. 通信学报, 2021, 6: 182-194.
ZHANG Hongbin, YIN Yan, ZHAO Dongmei, et al. Network security situational awareness model based on threat intelligence [J]. Journal on Communications, 2021, 6: 182-194.

作者简介:

郭世泽(1969-),男,教授,博士生导师,研究方向:网络安全信息安全,E-mail: szguo.uestc @gmail.com。

张磊(1989-),男,博士研究生,研究方向:人工智能与网络安全,E-mail: zhanglei@aeu. edu.cn。

潘雨(1990-),女,博士,工程师,研究方向:社团发现与机器学习,E-mail: panyu@aeu.edu. cn。

陶蔚(1991-),男,博士,助理研究员,研究方向:模式识别机器学习,E-mail: taowei@aeu.edu. cn。

白玮(1983-),男,博士,讲师,研究方向:网络安全、网络运维脆弱性,E-mail: baiwei_ lgdx@126.com。

郑奇斌(1990-),男,博士,助理研究员,研究方向:数据挖掘机器学习,E-mail: zqb1990@hot mail.com。

刘艺(1990-),男,博士,助理研究员,研究方向:机器学习、进化算法,E-mail:abertliu20th@ 163.com。

潘志松(1973-),**通信作者**,男,教授,博士生导师,研究方向:模式识别机器学习,E-mail: panzhisong@aeu.edu.cn。

(编辑:刘彦东)