Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2022, 11(2), 158-163
Published Online May 2022 in Hans. http://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2022.112017

Hans汉斯

面向人工智能的军队网络安全试验鉴定研究

王　军[1]，慈　颖[2]，胡　伟[3]，张　拓[3]

[1]西安交通大学社会智能与复杂数据处理重点实验室，陕西　西安
[2]北京跟踪与通信技术研究所，北京
[3]兴唐通信科技有限公司，北京

收稿日期：2022年3月7日；录用日期：2022年5月18日；发布日期：2022年5月25日

摘　要

信息网络技术作为武器装备的力量倍增器，在现代战争中地位与作用凸显，防止针对信息网络系统脆弱性的攻击，确保武器装备网络安全也变得尤为重要，试验鉴定技术是确保网络安全的重要手段，而随着大数据、人工智能等技术快速发展，基于智能化的网络安全与攻防技术对国家、军队信息网络安全提出了全新的挑战，也对试验鉴定工作提出更高要求。本文首先对“网络安全试验鉴定”概念进行阐述；其次，探讨了美军网络安全试验鉴定工作的内容与流程，最后，从人工智能为网络安全试验鉴定能力带来的重大挑战，引申出我军发展人工智能网络安全试验鉴定的几点做法和建议。

关键词

人工智能，网络安全，试验靶场，网络攻防，网络安全试验鉴定

Research on the Military Cyber Security Test and Evaluation for Artificial Intelligence

Jun Wang[1], Ying Ci[2], Wei Hu[3], Tuo Zhang[3]

[1]Key Laboratory of Social Intelligence and Complex data Processing, Xi’an Jiaotong University, Xi’an Shaanxi
[2]Tracking and Communication Technology Research Institution, Beijing
[3]Xingtang Telecommunications Technology Co. Ltd., Beijing

Received: Mar. 7th, 2022; accepted: May 18th, 2022; published: May 25th, 2022

Abstract

As a force multiplier of weapons and equipment, information network technology has a prominent position and role in modern warfare. It has become particularly important to avoid attacks on the

文章引用: 王军, 慈颖, 胡伟, 张拓. 面向人工智能的军队网络安全试验鉴定研究[J]. 人工智能与机器人研究, 2022, 11(2): 158-163. DOI: 10.12677/airr.2022.112017

vulnerability of information network systems and ensure the network security of weapons and equipment. It is an important means of network security, and with the rapid development of big data, artificial intelligence and other technologies, intelligent network security and attack and defense technologies have posed new challenges to the national and military information network security, and also put forward higher requirements for test evaluation. This paper first expounds the concept of "network security test and evaluation", and examines the content requirements of the US military's network security test and evaluation work. Secondly, based on the major challenges of network security test and evaluation capabilities by artificial intelligence, it is extended to the development of artificial intelligence networks in our military, with several practices and suggestions for safety test identification.

Keywords

Artificial Intelligence, Network Security, Cyber Range, Network Attack and Defense, Cyber-Security Test and Evaluation

1. 前言

过去几十年中，随着人类文明从工业社会步入信息社会，信息技术得到了越来越普遍的使用，尤其在武器装备的迭代中，军事战争形态的整体展现表现为由机械化向信息化不断转变的趋势。网络安全成为装备信息化的一个重要指标，网络安全涉及应用数学、密码技术、信息论、网络技术、计算机科学、通信技术等多种学科。美国政府最早启动研究与发展网络安全试验鉴定工作，不仅推动了网络安全手段与安全试验鉴定相关技术的发展，加强了联邦政府信息系统的安全管理，同时也促进了美军武器系统网络安全建设与发展，目前，网络安全试验鉴定技术已成为世界各国学术界与产业界共同关注的焦点。

2. 网络安全试验鉴定概述

网络安全试验鉴定(T&E)过程是系统工程过程(CSEP)的有机组成部分，它的作用是确定系统性能水平，帮助研制者纠正缺陷，同时也是决策过程的重要环节，为支持权衡分析、降低风险和完善需求提供数据支撑。通过网络安全风险评估、等保测评、网络攻防和渗透性测试等手段进行试验鉴定，对系统技术性能、技术规范的实现情况和系统成熟度进行评估，以确保系统的安全性、适用性和可靠性。

网络安全试验鉴定由于对象环境多变、系统复杂、对象多等因素，对试验鉴定技术提出的要求越来越高。随着人工智能、大数据新技术的不断迭代，层出不穷的网络空间新型攻击技术也花样迭出，这些新型攻击技术更加智能化、自动化、体系化，给日渐信息化、网络化的武器装备带来新的安全威胁，新型网络信息攻击技术大大增加了网络安全试验鉴定的层次与难度。未来战争模式将走向体系与体系间的对抗，以及高新技术和网络攻击技术的快速发展，直接促进了新型武器装备和先进的网络攻防技术的发展，也使试验鉴定技术向体系化、智能化方向发展。

3. 美军网络安全试验鉴定内容

近半年来，新冠疫苗普及大大减轻疫情对美军影响，美军开始围绕大国竞争缔造面向未来的网络能力，持续推进“联合全域指挥与控制”(JADC^2) [1]、“国防太空架构”等网络项目，决心打造一支拥有

高弹性全面通信网络和多域攻击能力的先进网络，以应对他国威胁，也将工作重心从反恐转向应对中国。JADC^2 宗旨是通过网络连接所有部队和装备，以共享情报、监视及侦查数据，并利用人工智能来处理这些数据。去年六月，美国副国防部长宣布启动“人工智能与数据加速”(AIDA)计划，该计划向 11 个作战司令部分别派出一支人工智能团队和一支数据团队，以帮助进行决策分析、实现数据自动化流动，以及开发用于简化决策流程的人工智能工具[2]。美军国防部也同步出台了 5000.89 指示《试验与鉴定》，旨在强力推进网络安全试验鉴定工作[3]。在试验鉴定法规方面，陆续出台发布了 15 项网络安全试验鉴定相关政策法规、指导文件和备忘录，将网络安全试验鉴定工作体系化。美军将网络安全试验鉴定内容划分为：认识网络安全需求、表征网络攻击面、协同脆弱性确认、对抗性网络安全研制试验鉴定、协同脆弱性与侵入评估和对抗性评估六个阶段[4]。如下图 1：

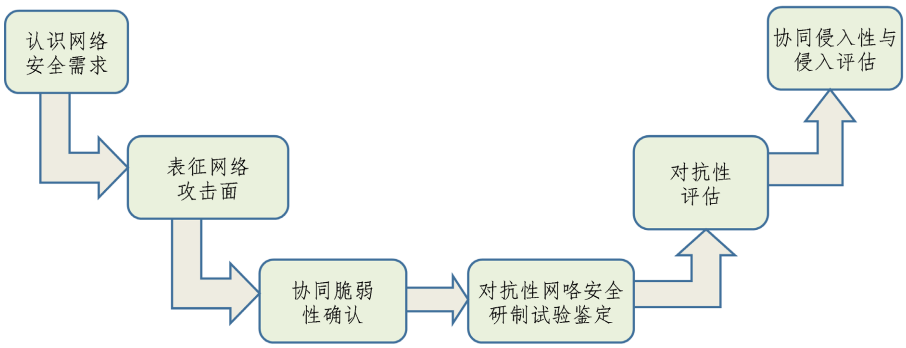

Figure 1. The process based on US military cyber-security test and evaluation
图 1. 美军网络安全试验鉴定流程

网络安全试验鉴定前四个阶段主要为研制试验鉴定提供支持。后两个阶段为武器系统的作战试验鉴定提供支持。

第一阶段：认识网络安全需求

该阶段主要了解项目或装备的网络安全需求，明确实验环境、方法、优先级等因素信息，并为开展网络安全试验鉴定制定一个初始试验方法和试验计划。具体工作分别如下：编制网络安全需求清单，确认网络威胁，在试验鉴定主计划中详细描述网络安全活动，制定初始研制鉴定框架，制定初始作战试验鉴定框架，风险管理框架成果文件与试验鉴定主计划的联系，准备研制试验鉴定分析，为征询建议书提供输入。

第二阶段：表征网络攻击面

该阶段，重点针对系统中硬件、软件、通信等相关组成单元构建攻击面，进一步确认网络攻击依赖攻击面利用可能存在的武器系统漏洞，目的是规划试验以鉴定安全漏洞是否会持续被利用。主要任务包括：确认网络攻击面，审查风险管理框架文件协助确认攻击面，分析攻击面，了解角色和职责，考虑主机环境，规划试验策略。

第三阶段：协同脆弱性确认

该阶段细化试验计划和实施脆弱性试验。这些试验和分析的完成为关键设计审查(CDR)提供支持，同时为关键设计审查提供反馈，作为试验准备审查(TRR)的输入，并为试验准备审查做准备。值得注意的是，脆弱性试验可能由一个或者多个试验事件组成。协同脆弱性确认阶段在武器系统采办全寿命周期过程中，主要是在工程与制造研发阶段实施，为系统设计人员、研发人员、工程人员提供反馈信息，从而为提高武器系统分弹性而减缓系统的脆弱性。试验、分析、修正、再试验，是一个不断重复的过程。

第四阶段：对抗性网络安全研制试验鉴定

该阶段是指利用逼真的网络空间威胁开发技术，在有代表性作战环境与任务背景下，对武器系统进行网络安全评估。利用脆弱性报告、安全评估报告和研制试验鉴定文件，由研制试验小组进行网络安全杀伤链分析，以判断若潜在网络攻击者进入被试系统他将做什么，以及被试系统将如何应对这样的攻击，所开展的对抗性评估试验，还包括模拟项目验证的能力文件中描述的网络空间威胁。

第五阶段：协同脆弱性与渗透评估

该阶段主要是制定试验计划，内容包括试验目标、完成试验所需数据、试验资源、交战规则、试验流程及测试用例等，并形成相关文档。在完成试验条件准备后，在具有代表性的作战环境中，针对系统关键功能，使用现实的威胁利用测试发现此前未知的关键脆弱性漏洞，并确定其对系统的影响，以评估任务范围内的系统作战弹性以及网络生存能力。

第六阶段：对抗性评估

该阶段针对训练有素且装备有现有系统的作战单位，设计并执行具有代表性的网络威胁行为，并根据试验结果分析网络威胁对关键作战任务产生的影响以及系统防御能力的有效性。

4. 人工智能为网络安全试验鉴定能力带来重大挑战

人工智能技术的自主化、智能化程度越高，在网络安全中得到进一步应用，基于机器学习的常见算法如决策树、支持向量机(SVM)等,不同算法可解决不同类型问题，需根据具体场景选择合适算法，如基于决策树的恶意软件检测系统，SVM 算法的新型网络入侵检测和分析系统等。自 Hinton 等人在 2006 年提出深度学习这一概念以来[5]，基于 RNN、CNN 算法的网络攻击检测技术也得到了极大提升。算法与数据是人工智能发展的核心关键，掌控的越多，供人工智能应用的资源就越多，就能取得更多突破性进展，目前，国内外在竞相推动人工智能技术发展，但是，这些算法本身存在脆弱性和不完整性，传感器欺骗、数据投毒、软件缺陷等，网络安全都会给人工智能在应用过程中造成安全威胁。

网络攻击越来越智能化、越来越隐蔽。采用人工智能技术，网络漏洞更易被挖掘，各种恶意软件更便捷使用，从而造成更严峻的安全威胁。美国国家漏洞数据库(NVD)、国家信息安全漏洞库(CNNVD)披露的漏洞越来越多，人工智能技术通过可执行文件、文档、源码等输入参数数据提取有价值信息，利用状态求解、路径发现获取利用案例，生成漏洞利用的程序，为漏洞的挖掘和利用提供了无人干预条件下自动化搜寻、自动生成、自动攻击的基本能力。另外，传统的攻击一般会在系统留下痕迹、易被追溯，容易被发现。人工智能时代，利用智能化技术，可以对攻击行为进行伪装和隐藏，如恶意代码、恶意程序可以通过内嵌神经网络模型，大幅度提高攻击行为的隐蔽性。

网络攻击对抗博弈愈演愈烈。网络安全是一个攻防博弈的过程。人工智能在处理海量数据、多元异构性数据、知识学习等方面优势巨大，攻击者会使用人工智能技术构建规模更大、结果更严重的攻击，而防御者则会利用人工智能技术提升网络的安全性，这一过程，促使网络安全攻防程度愈演愈烈。如利用算法本身脆弱性，图像识别神经网络容易被生成的和原样本高度相似的对抗样本迷惑，造成错误识别[6]。基于生成对抗网络(GAN)的 Mal GAN 算法可以使用一个替身检测器来适配黑盒恶意软件检测系统，该算法生成的恶意代码能够绕过基于机器学习的检测系统[7]。

5. 我军发展人工智能网络安全试验鉴定建议

人工智能技术既提升了网络安全，也带来新的风险和挑战，我们国家在人工智能技术上和国外存在代差，信息系统安全试验鉴定技术也存在短板，2021 年，美国参议院在其官网发布了《2021 年美国创新竞争法案》，重点提到推进“印太战略”，致力与中国的竞争，大力投资人工智能领域[8]。结合外国、

外军的先进思想和做法，弥补自身的短板与不足，需要在以下几个方面加强：

一、构建动态可扩展网络安全试验鉴定知识大脑。充分利用人工智能技术在处理海量数据、多源异构、实时动态数据方面的优势，构建试验鉴定知识大脑，提升新技术条件下的试验鉴定能力。具体而言，可分为网络安全知识库、场景知识库和试验鉴定工具知识库，针对海量文档化数据，通过知识图谱进行类别梳理，归类，分析，对过往文档进行自主学习，发现文档漏洞，同时模型本身能够存储经验。另一方面，可以采用增量学习的方法，面对新的文本数据，能够自主更新知识库。针对机器学习方法，进一步的探究其学习模型的可解释性，提升效能。网络安全知识来源广泛，包括安全事件报告、安全论坛、告警数据、病毒库、漏洞库等，为构建大规模网络安全知识大脑，需要从不同来源的网络安全数据中抽取知识，并进行有效融合，针对半结构化和非结构化数据知识，构建本体模型，结合双向循环神经网络学习方法，进行提取、标记，对未被识别的本体进行人工抽取，从而确保本体模型基于三元组的知识逻辑正确性，实现动态可扩展，为试验鉴定知识大脑知识库提供强大的自学习能力。

二、建立动态网络安全试验鉴定评估指标体系。通过对海量模糊、非线性、异构数据进行自动化的分类聚合与关联分析，全面感知网络安全威胁，自主学习认知网络空间态势；主动生成与快速调整网络威胁防御策略，在与自主化、智能化网络空间攻击手段的攻防博弈中不断学习演进，逐渐形成适应性强、反应迅速灵敏的网络空间安全防御“智慧”。参考卷积神经网络的分层结构，通过对数据进行归一化预处理，将原始的网络安全评估基础指标数据映射成卷积神经网络的输入数据向量，作为人工神经网络的输入数据，然后，基于经典神经网络的深度卷积神经网络，设计适于网络安全评价指标体系构建应用的人工神经网络参数，构造能够实现网络安全评级的最优人工神经网络分类模型，从而实现自动化的即时网络安全测评等级分类，同时达到较高精度的有效分类的结果。整体可以分为四个部分:网络安全试验鉴定数据预采集、数据预处理、深度学习核心训练、最终试验鉴定指标体系输出。实验和调整以及算法核心的关键则放在了深度学习的深度以及具体实现性技术的优化上。在训练时因实验数据较少，用同样的数据组进行重复训练，运用大量数据进行批量单次训练，因为卷积神经网络的训练波动较大，当数据集训练完成后或者连续数次反向传播后精确度依旧较高，最终，建立具备弹性、自学习力的试验鉴定评估指标体系，如下图 2 所示：

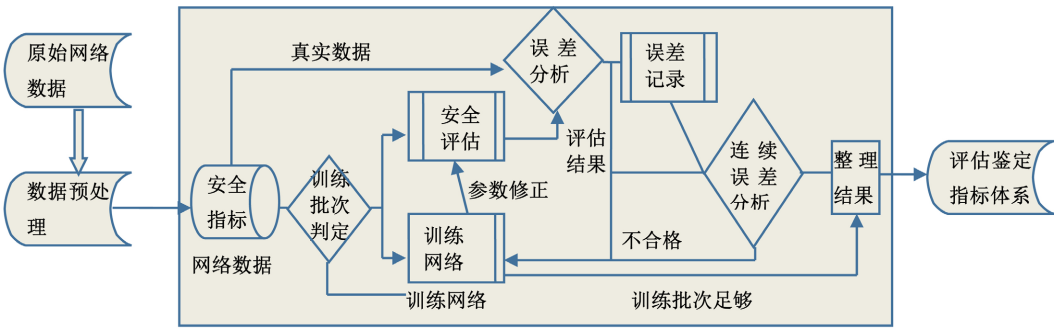

Figure 2. The process based on convolutional neural network test and evaluation
图 2. 基于卷积神经网络试验鉴定流程

三、重视试验鉴定人才队伍建设。伴随先进技术的大量引入以及装备新系统的不断开发采用，试验鉴定复杂度成倍增加，对人员的要求也更高，必须训练有素的专业技术人员，根据网络安全态势采取相应的安全措施，需要既懂武器装备又精通网络攻防对抗的专业人才，进行高技术条件下的网络安全试验和鉴定工作。为加强人员队伍建设，需要组建先进网络对抗小组，成立专门网络评估组织，通过专业技

术培训，同时取得相应资质，从而应对新技术条件下的试验鉴定能力和更好完成数量庞大的网络安全试验鉴定任务。

6. 总结

本文通过对网络安全试验鉴定知识的梳理，分析了人工智能技术在网络安全试验鉴定中存在的挑战和机遇，提出构建动态可扩展网络安全试验鉴定知识大脑，通过深度学习卷积神经网络算法，建立动态网络安全试验鉴定评估指标体系，为新形势下军队装备网络安全试验鉴定工作提供参考。

参考文献

[1] 陈彩辉, 线珊珊. 美军“联合全域作战(JADO)”概念浅析[J]. 中国电子科学院学报, 2020, 15(10): 917-921.

[2] 裴晔晔, 倪得晶, 陶智刚, 赵宇. 美军国防信息基础设施体系建设发展与启示[J]. 指挥信息系统与技术, 2012, 12(6): 7-13.

[3] 刘映国, 薛卫, 谢伟朋, 郑超. 美军国防采办政策改革及启示[J]. 国防科技, 2021, 42(6): 64-68+76.

[4] 刘映国. 美军网络安全试验鉴定[M]. 北京: 国防工业出版社, 2018.

[5] Lecun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. https://doi.org/10.1038/nature14539

[6] Guo, Z.Q., Hu, W.X., Zhang, C.J., *et al.* (2020) Gradient Shileding: Towards Understanding Vulnerability of Deep Neural Network. *IEEE Transactions on Network Science and Engineering*, **8**, 921-932.

[7] Hu, W.W. and Tan, Y. (2021) Generating Adversarial Malware Examples for Black-Box Attacks Based on Gan. https://arxiv.org/pdf/1702.05983.pdf

[8] 管传靖. 安全化操作与美国全球供应链政策的战略性调适[J]. 国际安全研究, 2022, 40(1): 73-99.