 DOI: https://doi.org/10.12349/iser.v6i10.7913

Research on Automated Detection Technology of Computer Network Security Vulnerabilities Empowered by Artificial Intelligence

Yuyang Xu

Northwest Normal University, Lanzhou, Gansu, 730000, China

Abstract

Network security vulnerabilities, as important entry points for cyber attacks, pose serious security threats and economic losses to individuals, enterprises, and even countries. Traditional vulnerability detection methods have gradually exposed problems such as low efficiency, low accuracy, and insufficient detection capabilities for new types of vulnerabilities when facing increasingly complex and massive network data. The rise of artificial intelligence technology provides a new solution for the automated detection of computer network security vulnerabilities. This paper deeply studies the automated vulnerability detection technology empowered by artificial intelligence, elaborates on the relevant technical principles, methods, and advantages in detail, and looks forward to the future development trends, aiming to provide strong support for improving the level of computer network security protection.

Keywords

Artificial Intelligence; Computer Network Security; Automated Vulnerability Detection; Machine Learning; Deep Learning

人工智能赋能的计算机网络安全漏洞自动化检测技术研究

徐煜洋

西北师范大学，中国 · 甘肃 兰州 730000

摘　要

网络安全漏洞作为网络攻击的重要切入点，给个人、企业乃至国家带来了严重的安全威胁和经济损失。传统的漏洞检测方法在面对日益复杂和海量的网络数据时，逐渐暴露出效率低、准确率不高以及对新型漏洞检测能力不足等问题。人工智能技术的兴起为计算机网络安全漏洞自动化检测提供了新的解决方案。本文深入研究人工智能赋能下的漏洞自动化检测技术，详细阐述相关技术原理、方法及优势，并对未来发展趋势进行展望，旨在为提升计算机网络安全防护水平提供有力支持。

关键词

人工智能；计算机网络安全；漏洞自动化检测；机器学习；深度学习

1 引言

依据行业报告，2024 年全球范围内因漏洞致使的网络攻击事件较上一年度增幅达 32%，所引发直接经济损失超越 5000 亿美元。传统漏洞检测方式，因其对人工规则的倚赖以及响应滞后等状况，无法面对当下海量数据与新型漏洞的挑战。人工智能凭借自身强大的数据处理与自主学习能力，在此种情形下，为网络安全漏洞自动化检测开拓出全新路径。对该技术展开深入探究，不但能够弥补传统检测手段的不足，且对于稳固网络安全防线、确保数字经济健康发展而言，具备重要的战略价值。

【作者简介】徐煜洋（2004–），中国浙江台州人，本科，从事软件工程研究。

2 计算机网络安全漏洞概述

2.1 漏洞类型

2.1.1 软件漏洞

软件开发期间，鉴于程序员疏失、逻辑设计瑕疵或安全考量欠妥等缘由，会产生各类漏洞 [1]。像缓冲区溢出漏洞，程序向缓冲区写入数据时，若超越缓冲区容量，或许致使数据覆写相邻内存区域，攻击者可借此漏洞更改程序执行流程，植入恶意代码。Web 应用程序里的 SQL 注入漏洞，倘若对用户输入数据未予严格筛选，攻击者能于输入字段插入恶意 SQL 语句，非法获取、窜改或删除数据库数据。

2.1.2 网络协议漏洞

在网络协议的设计及实现进程中同样可能出现漏洞。TCP/IP 协议栈内便存有些许漏洞，以 SYNFlood 攻击为例，此攻击借由 TCP 三次握手流程里的漏洞，借发送诸多伪造的 SYN 请求包，将服务器资源耗尽，致使正常用户无法构

建连接，触发分布式拒绝服务（DDoS）攻击。

2.1.3 系统配置漏洞

安全隐患或因系统管理员配置服务器、网络设备等时设置欠妥而遗留。诸如，弱密码策略易致用户密码遭破解；开放不必要端口与服务，为攻击者提供入侵路径；权限设置不合理，或致用户权限过高或过低，对系统安全及正常运行造成影响。

2.1.4 人为因素漏洞

网络安全漏洞的产生，或源于人为因素[2]。员工安全意识匮乏，对不明链接随意点击，对未知文件贸然下载，恶意软件便易乘虚而入。内部人员若故意泄露敏感信息，或实施违规操作，网络安全亦会遭受严重威胁。

2.2 漏洞危害

2.2.1 数据泄露

攻击者凭借漏洞可将其获取系统内敏感数据，像用户个人信息（姓名、身份证号、银行卡号等）、企业商业机密、政府机密文件等[3]，导致个人隐私遭侵犯、财产受损失。甚至导致企业面临法律风险、声誉受损，影响国家安全稳定。

2.2.2 系统瘫痪

攻击通常通过漏洞发起，比如 DDoS 攻击，能够致使服务器或网络设备因资源的耗尽，而无法正常运作，进而引发系统瘫痪的状况。企业的业务运营，会因这一情形遭受严重影响，产生巨大经济损失[4]。至于一些关键基础设施，像电力、交通、医疗等领域的网络系统，系统一旦瘫痪，公众的生命财产安全或会受到危及。

2.2.3 恶意控制

系统控制权在攻击者借助漏洞获取后，恶意软件，像木马以及病毒等，便有于系统内被植入之可能。受害者设备的进一步掌控随之而来，更多恶意操作亦接踵而至。例如发起后续网络攻击，窃取更多信息，或者以僵尸网络一部分的身份参与大规模攻击活动。

2.2.4 经济损失

经济损失是网络安全漏洞危害的最终结果[5]。该范畴既包括直接层面的因素，如数据恢复成本、系统修复成本以及因业务中断而导致的收入损失，也涵盖间接层面的因素，例如企业声誉受损所引发的客户流失、市场份额下降等情况。

3 人工智能技术在漏洞自动化检测中的应用

3.1 机器学习在漏洞检测中的应用

3.1.1 分类模型用于漏洞识别

在漏洞检测领域，分类模型于机器学习范畴内具备显著意义。朴素贝叶斯（NB）、支持向量机（SVM）、决策树 / 随机森林（DT/RF）以及神经网络等常见分类模型，借由针对大量已知漏洞及正常数据样本展开研习，实现构建分类器的目的，进而判别新数据有无漏洞。以朴素贝叶斯分类器训练为例，首先，针对网络流量数据、系统日志数据或者代码数据等实施特征提取操作，将所提取特征当作输入，把对应的漏洞标签作为输出，借助训练促使模型获悉不同特征与漏洞间的概率联系。当新数据输入之时，模型依据习得的概率分布对该数据属于漏洞数据的可能性予以判断。支持向量机旨在寻觅一个最优分类超平面，借由分隔不同类别数据样本实现漏洞识别。决策树 / 随机森林通过构建树形结构，基于数据特征开展决策从而实现数据分类，随机森林经由集成多个决策树，进一步提升分类的精准性与稳定性。神经网络具备强大的非线性拟合能力，通过搭建多层神经元结构，能够自主学习数据中的复杂特征与模式，在漏洞检测方面呈现出较高准确率。于实际应用场景中，多种分类模型常被结合运用，先是借助简单快速的模型实施初步筛选，继而运用复杂且准确率高的模型进行精细分类，以此在检测效率与准确率之间寻得平衡。

3.1.2 回归模型预测漏洞相关指标

在漏洞检测领域，回归模型具备预测与漏洞相关指标之功用，如漏洞数量以及漏洞严重程度评分等。常用的回归模型涵盖线性回归与逻辑回归。借助建立因变量（例如漏洞数量）和自变量（诸如代码行数、函数调用次数这般的代码特征）间的线性关系，线性回归实现对漏洞数量的预测。逻辑回归适宜预测具备二分类属性的漏洞相关指标，比如漏洞是否易于被利用（是与否），借由对数据施行逻辑变换，搭建预测模型。深度学习中的部分模型，例如多层感知机（MLP）同样可用于回归任务，借由构建复杂的神经网络结构，其能更优地捕捉数据里的非线性关系，进而更精准地预测漏洞相关指标。例如，分析软件项目的历史数据，其中包含代码复杂度、开发周期、以往发现的漏洞状况等，凭借回归模型预测当前版本软件中或许存在的漏洞数量或某个特定漏洞的严重程度，为安全人员制定漏洞修复策略供应参考依据。

3.1.3 聚类分析发现潜在漏洞模式

聚类分析作为一种将数据对象予以分组，形成相似对象集合（簇）的进程。于漏洞检测范畴内，该分析方法能够用于潜在漏洞模式的识别。对网络数据抑或代码数据开展特征提取操作之后，借助诸如 K - Means 算法等聚类算法，把具备相似特征的数据汇聚为一组。若在某一簇当中察觉已知的漏洞样本，那么该簇里其余未知样本亦可能存在类似漏洞情况，为安全人员供给潜在漏洞线索。例如，针对大量系统日志实施聚类分析，把具备相似行为模式的日志归为一类，倘若其中一类日志涵盖已知的由漏洞引发的异常行为记录，那么该类里其他日志所对应的系统行为或许存在漏洞风险，安全人员能够针对这些潜在漏洞展开更深入分析，预先做好防范工作。

3.2 深度学习在漏洞检测中的应用

3.2.1 卷积神经网络（CNN）用于代码漏洞检测

在图像识别领域，卷积神经网络收获斐然成就。近年来，其亦广泛涉足代码漏洞检测领域。代码可视作具备特定结构与语义信息的特殊文本序列。将代码转译为适配 CNN

处理之格式，诸如对代码语法结构、语义信息等予以编码，使之呈现为图像或张量形式。借由卷积层、池化层与全连接层等架构，CNN 可自动萃取出代码内的局部与全局特征。卷积层里的卷积核于代码数据上滑移，提取不同位点的特征；池化层则用于特征降维，削减计算量且留存关键特征。全连接层对所提取特征加以分类，判定代码有无漏洞。以检测 C 语言代码的缓冲区溢出漏洞为例，CNN 能够习得代码中与缓冲区操作相关的特征模式，像数组声明、指针运算之类特征，借由对海量涵盖缓冲区溢出漏洞及正常代码样本的研习，可精准甄别新代码是否存在此类漏洞，与传统方式相比，显著提升检测的准确率与效率。

3.2.2 循环神经网络（RNN）及变体处理序列数据中的漏洞

具备序列特性的数据处理，循环神经网络（Recurrent Neural Network, RNN）及其变体，如长短期记忆网络（Long Short-Term Memory, LSTM）和门控循环单元（Gated Recurrent Unit, GRU），具有较高的适用性。网络流量数据、系统日志数据以及代码内函数调用序列等，皆具序列特征。循环神经网络（RNN）可处理序列数据中的时间依赖关系，借由隐藏层状态的循环传递，将之前输入信息予以记住。于漏洞检测领域，RNN 能够习得网络流量或系统日志里的正常行为模式，一旦出现偏离正常模式的异常序列，便可能检测出存在的漏洞或攻击行为。而长短期记忆网络（LSTM）与门控循环单元（GRU），解决了 RNN 处理长序列时易出现的梯度消失及梯度爆炸问题，对长距离依赖关系的捕捉更为出色。以分析系统日志的入侵检测场景为例，LSTM 可针对一段时间内的系统操作日志序列展开学习，精准识别攻击者借漏洞实施的一系列操作步骤，即便这些操作步骤在时间上存在一定跨度，亦能凭借其记忆机制准确检测，在复杂攻击场景下对漏洞利用行为的检测能力得以有效提升。

3.2.3 生成对抗网络（GAN）辅助漏洞检测

生成对抗网络，由生成器以及判别器构成。二者处于相互对抗且相互学习的状态。于漏洞检测场景下，生成器具备生成看似正常却可能隐匿潜在漏洞数据样本之能力。判别器的职责，是区分生成的数据、真实正常数据以及已知漏洞数据。借由这般对抗训练模式，判别器针对各类数据的区分能力得以持续提升，进而能够更有效地检测出真实数据里的漏洞。以网络流量检测为例，生成器生成模拟网络流量数据，这些数据或许涵盖攻击者利用漏洞实施攻击时的部分特征，然而却伪装成正常流量之形态。在与生成器的对抗进程中，判别器逐步洞悉真实正常流量、已知攻击流量以及生成的伪装攻击流量之间的差异，提高对真实网络流量中潜在漏洞利用行为的检测准确率。此外，生成对抗网络还可用于扩充漏洞检测数据集，借由生成更多不同类型的漏洞数据样本，缓解实际检测时数据集不足的状况，提升模型的泛化能力。

3.3 自然语言处理技术在漏洞检测中的应用

3.3.1 代码语义分析识别漏洞

对于代码漏洞检测时，可运用自然语言处理技术内的语义分析手段。代码虽为形式化语言类型，却也蕴含特定语义信息，和自然语言存在相似特征。借助词法、语法以及语义分析于代码之上，代码意图与结构得以理解，潜在安全风险亦能识别。像剖析代码里函数调用关系、变量作用域、控制流及数据流等语义信息，以此判定代码是否依从安全编程规范。若代码中出现对敏感函数的不恰当调用情形，比如在未实施充分输入验证状态下调用可能引发缓冲区溢出的函数，语义分析可察觉此类潜在漏洞。自然语言处理中的语义相似度计算方式，可将待检测代码同已知安全代码模式或者漏洞代码模式开展语义匹配，对代码是否存在漏洞风险加以判断，使漏洞检测的准确性与全面性得以提升。

3.3.2 漏洞报告文本挖掘辅助检测

在漏洞检测过程中，存在大量的漏洞报告文本，这些文本包含了关于漏洞的发现、特征、影响以及修复建议等丰富信息。自然语言处理范畴内的文本挖掘技术，如信息抽取、文本分类、主题模型等，可从这些漏洞报告文本中提取有价值信息以辅助漏洞检测。借助信息抽取技术，从漏洞报告提取漏洞类型、位置、触发条件等关键信息，实现漏洞知识库的构建。文本分类技术能够对新漏洞报告予以分类，判定其所属漏洞类别，利于安全人员迅速明晰漏洞性质。主题模型可发觉漏洞报告中的潜在主题，挖掘不同漏洞间的关联关系。例如，针对大量 Web 应用漏洞报告施行主题模型分析，发现某些类型漏洞（诸如跨站脚本攻击与 SQL 注入漏洞）于特定 Web 开发框架或环境中更易出现，为在该框架或环境下开发的应用程序开展针对性漏洞检测提供参照，提升检测效率与效果。

4 结语

计算机网络安全漏洞自动化检测领域，因人工智能技术之融入，获致突破性进展。其中，机器学习借分类、回归以及聚类模型，实现对漏洞精准识别与指标预测；深度学习凭借 CNN、RNN 连同 GAN 等模型，成功攻克序列数据处理与复杂漏洞挖掘难题；自然语言处理技术，对代码语义与漏洞报告内价值信息加以深度挖掘。此三者相互配合，使得检测效率与准确率显著提升。不过，当下技术所面临之挑战亦不容忽视，诸如模型泛化能力欠佳、对抗性攻击防御薄弱等问题，仍亟待取得突破。

参考文献

[1] 邱峻. 计算机网络安全漏洞及防范策略探析 [J]. 数字技术与应用, 2025, 43 (06): 71-73.

[2] 白子文. 计算机网络安全的漏洞检测与对策分析 [J]. 电子技术, 2025, 54 (01): 286-287.

[3] 赵刘葳. 计算机网络安全漏洞检测与对策分析 [J]. 电子技术, 2025, 54 (01): 290-291.

[4] 宁晓斐,马俏. 计算机网络安全漏洞及防范措施解析 [J]. 通信管理与技术, 2024, (06): 46-47.

[5] 齐德林. 云计算环境中计算机网络安全威胁与漏洞分析研究 [J]. 电脑知识与技术, 2024, 20 (19): 94-96. DOI:10.14004/j.cnki.ckt.2024.1038.