

云数据中心内部网络安全防御关键技术

Yunshuju Zhongxin Neibu Wangluo
Anquan Fangyu Guanjian Jishu

■ 王欢 李华 陈占芳 赵建平 底晓强 冯欣 著



国防工业出版社
National Defense Industry Press

云数据中心内部网络 安全防御关键技术

王欢 李华 陈占芳 赵建平 底晓强 冯欣 著

国防工业出版社

·北京·

内容简介

本书主要介绍了以网络安全态势感知为核心的云数据中心网络主动防御技术,分别介绍了在网络安全态势感知的察觉、评估、预测以及控制阶段采取相应手段和方法,并将一系列智能算法应用于防护过程中。

本书主要面向计算机网络、网络与信息安全、数据中心自动化运维等领域的学者和技术人员,希望能够帮助读者在研究工作中有效整理研究思路、探索研究方法。

图书在版编目(CIP)数据

云数据中心内部网络安全防御关键技术/王欢等著.

—北京:国防工业出版社,2018.12

ISBN 978-7-118-11809-4

I. ①云… II. ①王… III. ①计算机网络-网络安全
全 IV. ①TP393.08

中国版本图书馆 CIP 数据核字(2018)第 299135 号

※

国防工业出版社 出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

三河市众誉天成印务有限公司印刷

新华书店经售

*

开本 710×1000 1/16 印张 7 字数 117 千字

2018 年 12 第 1 版第 1 次印刷 印数 1—1500 册 定价 60.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777

发行邮购:(010)88540776

发行传真:(010)88540755

发行业务:(010)88540717

前 言

随着云计算技术的发展,基于云计算环境的数据中心也越来越多地被广大运营商采纳。云数据中心存储着大量的核心数据,是各种业务系统的基础。云数据中心网络作为云内部网络,是云平台的核心支撑,备受攻击者的关注。云数据中心不仅要解决传统的安全问题,而且接受针对云环境的新型攻击考验,面临着前所未有的安全挑战。为此,如何实现对云数据中心网络的安全防护成为一个亟待解决的问题。传统的网络安全防护方法,大多数是通过防火墙等安全设备尽可能识别攻击行为,是一种被动式防御方法,不仅不能对未知攻击行为进行有效监测,而且无法对云数据中心的安全状况进行整体掌控,已经无法适应云数据中心的安全防护工作。为此,急需对云数据中心网络系统进行以网络安全态势感知和最优防御为核心的新型主动防御研究。本书针对云数据中心主动防御中存在的主要问题,深入研究网络攻击要素识别、脆弱性评估、安全态势评估预测等主动防御的核心问题,并实施最优防御,以保证云数据中心对外提供安全、可靠的服务。

以网络安全态势感知和最优防御为核心的主动防御技术,涉及安全要素提取、攻击识别、脆弱性评估、态势评估预测以及安全控制等多方面内容。通常情况下,态势感知的安全要素是从不同的安全设备中获得的,然后通过各类融合手段进行消重,所以本书未将安全要素提取作为研究内容,而是对攻击要素识别、网络脆弱性分析、态势评估预测以及最优策略选取进行了重点研究。

本书采用基于径向基神经网络的攻击要素识别实现对攻击行为的分类,基于攻击图模型的网络脆弱性分析完成对网络脆弱性的评估并确定关键脆弱性环节,二者是网络安全态势量化基础,分别为态势指标的攻击性态势和脆弱性态势提供计算依据,支撑态势评估。基于 GA-LSSVM 时间序列分析的网络安全态势预测是对网络安全态势量化与评估的进一步深化,以历史的态势数据为基础,实现对短期安全态势的预测,掌控态势的发展趋势。基于非合作博弈与粒子群优化的最优防御是在态势感知和预测结果之上,实施最优的安全防护策略,以对安全风险进行有效控制,实现云数据中心网络的“适度安全”。

针对以上研究过程和研究成果,本书提供了实验仿真部分,对研究工作中的

相关概念、方法以及模型进行验证。在陈述过程中,采用了示例与内容相结合的方式帮助读者理解相关的研究内容,并通过各章的引言,支持读者根据个人的实际情况对本书内容进行阅读上的取舍。

本书的研究内容和研究思路还获得了长春理工大学信息化中心老师给予的支持和帮助。由于作者水平有限以及本书成稿仓促,难免有不足之处,欢迎广大读者批评指正。

著者

目 录

第 1 章 绪论	1
1.1 概述	1
1.1.1 背景	1
1.1.2 研究意义	2
1.2 网络安全防护相关技术	3
1.2.1 网络威胁识别分类	3
1.2.2 脆弱性评估	4
1.2.3 网络安全态势评估	5
1.2.4 网络安全态势预测	6
1.2.5 安全加固策略选择	7
1.3 现有技术面临的问题	8
1.4 本书的内容及结构安排	10
1.4.1 主要内容	10
1.4.2 结构安排	13
第 2 章 基于径向基神经网络的攻击要素识别	15
2.1 引言	15
2.2 径向基神经网络	15
2.2.1 径向基神经网络模型	16
2.2.2 径向基神经网络解决分类问题的理论基础	17
2.3 基于径向基神经网络的分类模型	17
2.4 基于改进 K-均值算法的输入样本聚类	18
2.4.1 改进的基本思想	19
2.4.2 算法具体实现	19
2.4.3 算法复杂度分析	22
2.5 基于梯度—强化学习的网络参数训练	22
2.5.1 梯度算法	22
2.5.2 基于 Q 学习的学习率调整方法	24

2.6	基于 OLS 的隐含层至输出层权值训练	25
2.6.1	OLS 训练算法基本思想	25
2.6.2	算法具体实现	27
2.7	实验分析	28
2.7.1	实验数据准备	28
2.7.2	攻击要素分类实验	28
2.7.3	比较实验	32
第 3 章	基于攻击图模型的网络脆弱性分析	34
3.1	引言	34
3.2	基于攻击图的脆弱性分析模型	34
3.2.1	脆弱性评估框架	34
3.2.2	脆弱性攻击图模型化表示	36
3.3	基于 CVSS 的脆弱性危害量化	37
3.4	脆弱性攻击图生成	39
3.4.1	脆弱性节点遍历匹配	39
3.4.2	攻击图生成	40
3.4.3	算法复杂度分析	41
3.5	最佳攻击路径确定	42
3.5.1	最大损失流算法	42
3.5.2	路径搜索	43
3.5.3	多目标增广路排序	44
3.5.4	算法复杂度分析	45
3.6	实验仿真	45
3.6.1	实验环境	45
3.6.2	实验场景的攻击图	46
3.6.3	攻击路径搜索性能分析	46
3.6.4	指标量化结果分析	48
第 4 章	基于 AHP 的网络安全态势量化与评估	53
4.1	引言	53
4.2	基于 AHP 的网络安全态势评估模型	53
4.2.1	态势评估框架	53
4.2.2	评估模型形式化表示	55
4.3	基于 D-S 证据理论的攻击要素融合	56
4.3.1	告警数据的模糊化表示	56

4.3.2	基于 D-S 证据理论的攻击要素融合	57
4.4	态势评估指标的量化计算	58
4.4.1	风险性态势	58
4.4.2	基础运行性态势	59
4.4.3	破损性态势	59
4.5	态势评估指标权重因子确定	59
4.5.1	构造成对比较矩阵	59
4.5.2	特征向量计算	60
4.5.3	最大特征值计算	60
4.5.4	一致性校验	61
4.6	态势指标的聚合	61
4.6.1	评价指标聚合	61
4.6.2	节点态势聚合	61
4.7	实验分析	62
第 5 章	基于 GA-LSSVM 时间序列分析的网络安全态势预测	69
5.1	引言	69
5.2	基于 LSSVM 的预测模型	69
5.2.1	模型定义	69
5.2.2	核函数选择	71
5.2.3	数据归一化处理	71
5.3	基于 GA 的 LSSVM 联合参数优化	72
5.3.1	染色体编码	72
5.3.2	染色体操作	73
5.3.3	适应度函数	74
5.3.4	染色体生成	74
5.3.5	联合优化算法	75
5.4	基于 GA-LSSVM 的预测方法	76
5.5	仿真实验分析	76
5.5.1	实验环境	76
5.5.2	评价方法	77
5.5.3	结果分析	77
第 6 章	基于非合作博弈与粒子群优化的最优防御	81
6.1	引言	81
6.2	非合作攻防博弈模型	82

6.2.1	模型定义	82
6.2.2	模型假设	82
6.3	攻防收益量化及计算方法	83
6.4	攻防博弈的纳什均衡	84
6.5	基于 PSO 的最优攻防策略选取算法	85
6.5.1	基于 PSO 的最优攻防策略选取算法	85
6.5.2	算法分析	90
6.6	实验分析	90
6.6.1	实验环境	90
6.6.2	攻防决策实验	91
6.7	进一步讨论	94
参考文献		95

第 1 章 绪 论

1.1 概 述

1.1.1 背景

云计算以互联网为依托,是一种能够为用户按需快速提供高质量资源的新型计算模式^[1],具有成本低、管理便捷和可靠性高等特点。当前,云计算已被业界公认为将对人类社会产生深远影响的革命性技术。世界各国都已经将其作为重点发展的战略领域,并建设了自己的云计算项目,如美国的“星云计划”、日本的“霞关计划”、英国的“G-cloud 计划”等。为了占领云计算技术这一制高点,我国也颁布了《国务院关于加快培育和发展战略性新兴产业的决定》重点支持云计算产业的发展。我国的政府机关、企事业单位和科研机构也正在紧锣密鼓地进行着一系列的云计算项目,如北京“祥云工程”行动计划、上海市的“云海计划”、阿里巴巴的“云服务平台”、中国移动的“Big Cloud”等。

虚拟化、数据托管和外包服务是云计算的特点。这些特点也使得云计算面临着前所未有的安全挑战,主要表现为拒绝服务攻击、篡改消息、伪造身份等传统的攻击,以及由多点多租户共享带来的侧通道攻击、高级持续性威胁等新型攻击。虽然国内外专家学者已经对云安全进行了大量研究,但主要集中于数据的加密、访问控制和行为认证等方面,对云数据中心内部网络的安全防护技术研究还处于起步阶段,成果较少。

云数据中心是支持云平台的基础,存储着大量的核心数据,是一种基于云计算架构,具有耦合度低、设备虚拟化、自动化程度高等特点的新型数据中心。为了更好地为广大用户服务,国内的运营商也正在逐步地将传统的数据中心向云数据中心过渡。截止到目前,我国已经建立了 10 万台服务器以上的云数据中心 60 余个,价值 6000 多亿元。云数据中心网络作为云内部网络,是云平台的核心支撑,具有较强的扩展性和自治性^[2],这给传统的网络安全防护体系造成了极大的冲击,使其无法直接用于云数据中心。同时,由于云数据中心网络结构也由传统的层次式发展为递归式,出现了一系列新的网络结构。如果对这些新的网

络实例进行安全防护,就要结合其特点,建立一套新的适用于云数据中心的安全防护体系。由此可知,研究面向云数据中心网络的安全防护体系已经刻不容缓,势在必行。

1.1.2 研究意义

云数据中心网络通过高速的物理链路将网络内的交换设备、服务器以及存储设备互连起来,利用虚拟化技术构建共享资源池,为数据中心的应用和服务提供数据传输、存储和计算功能^[3]。由于数据中心网络的这些特点和存储着业务系统的核心数据,因此备受攻击者的关注。同时,随着互联网技术的发展,针对云数据中心特点,也涌现出许多新的攻击方式和手段^[4-7]。如自“棱镜门事件”后,出现的高级持续性威胁(Advanced Persistent Threat, APT)。APT攻击与传统的攻击不同,多采用多步攻击方式进行。攻击者在发动攻击前,潜伏在目标网络内部,长期监测和分析系统内的网络流量与不同主机间的漏洞关系。当需要发起攻击时,攻击者通过多步行动同时进行,控制更多的内部资源,有计划、有步骤地进行渗透,最终破坏目标网络,窃取数据。此攻击具有高度的隐蔽性,很难监控和预防。由此可见,云数据中心面临着巨大的安全威胁。传统安全防御方法是采用防火墙、入侵检测系统、流量审计系统等对特征明显的攻击进行识别,属于被动式的防御,既不能对多阶段的隐蔽攻击进行监测,也不能对整个网络态势进行监控,无法适应云数据中心的安全防护工作^[8-10]。

为了保护云数据中心中的数据资产安全,解决日益突出的云数据中心网络安全防护问题,学术界和工业界认为,除了设计更加安全的系统,尽可能地识别攻击的发生,还需要对云数据中心网络系统进行以网络安全态势感知和最优防御为核心新型主动防御^[11-13]。主动防御技术能够动态地评估网络安全态势走向,并根据态势评估和预测的结果实施适当的安全防御策略,进而对网络安全态势进行有效的控制。因此,主动防御可以有效地遏制潜在威胁,从而降低潜在威胁带来的负面影响和巨大损失。进行主动防御的研究意义包括以下四个方面:

- (1) 监测未知的、多阶段的、隐蔽的攻击行为,系统、全面地识别云数据中心网络当前面临的安全风险;
- (2) 分析脆弱性及脆弱性之间的依赖关系对网络的影响,从网络整体性的角度评估网络的脆弱性,发现关键脆弱环节;
- (3) 能够将多种安全设备识别的攻击信息融合,并以此为依据对网络节点状态进行量化,实现对云数据中心网络安全态势的评估,从宏观上掌握网络安全状况,并对网络安全态势的发展趋势进行准确、有效的预测;

(4) 根据脆弱性评估和态势评估的结果,对云数据中心采取适当的安全防御策略,保证云数据中心安全、可靠运行。

1.2 网络安全防护相关技术

1.2.1 网络威胁识别分类

网络攻击要素识别是网络安全态势感知的重要环节。在网络系统内部会部署多种安全设备,对网络中的数据进行监测和采集。如果完成网络安全态势评估,就必须对采集到的网络流量进行攻击要素识别。网络安全要素识别主要是根据采集到数据的属性及特征,按照攻击特点将其分为正常、Dos、Probe、R2L、U2R 等类别,因此,对攻击要素进行识别是一个非线性数据的多分类问题。本书将目前对网络非线性数据进行多分类的方法研究进行了以下梳理。

文献[14]基于贝叶斯理论,对海量数据进行处理,提出了一种新的非线性数据分类模型及方法。模型的实现过程中,以非线性网络数据的属性为基础依据,进行数据的不同类别区分。此模型虽然实现了数据分类,但方法较为简单,分类结果存在较大的噪声。文献[15]提出了一种基于神经网络算法的海量非线性网络流量数据建模方法。该方法利用神经网络的泛化特性,将非线性数据映射到高维空间,使其线性可分。该方法具有抗噪性强特点,但在分类过程中,计算量大、复杂度高、训练速度慢。文献[16]提出了一种基于贝叶斯和神经网络相结合的网络流量数据分类模型。模型采用贝叶斯网络进行推理,利用神经网络进行学习训练,提高了分类的速度,具有较强的适应性,但由于神经网络对训练集的要求较高,故其存在准确率低的问题。文献[17]基于支持向量机(SVM)算法,建立了网络流量数据分类模型,并提出了一种新的分类方法。该方法在分类的初始阶段对训练样本进行聚类 and 压缩,并通过样本数据中聚类特征贡献的大小确定模糊因子,使用特征有效度建立核函数,优化分类精度。但该模型在处理海量数据过程中训练速度慢。文献[18]提出了一种新的基于置信规则库的组合分类模型,模型在实习的过程中,利用半定量信息为依据进行要素分类;通过有向无环图组合置信规则库,避免了分类过多导致精度下降的问题,但其模型分类速度较慢。文献[19]基于主成分分析(PCA)神经网络进行分类器设计,设计了一种适用于入侵检测的分类模型。模型对神经网络的学习算法进行了改进,在分类过程中,动态地确定类别数据的空间维数,提高了分类精度。但该模型需要提供丰富的先验知识,对先验知识依赖较高。

从上述研究成果可以看出,基于神经网络的分类方法以其较强的学习能力

和自适应性备受关注。反向传播(BP)神经网络和径向基函数神经网络都是非线性多层前向网络,它们都是通用逼近器,都能够实现非线性数据的多分类。但由于网络结构、训练算法以及逼近性能不同,也使得二者应用的场景不尽相同。BP神经网络采用BP学习算法,其算法本身就存在收敛速度慢和局部极小等不足。但基于RBF神经网络支持在线学习与离线训练方式,不仅能够动态地优化网络结构、确定隐含层数据中心,还可以根据训练过程调整扩展常数,具有较快的学习速度和较强自适应性,学习训练能力强,已经成功地用于非线性函数逼近、时间序列分析、数据分类、模式识别、信息处理、图像处理、系统建模、控制和故障诊断等领域。

1.2.2 脆弱性评估

网络脆弱性分析是指通过对网络中存在的脆弱性综合评判,分析网络可能遭受的安全威胁以及预测攻击者利用脆弱性可能发动的攻击路径,从而反映网络的安全状态。网络系统的脆弱程度直接关系到攻击者能否轻易入侵到系统内部,并对系统进行破坏和控制,影响着网络安全态势的发展。因此,脆弱性分析已经成为态势感知中的重要环节。

传统脆弱性评估方法大都针对网络节点的孤立脆弱性的评估。但是,在网络系统中,由于网络威胁的传递性,威胁不仅会来自直接相连网络节点,而且可能来自网络深处间接连接的网络节点。攻击者可以利用网络中节点间的连接关系,将分布在不同节点上的若干独立脆弱性组合起来,联合攻击目标。独立脆弱性可能不会对网络造成严重危害,但多个脆弱性组合起来可能对网络造成严重伤害。因此,对网络安全风险评估不仅考虑独立脆弱性,而且考虑脆弱性间的关联关系。

近年来,国内外很多专家学者对整体网络脆弱性评估模型开展了研究,并取得了一些成果。文献[20]利用贝叶斯网络分离的思想,建模脆弱性之间的关系,提出了一种基于贝叶斯网络的脆弱性评估模型。通过在相关性节点间加入分离节点,将原有关系分解为相关性节点与分离节点的关系,并使用条件概率解决节点相关性导致的概率计算错误问题。文献[21]提出了一种基于攻击图的脆弱性分析模型。模型将安全概率攻击图作为攻击图概率。通过最大可达概率和删除不可达路径来解决循环路径概率重复计算问题。以脆弱性评分系统(Common Vulnerability Scoring System, CVSS)^[22]为基础,进行安全的概率计算。模型能够适应大规模网络的脆弱性评估,时间复杂度小。但该方法对渗透之间的相关性对网络安全概率的影响没有做出说明。文献[23,24]通过脆弱性分析,提出按照修复集的思想对网络系统进行安全加固。通过对脆弱性以及其节

点关系,构造初始条件修复集,并通过布尔表达式的析取范式的计算,对初始条件进行修复,最后通过最优化理论求解最佳安全策略。该方法只是单纯地考虑主机节点间的关系,而没有考虑脆弱性本身和其利用关系。此外,该方法的复杂度为幂指数级,不是大规模网络场景。文献[25]基于 MulVAL 和攻击图建立网络脆弱性评估系统。系统对 PageRank^[26] 算法进行改进,提出了一种能够对大规模攻击图分析的 AssetRank 算法。算法按照节点的重要性将其分类,并采用不同的颜色标记不同等级,方便安全人员处理脆弱性严重的问题。该系统虽然实现了脆弱性分析和访问控制,但对主机的最弱等级划分和脆弱性利用难易程度方面缺乏考虑。文献[27]提出了一种基于攻击图的脆弱性评估模型及方法。模型将攻击和初始条件进行加权和拆分处理,将攻击和初始条件的修复问题转换为最小 S-T 割集问题,并用搜索路径的方式进行关键脆弱性求解,其求解时间复杂度为 $O(M+N)$,能够实现脆弱性的评估和安全加固;但该方法在最佳弥补集的求解过程中没有考虑重复策略的消解问题。文献[28]对攻击图模型进行了扩展,设计了状态攻防图模型,给出了状态攻防图的生成算法,算法在搜索过程中,自动过滤与目标无关的脆弱性,缩小了问题的解空间,提高求解速度;但缺乏对脆弱性的量化和主机之间依赖关系的考虑,评估准确度不高。文献[29]提出了一种基于网络中心性的脆弱性评估模型,模型以 CVSS 为基础,对攻击者利用脆弱性的成本进行量化,并根据量化结果利用攻击图技术分析渗透路径,采用攻击图节点的介数和节点连通度相结合的方法,对关键脆弱性分析,实现对网络脆弱性的评估;但该模型在关键路径求解的过程中复杂度过高。

1.2.3 网络安全态势评估

网络安全态势评估是以网络中检测到的安全事件为基础,通过建立相应的指标体系,实现对安全态势值的量化,以完成对网络安全状态进行评定。本书将国内外对网络安全态势评估的主要研究进行了梳理。

文献[30]将攻击能力增长作为最终目标,建立一种新的网络安全分析模型。模型提高了攻击图的准确度,使路径分析更加精确,但模型在量化环节较为薄弱。文献[31]建立了基于 D-S 证据理论的多源信息融合态势评估模型。模型使用 D-S 证据理论融合多种检测设备的告警信息,并通过态势因子和节点态势的综合计算得到网络的安全态势。此模型实现了对态势的评估,但由于 D-S 证据理论在规则合成过程中会出现违背常理的情况,故结果有些时候会不符合实际。文献[32]基于信息融合的思想,建立基于多传感器融合的态势评估模型。模型在对异质传感器融合的过程中,使用 SVM 作为融合引擎,并通过特征简约,提高数据的融合效率。文献[33]则基于马尔可夫博弈分析理论建立了态

势感知模型。模型通过马尔可夫方法建模威胁传播的过程,分析传播过程对网络态势的影响。进而选择相应的加固策略,能够有效地阻止威胁,实现态势风险控制。文献[34]根据生物免疫的原理,建立了基于免疫原理的态势评估模型。模型通过提高免疫力方法进行风险控制,采用活力自学习方式对告警信息处理。模型适用于无线自适应网络态势感知,但是缺乏较为完全的定量分析体系。文献[35,36]利用神经网络对态势源数据进行训练,划分安全要素的种类,并通过隐马尔可夫模型建模态势变化过程,实现对态势的感知和预测。文献[37]提出了一种基于分布式入侵检测数据融合的态势评估模型及方法。模型利用入侵检测系统(IDS)的分布式传感器采集告警信息,然后通过数据融合,消除重复的告警信息。通过数据挖掘进行要素提取,实现态势感知。该模型虽然实验效果良好,但缺乏原型系统实现。文献[38]详细地阐述了适用于局域网攻击检测、态势评估和响应评估的 SSARE 系统的实现过程及原理。该系统虽然能够实现态势评估和安全决策,但数据来源单一,容易产生告警误差。文献[39]通过在网络设备上部署异常检测代理的方式,建立了基于异步网络流的态势评估方法。方法利用代理检测异常信息,然后通过控制中心合成态势曲线,实现态势评估。但该方法只是考虑了攻击,而没有考虑网络脆弱性对态势的影响。文献[40]提出了基于报警信息的安全态势评估模型。模型从 Honeynet 采集网络的行为信息,并使用 Bro 的工具提取告警,然后构建态势走向曲线。该模型虽然构造态势曲线,但只有在大规模遭受攻击或感染病毒时才能起作用,适用场景单一。文献[41]提出了一种层次化的网络态势评估模型。模型按照自上而下、先局部后整体的思路对主机态势进行评估,并通过服务、主机、网络 3 个层次计算整个网络的安全态势,该模型能够很好地进行态势评估,但是模型的评价指标过于单一,计算体系过于简单,量化不够客观。

1.2.4 网络安全态势预测

网络安全态势预测是根据当前以及历史态势值对未来一段态势变化情况进行预测,使网络在受到攻击前能够采取应对措施。为了实现网络安全态势预测,国内外众多学者对网络安全态势预测进行了深入研究。文献[36]以隐马尔可夫模型为理论基础,建立网络安全态势预测模型 HMM-NSSP,通过对状态观察进行态势预测,并针对本模型提出了一种新的预测算法,模型可以很好地适应实际生产环境。文献[42]为了尽可能准确评估和预测网络安全状态,在研究量子粒子群算法(QPSO)的基础上,探索影响算法全局收敛性能的因素,形成一种基于进化策略的改进 QPSO。文献[43]利用集对分析的方法,并结合可信网中多源数据的确定和不确定性问题,针对可信网络连接(TNC)框架,进行了网络态势

感知(Cyberspace Situation Awareness ,CSA)体系的研究,并提出了一种预测方法(SPSAF)。文献[44]将BP神经网络应用到态势预测领域,建立了似然预测方法,将网络安全态势序列作为训练序列,并在训练过程中自动调整参数权值,解决了态势数据处理的局限性和依赖专家权值的问题,但该方法训练过程复杂,收敛慢。文献[45]利用隐马尔可夫模型进行用户行为检测,对用户行为模式和轮廓的表示形式进行了改进,并采用序列匹配法降低运算量,基于状态序列的概率对行为进行判决。该方法可以实现用户行为的识别,但没有给出原型系统的实现过程。文献[46]利用时空分析的方法建立态势预测模型,安全要素的提取从攻防双方扩展到网络环境中,从空间维度的角度分析安全要素及其相关关联,计算网络安全态势。但该模型计算时间复杂度较高。文献[47,48]均采用神经网络的方法建立了态势预测模型,利用神经网络对数据进行前期训练,但模型对训练样本的数量和规模要求较高,其结果的精确度受训练网络的影响较大。与文献[47,48]不同,文献[49,50]则通过数据挖掘的方法建立预测模型。模型对历史数据中潜在模式挖掘,来寻找态势的变化趋势。此方法可以实现态势预测,但数据挖掘的过程中存在规约、模式发现等理论问题,基于数据挖掘的预测模型还不是完全成熟^[51]。此外,文献[52]以ARIMA模型为理论基础,设计了基于多维分析的态势预测系统,并实现了预测功能;但该系统的预测结果容易受到实验前提假设影响,预测精度存在波动现象。文献[53]则根据网络态势预测特点,直接通过ARIMA模型进行态势预测;但缺乏平稳性实验过程,实验精度容易受到外界影响,出现波动。

1.2.5 安全加固策略选择

在网络安全防护过程中,由于管理员可支配的资源和个人能力有限,不可能将每一个脆弱环节都消除,也不可能对每一种攻击进行有效的防御。如何平衡网络安全中的风险和投入,提高防御成本的有效性,利用现有资源做出最佳的防御决策,实现网络“适度安全”,已经成为亟待解决的问题。为此,就产生了网络安全防护策略的最优选取问题。

为了实现网络“适度安全”目标,国内外专家学者对限定条件下的网络安全最优防护策略选取问题开展了大量研究。文献[54]使用转台攻击图,对网络中的关键节点进行最优防护策略的求解,并利用防护策略对每一步攻击进行阻断,实现对网络节点的安全防护;但该方法没有考虑全局的最佳防护,容易陷入局部最优解问题。文献[55]通过防御措施成本量化,计算每一步防护的成本,然后利用攻击图计算攻击路径上的最小防护成本,并通过选择其对应的防护策略,破坏实施攻击的前提条件,从而实现安全防护;该方法效果虽然好,但是计算复杂

度高,不适应大规模的网络环境。文献[56]采用攻击图方法计算每一个攻击的源头,并利用现有资源对攻击的源头进行阻止。这种基于攻击图的网络代价加固方法,虽然能够消除所有的攻击源头,但是其开销很大,存在成本过高的问题。文献[57]将网络安全防护策略选取转化为多目标优化问题,将每一步防护策略的选取映射为一个目标,然后通过多个目标的优化求解,计算最优策略。此方法能够实现最优策略选取,但当攻击步骤过多时,目标数量会增多,计算复杂度会增大,不适合复杂的攻击场景。文献[58]基于主机漏洞对网络节点的防御成本进行计算,建立了成本估计模型。模型在评估中通过利用漏洞消耗成本,确定漏洞的危害程度。该模型虽然实现了利用漏洞的成本计算和策略选取,但其漏洞计算体系过于简单,不能够客观地评价漏洞对系统的影响。文献[59]利用攻防博弈模型建模攻防状态的变化,通过量化攻防双方的收益成本,求解博弈过程中的纳什均衡,实现最优策略的选取。该模型虽然实现了攻防博弈的策略选取,但处理方法过于简单,缺乏对网络安全攻防过程中的攻防双方状态不断转换的思考,且在策略选取的过程中容易陷入局部最优解。文献[60]根据攻击图规模的不同,分别建立了大小规模攻击图,利用弥补集求解防御策略,并在求解二者最优弥补集过程中,提出了精确和近似求解方法,解决了不同规模情况下最优策略的求解问题。文献[61]通过利用攻击威胁的概率构造攻击图,建立安全防护策略的攻击图模型。模型将攻击威胁可能发生的概率作为路径的权值,并根据攻击的分支建立图的转换关系。模型在真实网络中进行了验证,效果良好;但该模型单纯地考虑威胁的概率,而没有考虑攻击场景的变化和资源限制。文献[62]将粗糙图和博弈论相结合,利用粗糙图建模攻击的变化过程,通过动态博弈,针对攻击动作求解最优策略,实现了动态、实时的防护策略选取。文献[63]针对云滴冻结攻击,提出了一种内部攻击检测的博弈优化模型,但该模型只针对云环境内部分布式拒绝服务(DDoS)攻击进行检测,没有考虑其他的攻击特点,适用面较窄。文献[64]阐述了各类博弈论模型在网络安全领域的应用。文献[65]提出了一个基于 Petri 网的随机博弈模型,解决了网络中的动态复杂博弈问题。文献[66]利用贝叶斯网络,对多个层级建模,推断系统可能达到的状态,建立了一个非标准的博弈框架,对复杂的分布式 DDoS 攻击进行评估,但该方法没有考虑到其他的网络攻击,适用场景单一。

1.3 现有技术面临的问题

从上述的国内外研究学者对云数据中心网络进行的主动防御研究中可以看出,云数据中心网络仍然面临 5 个方面的主要挑战: