# 基于深度强化学习的智能网络安全防护策略研究

任海娟
(山西工商学院,山西 太原 030000)

**摘要**:深度强化学习作为统计学习常见算法,将其应用于智能网络安全防护设计环节将取得显著效果,以此规避智能网络安全风险。文章简要分析智能网络安全常见问题,根据对问题的分析研究,总结智能网络安全防护优化设计具体目标,经由设计网络状态集合、细化网络动作集合、规范设计回报函数、强化数据分析功能等设计步骤的落实,智能网络将充分发挥安全服务作用,由此维护网络安全。本文提出了可行性措施,期待提升网络安全防护的有效性。

**关键词**:深度强化学习;智能网络;安全防护
**中图分类号**:TN325 **文献标志码**:A

## 0 引言

深度强化学习的本质是在分析中寻求最大化回报行为的综合流程。此算法的实施,能为智能网络安全防护系统的合理设计给出新指引。技术人员通过对智能网络体系中交互环境、状态信息以及智能动作关联性的分析,促进动作与状态的科学匹配,以便提出合理的安全防护决策。未来,相关技术人员要加强对安全防护渠道的深入研究,从而让人工智能网络在智能制造等多行业中彰显突出价值。

## 1 智能网络安全常见问题

### 1.1 网络攻击

人工智能网络是依托人工智能技术研发的网络系统,其常见网络安全问题包括网络攻击,即对抗性攻击。技术人员对神经网络算法中对应主体进行修改,决策执行过程将出现变化。如麻省理工学院曾修改玩具龟参数,而后在神经网络算法下将其纳入集合中,从而引发了网络出错。人工智能网络虽然在运行期间具有便捷性特征,但无法完全避免网络攻击的风险。源于网络攻击背景,智能网络视觉功能将减弱,从而无法在正常的网络服务中表现出稳定性、准确性的优势。因此,参数的调整极易加剧对抗性网络攻击发生的风险[1]。

### 1.2 软件威胁

智能网络平台运行期间会形成具有威胁性的应用软件,此类软件往往具有恶意负载隐匿性,这会导致软件在启动期间极易被窃取隐私信息,引发安全风险事件。有研究表明,危及智能网络安全的主要问题在于软件程序安全等级偏低,在软件携带病毒后,会影响软件程序正常功能。因此,软件威胁也是智能网络安全常见问题,技术人员须从软件安全监测层面应对不良风险。

### 1.3 数据中毒

基于深度强化学习研发的智能网络平台,也会产生数据中毒问题。深度学习算法虽然能够提供智能辅助服务,但在道德标准以及人类思维特征上未建立明确依据。这会致使智能网络在交互服务中,无法参照用户思维特征提供所需数据,甚至提出的交互语言违反既定条件。在智能网络活动中,深度强化学习算法导向下也会形成带有病毒的数据。此类数据在传递中将侵害平台安全,尤其是往年在智能网络服务中形成的面部识别功能,在遭受数据中毒风险后,系统将无法顺利识别人脸面部。鉴于此,智能网络安全问题是现如今改进智能网络系统的重要要求。智能网络安全防护决策既有着深刻的现实意义,也对我国网络领域智能化发展具有促进作用。

## 2 智能网络安全问题防护目标分析

### 2.1 完善网络防护架构

针对上述常见的智能网络安全问题,技术人员提出安全防护措施。其中,对于网络攻击问题,技术人员需要从网络防护架构的有效完善层面阻断攻击动作。实际上,之所以智能网络会形成对抗性攻击,主要源于网络系统架构设计缺少合理性,进而影响系统运行的安全性。技术人员在设计安全防护系统时,需要依据局域网,联合路由器重新树立网络体系,促使网络系统在信息传播中能够顺利地将隐私信息与常规网络资源区分开来,这样也更为直接地应用交换式集线器,自此消除网络攻击风险。

### 2.2 强化计算机软件监测

要想保证智能网络免遭软件威胁,系统还需要针对计算机中安装的软件进行有效监测,以便从监测方向上强化安全防护效果。关于计算机软件监测事项的落实,技术人员可以利用访问控制、网络权限控制、网络服务器监测等多种方法,严格筛选出带有威胁性的恶意软件,进而确保智能网络系统处于安全运行状态。软件威胁的形成与软件运行动态有关,若能及早

**作者简介**:任海娟(1986— ),女,山西太原人,实验师,学士;研究方向:大数据和计算机网络。

发现软件风险,便可以提早预防安全隐患。比如可以将智能网络服务对象划分为 3 类,即特殊用户、普通用户以及审计用户,而后为其发放不同控制权限,以免因权限混乱而无法发挥软件的实践作用[2]。

### 2.3 启动数据备份机制

在智能网络安全防护过程中,若产生数据中毒问题,也会影响数据输送质量。因此,在确定安全防护目标时,系统应积极启动数据备份机制,对智能网络系统中的有用数据进行备份处理。后期即使出现数据窃取或数据中毒情况,也能及时从备份数据中查询,以免削弱数据应用实效性。参照上述智能网络安全问题,经分析后确定安全防护目标,以此利用深度强化学习算法建立新型智能网络安全防护系统。

## 3 基于深度强化学习的智能网络安全防护解决对策

### 3.1 设计网络状态集合

深度强化学习算法在实际应用中,主要是从回报学习最优求取过程中获取最大期望回报值的行为,由此运用行动与状态的相关性,促使智能网络实现安全等级的合理提升。其中,深度强化学习算法具体包含动作(Action)、智能体(Agent)、回报、环境等元素,形成架构流程。经过对集合体元素的综合分析,能够顺利获取累计期望回报值($R$),具体可以借鉴下述公式加以分析,即 $R = E(\sum_{t=0}^{T}\gamma^{t}r_{t})$,该公式对应的 $E$、$r$、$\gamma$ 分别表示的是数学期望、回报(奖励、惩罚)、折扣因子(归属于 0~1 的范围内)。而 $T$ 则属于时间序列,$t$ 为现下记录的具体时刻。在智能网络安全防护中,智能体可以凭借环境状态对动作给出新要求,使其在数学函数运算中出具全新的执行指令,借此在智能体输入后判定当前状态的匹配度。这类智能网络安全防护系统的设计在应用深度强化学习算法进行改进后,实际上是以云计算技术为主体,建立模拟仿真场景,然后将环境数据与现实网络进行连接,便于系统在真实的网络服务中获取最优决策。在具体设计阶段,技术人员需要先行设计网络状态集合,从而在智能体终止迁移学习时,寻求最佳网络环境,预防网络安全风险[3]。

在设计智能网络安全防护系统阶段,技术人员应当注重状态集合的优化设计。作为涵盖智能网络状态的集合体,系统应参照状态结合映射网络动作,并从状态识别中预估网络安全风险程度,继而在攻守双方建立对应的对抗单元。关于状态集合的有效设计,技术人员具体需要从智能网络攻击的“攻击”“防守”两个部分整理状态元素。前者可以将网络攻击种类、攻击范围以及攻击轨迹、攻击来源、攻击方法与攻击速度纳入攻击单元集合中。后者以智能网络安全服务等级、安全域、网络应用服务、结构元素与网络策略等多项元素为主[4]。根据有关状态元素集合的分析,系统可从中全方位知晓安全风险等级,之后以网络动作的调整下达正确的网络服务指令。考虑到网络状态集合中的元素较为多样,所以在设计状态集合时,技术人员还要充分借助攻守单元的状态信息,做好划分归类工作。如智能网络系统连接的路由器、计算机设备遭受攻击,系统可从状态集合的防守单元中,查询到适宜的对抗元素,以达成安全防护效果。

### 3.2 细化网络动作集合

技术人员在设计智能网络安全防护系统期间,还需要针对网络动作集合进行细化设计。因网络状态与网络动作本身存在映射关系,根据上文提出的多项状态元素,建立动作集合。系统按照动作集合中动作要素的类别向网络空间传输有用信息。对应的 Action 结构包含 4 种类型的元素。在“what”中泛指网络操作动作与网络行动动作,“who”对应网络用户与网络服务机构,“where”则以网络服务、网络设备、空间位置、地理位置为主。“when”表示网络动作相对应的时间,此 4 项组合成动作集合。

此外,要想基于深度强化学习算法优化网络安全防护效果,技术人员还应建立网络安全态势感知结构。以金融领域应用的智能网络平台为例,技术人员在安全防护设计中,可以打造集网络安全检测、智能防护、全面预警特征于一体的综合防护机制。系统还可以联合金融平台,对平台上展现的金融数据进行追踪记录,并且排查可能存在的安全隐患,便于在增强金融智能网络安全防护能力的基础上,使金融机构具备可靠且完善的安全保障。为保证智能网络动作在安全防护中体现出实时防范价值,技术人员还应充分利用大数据技术及时评估数据风险性,以便在动作元素的迁移学习中维护智能网络运行安全。

### 3.3 规范设计回报函数

基于深度强化学习的智能网络安全防护系统的优化设计,技术人员应从函数设计方面提出可行性设计决策。此次研究依托深度强化学习算法中对应的回报值,对交互场景中的参数进行持续调整,确定高回报值动作。该函数在深度强化学习理论中扮演着导向角色,系统能在函数引领下求取最大回报值,然后获取与之对应的动作元素。在智能网络运行中,智能体按照对应网络动作确定回报情况时,若某时刻下网络攻击遭受的威胁度超出阈值范围,此时对照下回报动作以奖励为主。若未至阈值范围,以惩罚回报为最终结果。若某时刻下威胁度刚好与阈值相同,则不提出具体的回报指令,从而系统在回报函数计算引导下顺利得出回报动作[5]。

根据相关研究,攻击威胁度可以划分为攻击力度、攻击行为以及攻击能力 3 种类型。技术人员可以参照不同攻击威胁度类别中的相关信息评估现有时刻下的智能网络攻击威胁度。比如在攻击行为中涵盖攻击对象重要程度、攻击位置以及攻击目标部分,系统从有方向性的判定网络威胁度。以上述回报动

作的归属范围给出适合的网络动作。在网络动作下,还可以对网络状态风险性的高低情况进行预判。经过此种系统的运行,智能网络安全性将得以增强。据此,深度强化学习算法的引进具有实践应用意义。

### 3.4 强化数据分析功能

在改进智能网络安全防护系统的环节,技术人员应借助深度强化学习算法,有效强化系统的数据分析功能,从中精准识别危险数据,剔除不安全数据,为用户提供更优质的智能网络服务[6]。为了提高智能网络安全防护等级,系统应在原有基础上对数据分析功能进行优化处理。现如今,智能网络与传统网络空间呈现增扩趋势,对应的网络安全风险也会有所升高。如果单纯按照传统算法处理网络安全问题,显然处理进度与智能网络适用性不符。而在深度强化学习算法指引下,系统可以利用既定安全防护目标,结合状态元素集合与动作状态集合,快速识别可能引发安全后果的数据信息,之后按照回报值(奖励、惩罚),对安全威胁度进行评估,提出智能修复网络系统的决策,保证经过动作与状态的双向辅助,智能网络安全防护作用将得以增强。在验证智能网络安全防护系统数据分析功能强度时,技术人员可以利用 Agent 训练程序予以评估。其中设定的循环结构以判定样本是否归属于构造数据集为基础,若显示“是”,则进入终止循环部分。若为“否”,则通过虚拟网络空间环境,对其网络安全态势进行测定,充分利用参数梯度调节方式获取对应的网络动作,如若在程序运行中样本数据存在攻击威胁度,则进行参数修正,最终保障智能网络呈现较高的安全水平。

## 4 结语

笔者提出的问题多与网络防护架构、计算机软件监测质量以及数据备份功能有关。要想达成智能网络安全防护目标,相关技术人员须充分利用深度强化学习算法,设计新系统,便于在此系统辅助下,可以全方位探寻智能网络安全风险,基于数据分析法,提高网络服务质量。

## 参考文献

[1]陈洪超.大数据背景下计算机网络安全防范策略研究[J].机械设计,2021(12):160.
[2]谭俊杰,梁应敞.面向智能通信的深度强化学习方法[J].电子科技大学学报,2020(2):169-181.
[3]卢宛芝,丁要军.基于半监督多视图特征协同训练的网络恶意流量识别方法[J].通信技术,2022(4):513-518.
[4]潘晔,刘媛.基于防火墙技术的计算机网络安全防护研究[J].网络安全技术与应用,2022(8):6-8.
[5]吉红清.信息化时代计算机网络安全防护技术[J].数字技术与应用,2022(6)234-236.
[6]刘晓影,王淮,乌吉斯古愣.基于复杂网络的多维网络安全威胁评估模型[J].通信技术,2021(8):1969-1974.

(编辑 王永超)

# Research on intelligent network security protection strategy based on deep reinforcement learning

Ren Haijuan
(Shanxi Technology and Business College, Taiyuan 030000, China)

**Abstract:** As a common algorithm of statistical learning, deep reinforcement learning will achieve remarkable results when it is applied to the design of intelligent network security protection, so as to avoid the risk of intelligent network security. This paper briefly analyzes the common problems of intelligent network security. Based on the analysis and research of the problems, it summarizes the specific objectives of intelligent network security protection optimization design. Through the implementation of design steps such as designing network state sets, refining network action sets, standardizing design return functions, and strengthening data analysis functions, intelligent networks will give full play to the role of security services, thus maintaining network security. The feasibility measures proposed in this paper, expect to improve the effectiveness of network security protection.

**Key words:** intensive learning; intelligent network; safety protection