融合深度特征与强化学习的工控协议模糊测试方法*

宗学军[1,2]，孙俊辉[1,2]，何　戡[1,2]，史洪岩[1,2]，连　莲[1,2]，宁博伟[2,3]
（1. 沈阳化工大学　信息工程学院，辽宁　沈阳　110142；2. 辽宁省石油化工行业信息安全重点实验室，辽宁　沈阳　110142；
3. 沈阳工业大学　人工智能学院，辽宁　沈阳　110870）

摘　要：针对工业控制协议漏洞挖掘存在协议语义理解不足、变异策略单一的问题，提出一种融合深度特征与强化学习的工控协议模糊测试方法——CTARFuzz。该方法通过 CTCA-Net 模型，提取协议结构与上下文特征，并引入注意力机制强化关键字段，提升测试用例多样性与接收率。结合 Actor-Critic 强化学习模型，以 CTCA-Net 模型的输出特征驱动 Actor 网络选变异策略生成用例，Critic 网络依据设备反馈动态优化策略，实现变异策略的自适应优化。实验在典型能源企业工业场景的攻防演练靶场上采用 Modbus TCP、EtherNet/IP 和 S7Comm 协议进行验证，结果表明 CTARFuzz 异常触发率优于其他方法，并拥有较高的接收率与多样性，在靶场多个设备中触发异常，验证了 CTARFuzz 的适用性与有效性。

关键词：模糊测试；工业控制协议；卷积神经网络；强化学习；时序卷积网络
中图分类号：TP393　　**文献标志码：**A　　**DOI：**10.19358/j.issn.2097－1788.2026.02.001

中文引用格式：宗学军，孙俊辉，何戡，等. 融合深度特征与强化学习的工控协议模糊测试方法［J］. 网络安全与数据治理，2026，45(2)：1－11.
英文引用格式：Zong Xuejun，Sun Junhui，He Kan，et al. Fuzzing test method for industrial control protocol based on deep feature and reinforcement learning［J］. Cyber Security and Data Governance，2026，45(2)：1－11.

Fuzzing test method for industrial control protocol based on deep feature and reinforcement learning

Zong Xuejun[1,2], Sun Junhui[1,2], He Kan[1,2], Shi Hongyan[1,2], Lian Lian[1,2], Ning Bowei[2,3]
(1. College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China;
2. Liaoning Key Laboratory of Information Security for Petrochemical Industry, Shenyang 110142, China;
3. College of Artificial Intelligence, Shenyang University of Technology, Shenyang 110870, China)

Abstract: The vulnerability mining of industrial control protocol mainly has the problems of insufficient protocol semantic understanding and single mutation strategy. A fuzzing test method of industrial control protocol-CTARFuzz based on deep feature and reinforcement learning is proposed. This method extracts protocol structure and context features through CTCA-Net model, and introduces attention mechanism to strengthen key fields, so as to improve test case diversity and acceptance rate. Combined with the Actor-Critic reinforcement learning model, the output characteristics of the CTCA-Net model are used to drive the Actor network to select the mutation strategy to generate use cases. The Critic network realizes the adaptive optimization of the mutation strategy according to the dynamic optimization strategy of the device feedback. The experiment is verified by Modbus TCP, EtherNet/IP and S7Comm protocols on the attack and defense drill range of typical energy enterprise industrial scenes. The results show that the abnormal triggering rate of CTARFuzz is better than other methods, and has a high acceptance rate and diversity. It triggers abnormalities in multiple devices in the range, which verifies the applicability and effectiveness of CTARFuzz.
Key words: fuzzing; industrial control protocol; convolutional neural network; reinforcement learning; temporal convolutional network

*基金项目：辽宁省科技重大专项项目（辽科办发［2025］77 号（3）-1）；辽宁省应用基础研究计划项目（2025JH2/101300012）；辽宁省科技重大专项项目（2024JH1/11700049）；辽宁省自然科学基金项目（2023-MSLH-273）

0 引言

工业控制系统（Industrial Control Systems，ICS）在现代工业自动化中至关重要，广泛应用于制造业、电力系统等领域[1]。ICS 通常由可编程逻辑控制器（PLC）、分布式控制系统（DCS）、远程终端单元（RTU）等[2]工控设备组成，设备之间通过工业控制协议（Industrial Control Protocol，ICP）进行通信和控制。随着工业互联网的发展，ICS 逐步向开放网络架构转型[3]，虽提升了系统的互联互通能力，却面临网络攻击威胁。例如，2025 年 5 月巴基斯坦对印度发动大规模网络攻击，导致印度国家电网工业控制系统受到攻击，使印度约 70% 的电网瘫痪[4]。

许多 ICP 设计之初并未充分考虑网络安全问题，其固有的脆弱性使得漏洞挖掘成为研究的重点[5]。模糊测试可通过变异协议报文并观察设备响应发现未知漏洞[6]，然而，传统的模糊测试在应用于 ICP 时面临着多样性不足和接收率低等问题[7]。

近年来，深度学习[8]和强化学习[9]在漏洞挖掘领域展现出强大潜力。Cheng 等[10]提出 MSFuzz，利用大型语言模型（Large Language Models，LLM）理解协议语法结构，生成符合协议规范的测试用例，但模型训练依赖有限的协议样本。Yang 等[11]提出 WG-GFuzz，利用生成对抗网络（Generative Adversarial Network，GAN）生成测试用例，但过度依赖特定协议的格式和状态特征且普适性不足。Che 等[12]提出了一种基于信息理论的模糊测试方法，通过协议结构解析算法和基于遗传算法生成测试用例，但对训练数据的质量和数量有一定依赖。Wanyan 等[13]提出了基于协议特征的变异方法，利用非关键字段的变异与测试用例组合技术，减少了冗余输入的生成，但接收率不足。

当前，针对工控协议的特征提取存在一些不足，单一的深度学习模型不能准确提取特征。卷积神经网络（Convolutional Neural Networks，CNN）[14]虽能捕捉协议字段局部组合模式，但无法建模长距离时序依赖。时序卷积网络（Temporal Convolutional Network，TCN）[15]可通过因果卷积与膨胀卷积覆盖长时序，但对关键语义字段关注度不足。因此，本文通过 CTCA-Net 模型提取特征。不同于单一模型的局限性，CTCA-Net 采用融合设计思路，CNN 捕捉协议报文的局部结构特征，TCN 建立字段间的长时序依赖关系，再通过注意力机制对关键语义字段进行强调，最终实现特征提取性能的提升。

综上，本文提出了一种融合深度特征与强化学习的工控协议模糊测试方法。本文主要贡献概括如下：

（1）提出 CTCA-Net 模型提取协议深层特征，解决传统方法对协议语义理解不足的问题，提升测试用例接收率与多样性。

（2）设计 Actor-Critic 强化学习框架，实现变异策略自主优化，解决传统变异策略单一问题，提升测试效率。

（3）采用 Modbus TCP、EtherNet/IP 和 S7Comm 协议评估 CTARFuzz 性能，与现有模糊测试方法相比，CTARFuzz 拥有较高的异常触发率，验证了其在不同协议与设备中的适配性及实用性。

1 相关工作

模糊测试[16]技术由 Miller 教授于 1988 年首次提出。其早期研究通过开发的模糊测试工具，将生成的随机数据输入 UNIX 实用程序[17]，结果发现有超过 25% 的程序会因处理随机输入而崩溃。

在 ICS 安全测试中，模糊测试可以检测工业协议中存在的潜在漏洞[18]。模糊测试流程如图 1 所示。

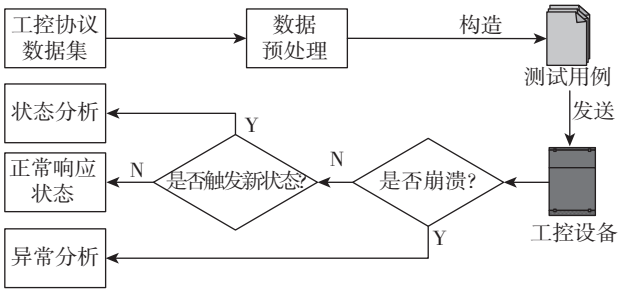

图 1 模糊测试流程

传统方法通常依赖人工设计的报文变异规则，变异的数据难以全面覆盖不同类型的潜在漏洞，存在诸多局限性[19-20]。因此，本文通过 CTCA-Net 模型对协议报文进行深层语义建模，提取关键字段及其时序依赖特征，结合 Actor-Critic[21]算法，在测试过程中根据环境反馈优化变异策略，提升模糊测试的效率与漏洞触发能力。

为明确 CTARFuzz 的创新性与技术优势，本文从特征提取方式、强化学习框架设计、变异策略选择逻辑及核心优化方法几个维度，与同类强化学习模糊测试方法 CTFuzz 进行系统性对比，具体如表 1 所示。

表 1　CTARFuzz 与 CTFuzz 技术对比

技术对比	CTARFuzz	CTFuzz
特征提取方式	使用 CTCA-Net 模型，提取协议语义特征与时序依赖特征、关键字段	字节数组直接映射，仅保留字节级静态结构，无语义挖掘
强化学习框架设计	使用 Actor-Critic 网络，策略生成与价值评估结合，多维度反馈	单网络 DQN，离散状态－动作映射，单一覆盖率相关奖励
变异策略选择逻辑	上下文感知，关联字段协同变异，关键字段优先	字节级通用变异，限制单一动作执行次数，无语义适配
核心优化方法	框架内动态价值评估，自适应调整	外部种子调度算法，模型未直接优化

2　融合深度特征与强化学习的工控协议模糊测试

2.1　整体概述

针对现有工业控制协议模糊测试方法在测试用例生成过程中存在的语义理解不足、变异策略单一等问题，本文提出一种融合深度特征与强化学习的工控协议模糊测试方法——CTARFuzz。该方法先利用 CTCA-Net 模型中的 CNN 捕捉报文局部模式，TCN 建模字段间长距离时序依赖，注意力机制[22]进一步突出关键语义字段，构建具备上下文感知能力的协议状态表示。

在此基础上，构建 Actor-Critic 强化学习框架。Actor 网络根据状态特征，从变异算子库中选择变异策略构建测试用例并发送至被测设备，Critic 网络依据设备响应评估测试效果，通过时序差分误差引导 Actor 网络策略的更新，实现变异策略的自适应优化。此外，引入多智能体[23]，增强框架通用性与跨协议适配能力，不同协议由独立的智能体进行训练与优化，提升整体框架的可扩展性。

整体架构主要包含数据预处理、特征提取、测试用例生成及策略优化等过程。如图 2 所示，从工控设备中获取协议报文进行预处理，经 CTCA-Net 提取特征。Actor 网络依据状态特征选择变异策略构造测试用例并发送给设备，Critic 网络依据设备的响应评估当前策略的效果，通过时序差分误差优化 Actor 网络参数。

2.2　数据预处理

为了适配模型的输入要求，需要对捕获的数据进行预处理，如图 3 所示。首先，过滤掉非目标协议的通信内容，然后筛选、剔除重复和无效的请求报文。利用 k-means[24]算法对收集的报文进行聚类，根据报文特征的相似性将其划分为多个簇，每个簇代表一种典型的通信行为或协议操作模式。之后在报文开头添加“SRT”作为开始标志，结束位置添加“END”作为结束标志。此外，短数据报文使用零填充，以实现统一的长度。最后，将协议数据统一为十六进制字符序列的形式，每个字符的取值范围为（0，1，2，3，4，5，6，7，8，9，a，b，c，d，e，f）。每个字符可以用一个长度为 16 的独热编码向量[25]来表示。

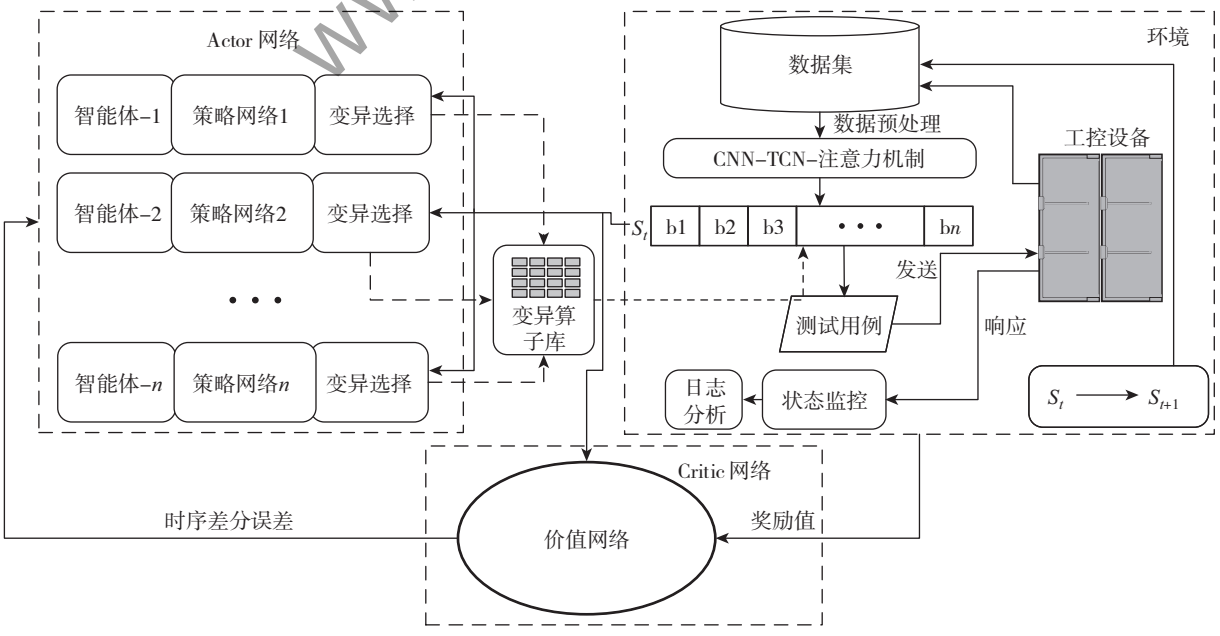

图 2　CTARFuzz 整体架构图

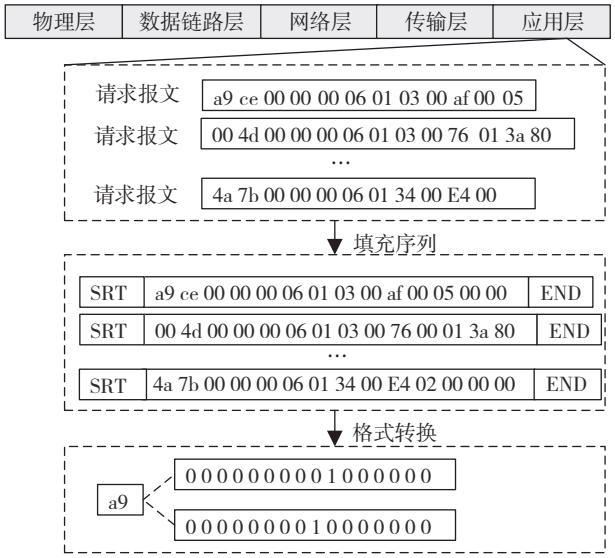

图 3　数据预处理

2.3　基于 CTCA-Net 模型的工控协议特征提取

ICP 具有特定的结构和语义信息，传统方法在面对复杂结构与语义的报文时，难以有效挖掘深层次特征。为此，本文利用 CTCA-Net 模型提取协议报文的空间结构特征与时序语义关系，如图 4 所示。

首先，CNN 提取局部字段组合的语义特征。通过多个一维卷积操作，利用大小为 k 的卷积核在序列上滑动，对相邻字段间的组合模式进行建模，提取出组合字段的潜在模式，生成局部特征表示序列。第 i 个卷积核在位置 j 的响应计算如式（1）所示：

$$c_j^{(i)} = \mathrm{ReLU}\left(< W^{(i)},\ X_{j:j+k-1} > + b \right) \tag{1}$$

其中，$W^{(i)} \in \mathbf{R}^{k \times d}$ 为第 i 个卷积核的权重，$X_{j:j+k-1}$ 为窗口内的输入切片，b 为偏置项，ReLU 为非线性激活函数。

获得初步的局部表示后，TCN 采用因果卷积与膨胀卷积进一步建模协议字段之间的时序依赖关系，使得当前时刻的表示仅依赖于其前文上下文，同时扩大感受视野以覆盖更长范围的上下文信息。其卷积过程如式（2）所示：

$$y_t = \sum_{i=0}^{k-1} W_i \cdot x_{t-d \cdot i} \tag{2}$$

其中，d 为膨胀率，k 为卷积核大小，$x_{t-d \cdot i}$ 为卷积权重。为保证深层结构的训练稳定性，引入了残差连接，如式（3）所示：

$$Y = f(X) + X \tag{3}$$

这一阶段可以有效建模协议报文中跨字段的上下文依赖关系，提取时间序列特征。

为了进一步增强模型对语义关键字段的感知能力，本文引入注意力机制对 TCN 输出的各时间步特征进行加权聚合。首先对每一时刻 t 的特征表示 $\boldsymbol{h}_t$ 计算注意力得分 e_t，如式（4）所示：

$$e_t = \boldsymbol{v}^{\mathrm{T}} \tanh\left(W_h \boldsymbol{h}_t + b_h \right) \tag{4}$$

其中，$\boldsymbol{v}$，W_h，b_h 为学习参数。

然后通过 Softmax 归一化为注意力权重，再加权求和得到最终的特征向量 $\boldsymbol{s}$。

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^{T} \exp(e_i)} \tag{5}$$

$$\boldsymbol{s} = \sum_{t=1}^{T} \alpha_t \boldsymbol{h}_t \tag{6}$$

其中，α_t 为第 t 个位置的注意力权重。

2.4　Actor-Critic 策略优化与奖励函数设计

为了实现测试用例生成过程中的变异策略自适应，本文引入强化学习的 Actor-Critic 架构并构造相应的奖励函数。

具体而言，Actor-Critic 网络将 CTCA-Net 模型提取的协议状态向量 $\boldsymbol{s}$ 作为当前环境状态的输入，Actor 网络根据该状态向量输出一组变异算子的概率分布，然后从动作空间中选择具体的变异算子 a 构造测试用例。Actor 网络的策略函数如式（7）所示：

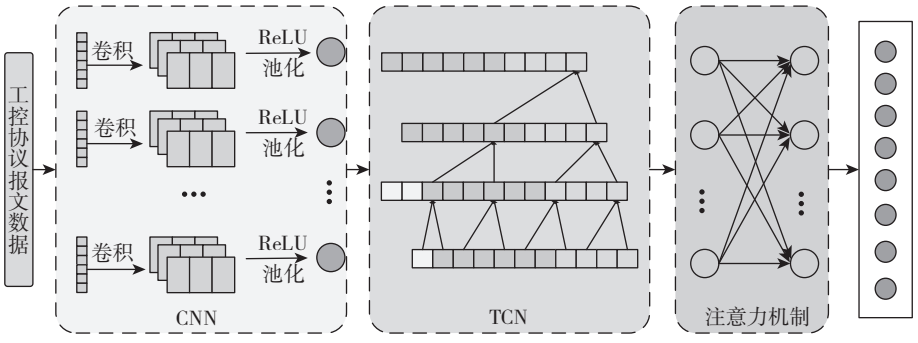

图 4　工控协议特征提取

$$\pi(a_i \mid \boldsymbol{s}) = \frac{e^{(W_i \boldsymbol{s}+b)}}{\sum_{j=1}^{n} e^{z_j}} \tag{7}$$

其中，$\pi(a_i \mid \boldsymbol{s})$ 表示在状态 $\boldsymbol{s}$ 下选择第 i 个变异动作 a_i 的概率，$W_i \boldsymbol{s}$ 表示状态 $\boldsymbol{s}$ 在动作 i 方向上的评分，b 为偏置项。

构造测试用例以后发送至设备，Critic 网络根据设备响应信息，评估当前状态下策略的价值，并根据奖励函数计算相应的状态值函数 $V(\boldsymbol{s})$，状态值函数$V(\boldsymbol{s})$如式（8）所示：

$$V(\boldsymbol{s}) = E_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t r_t \mid \boldsymbol{s}_0 = \boldsymbol{s}\right] \tag{8}$$

其中，π 表示当前策略，r_t 表示 t 时刻的即时奖励，γ 为折扣因子。

奖励函数综合考虑测试用例的接收率（TCA）、多样性（DTC）与异常触发能力（ATR）三个性能指标，如式（9）所示：

$$R = \lambda_1 R_{\mathrm{t}} + \lambda_2 R_{\mathrm{d}} + \lambda_3 R_{\mathrm{a}} \tag{9}$$

其中，R_{t} 表示接收率的奖励值，R_{d} 表示异常率的奖励值，R_{a} 表示测试用例多样性的奖励值，λ_1，λ_2，λ_3 表示权重。

对于 Actor 网络训练的目标是寻找一个最优策略去最大化奖励，最优策略的定义如式（10）所示：

$$\pi^*(a \mid \boldsymbol{s}) = \underset{\pi}{\operatorname{argmax}}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] \tag{10}$$

其中，$\pi^*(a \mid \boldsymbol{s})$ 表示某一状态 $\boldsymbol{s}$ 下选择动作 a 的最大化奖励，$\underset{\pi}{\operatorname{argmax}}$表示寻找能最大化累计奖励的策略，$\sum_{t=0}^{\infty} \gamma^t r_t$ 表示未来累计奖励，用来衡量策略的好坏。

在学习的过程中，需要不断和设备进行交互以提升 Actor 的性能，进而寻找最优策略，故采用策略梯度方法优化其参数。因此，对于 Actor 网络的优化目标函数如式（11）（12）所示：

$$J(\theta) = E_{\pi}[A(\boldsymbol{s},a) \cdot \log \pi_{\theta}(a \mid \boldsymbol{s})] \tag{11}$$

$$A(\boldsymbol{s},a) = r_t + \gamma \cdot V(s_t + 1) - v_{\pi^{(s)}} \tag{12}$$

其中，$A(\boldsymbol{s}, a)$ 为优势函数，用来衡量一个策略相对于平均水平的好坏；$\pi_{\theta}(a \mid \boldsymbol{s})$ 表示给定状态 $\boldsymbol{s}$ 下，采取动作 a 的策略概率分布；r_t 表示在当前状态 s_t 下执行动作 a 后获得的即时奖励；γ 为折扣因子。因此 Actor 网络的策略优化梯度如式（13）所示：

$$\theta \leftarrow \theta + \alpha \nabla_{\theta} J(\theta) \tag{13}$$

其中，θ 为学习率，α 为 Actor 网络参数。

Critic 网络的主要任务是评估当前状态的价值，并为 Actor 网络提供反馈信息。对于 Critic 网络的训练目标是最小化状态值函数和 TD Target（时间差分目标）之间的误差，即 TD 误差，通过这种方式来提高值函数的估计精度。因此，对于 Critic 网络的优化目标函数如式（14）所示：

$$J(\varphi) = \operatorname{argmin}\left[A(\boldsymbol{s}, a)\right]^2 \tag{14}$$

其中，φ 表示 Critic 网络参数。

由 Actor 网络的优化目标函数和 Critic 网络的优化目标函数的表达式可以得出，Actor-Critic 网络训练的损失函数如式（15）所示：

$$F_{\mathrm{loss}} = \lambda\left[A(\boldsymbol{s}, a)\right]^2 - J(\theta) \tag{15}$$

其中，λ 为超参数，用于平衡 Actor 和 Critic 网络的损失。

为实现 Actor 网络对变异策略的精准选择与执行，将定义的 13 种变异算子与动作空间一一映射，进行唯一动作编码，形成固定的动作空间映射表。动作编码采用整数连续编码方式，编码范围为 1 ~ 13，每个整数对应唯一变异算子，训练过程中保持编码与算子的对应关系不变，确保策略输出的一致性。

Actor 网络接收 CTCA-Net 输出的特征向量后，通过全连接层输出 13 维概率分布向量，向量中每个维度对应上述映射表中的动作编码。选择概率最高的动作编码，查询映射表得到对应的变异算子，调用该算子对原始协议报文进行变异处理，生成测试用例。

2.5 变异算子库

通过对漏洞库的总结，工控场景下常见的漏洞类型包括缓冲区溢出[26]、拒绝服务[27]、命令注入[28]等，造成这些漏洞的原因通常是程序或设备对异常报文未正确处理。通过对异常报文的结构与内容进行分析，筛选出对报文异常产生影响的字段取值并进行分类，最终存储为变异算子，如表 2 所示。

表 2　变异算子库

变异策略类型	变异算子
字段级变异	字段值边界（极值、越界、空值）
	字段类型混淆（枚举非法值、布尔转数值）
	长度字段篡改（矛盾值、超长值）
	字段顺序打乱（删除重复字段）
结构变异	畸形报文构造（截断、协议片段拼接）
	状态机破坏（跳过连接流程、重复状态文）
	协议层次干扰（VLAN 标签篡改）
数据内容变异	字符集攻击（ASCII 插入）
	数值型变异（NaN、随机噪声）
	字符串攻击（超长、特殊字符）
	二进制位翻转/随机字节插入
时序变异	会话 ID 篡改（无效/重复 ID）
	时序干扰（乱序发包、延迟响应）

为提升漏洞挖掘效率，将表 2 中四种变异策略类型的优先级由高到低依次划分为 P1 ~ P4。该逻辑以关键影响度 - 协议依赖性 - 异常触发概率为核心，优先聚焦协议核心语义字段，协议依赖性侧重结构与状态机相关算子，异常触发概率则依据历史漏洞数据筛选高效算子。执行时，Actor 网络优先选择高优先级中的变异算子，若连续多次未触发设备响应变化，则动态向下兼容低优先级算子，既聚焦关键环节又避免局部搜索盲区。

3 实验验证

3.1 实验环境

为验证所述方法的有效性和可行性，在典型能源企业工业场景的攻防演练靶场上进行测试，该平台包含多个厂家的 PLC，支持 EtherNet/IP、S7comm 等工控协议。

该靶场网络拓扑图如图 5 所示，包含从监控层到设备层的全流程关键节点与网络设备布局。

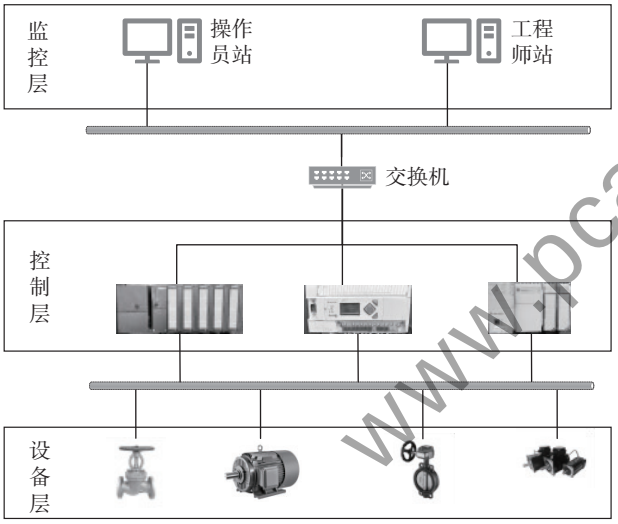

图 5　攻防演练靶场网络拓扑图

3.2 实验设置

3.2.1 实验硬件配置

实验中主要针对常用工控网络协议 Modbus TCP、EtherNet/IP、S7Comm 进行测试。采用 PyTorch 2.4.0 框架构建模型，训练环境为 Python 3.12.8，采用的主机硬件是 Intel Core i9-14900K CPU @ 3.00 GHz、Windows 11 操作系统、64 GB DDR5-6400 RAM NVIDIA GeForce RTX 4090 显卡。

3.2.2 模型参数配置

初始化的 CNN 模型采用 3 层一维卷积层，卷积核大小为 3 × 3，每层输出通道数均为 128，激活函数采用 ReLU。

初始化的 TCN 模型设置 3 层因果卷积，膨胀率分别为 1、2、4，卷积核大小为 3 × 3，每层输出通道数为 128，均采用残差连接与层归一化，确保梯度稳定传播。初始化注意力机制中注意力权重矩阵维度为 128 × 64，通过线性变换矩阵将 TCN 输出特征映射至 64 维空间计算注意力得分，最终输出特征向量维度为 256。

初始化的 Actor 网络采用 3 层全连接网络结构，其中隐藏层维度依次为 256、128，输出层采用 Softmax 激活函数以输出动作的概率分布，学习率设置为 0.001。Critic 网络采用 2 层全连接网络，隐藏层维度为 128，激活函数使用 ReLU，学习率设置为 0.005，训练时的批次大小为 64。

3.3 实验数据

3.3.1 公开数据集

SWaT 数据集（Secure Water Treatment Dataset）是由新加坡科技设计大学提供的一个针对水处理的工业控制系统安全性研究数字化仿真数据集，其中包含了 Modbus TCP、EtherNet/IP 等协议的正常操作数据和异常数据。

3.3.2 实验室数据

在实验室搭建典型能源企业工业场景攻防演练靶场，通过 Wireshark 工具对实际运行工况下的工控流量进行实时采集。Wireshark 能够精准捕获 S7Comm、EtherNet/IP 等主流工控协议的通信数据包。

3.4 评估指标

（1）测试用例的接收率（TCA）

测试用例的接收率表示有效测试用例的比例，即能够通过被测设备并被有效识别的测试用例数量，其定义如式（16）所示：

$$\mathrm{TCA} = \frac{N_a}{N_c} \times 100\% \tag{16}$$

其中，N_c 表示发送的测试用例总数，N_a 表示被测设备接收的测试用例数量。

（2）测试用例的多样性（DTC）

测试用例的多样性反映了在测试过程中能够覆盖协议的不同输入条件下的范围，其定义如式（17）所示：

$$\mathrm{DTC} = \frac{\sum_{i=1}^{M} N_i^{\,2}}{N^2} \tag{17}$$

其中，M 表示变异算子的总数，N_i 表示使用变异策略

M_i构建的测试用例，N 为总的测试用例数量。

（3）异常触发率（ATR）

在模糊测试中，重点是发现尽可能多的不同类型的异常或漏洞。通过发送给测试目标的所有测试用例中的异常数量来评估模型的性能。因此，触发的异常数量越多，越能反映异常的触发能力，其定义如式（18）所示：

$$\mathrm{ATR}=\frac{N_{\mathrm{q}}}{N_{\mathrm{c}}}\times 100\% \tag{18}$$

其中，N_{c} 表示发送的测试用例总数，N_{q} 示触发异常的测试用例数量。

3.5 消融实验

本节设计两组消融实验，分别验证加入 TCN 模型和注意力机制对提升模糊测试性能指标上的有效性。

3.5.1 时序特征提取有效性验证

为验证 TCN 模块可以提高测试用例的接收率，设计了第一组消融实验。对照组使用 CNN 结构提取协议报文特征，然后结合 Actor-Critic 框架构造测试用例，实验组则在此基础上加入 TCN 时序特征提取器，提取深层次的结构特征。实验以接收率为评价指标，使用相同的数据集训练同样的时间，记录在不同的测试设备上的接收率。

实验结果如图 6、图 7 所示，实验组相比于对照组，Modbus TCP 协议的接收率提高了 5.81%，EtherNet/IP 协议的接收率提高了 5.93%，S7Comm 协议的接收率提高了 6.85%。这表明 TCN 在建模协议字段的时序依赖关系方面具有更强的表达能力，能够更有效地构造结构合理、语义规范的测试用例，从而提升接收率，验证了加入 TCN 可以有效提升测试用例的接收率。

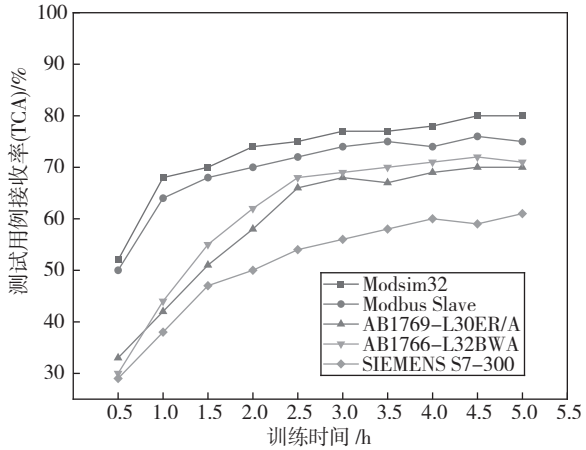

图 6　基于 CNN 的 TCA 结果

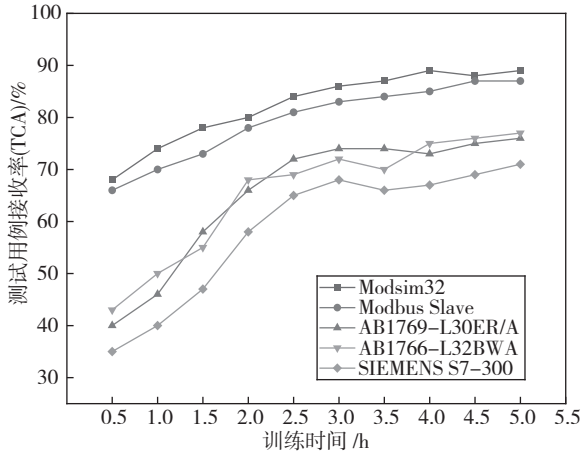

图 7　CTARFuzz 模型 TCA 结果

3.5.2 注意力机制有效性验证

为验证加入注意力机制可以提高测试用例的多样性，设计第二组消融实验。对照组同样使用 CNN 结合 Actor-Critic 框架，用于提取协议报文特征并构造测试用例，实验组使用 CTARFuzz，在特征建模过程中通过注意力机制增强对关键字段的识别能力。实验针对 Modbus TCP 协议进行测试，以测试用例中所覆盖的功能码类别数量作为多样性评估指标。

实验结果如图 8 所示。在训练初期，两组方法的多样性差异较小，但 CTARFuzz 凭借注意力机制的引导，能够更快地关注协议报文中的关键字段，使其构造的测试用例更具变异性，功能码类别数迅速增长。随着训练时间的延长，CNN-AC 模型逐渐趋于稳定，其策略网络陷入局部最优，导致功能码类别增长缓慢并最终停滞于约 7 类。相比之下，CTARFuzz 通过注意力机制能够动态调整，引导模型关注易被忽略但具有重要语义差异的字段区域，从而构造更加多样化的测试用例，在 5 h 训练时间内功能码类别拓展至 14 类。

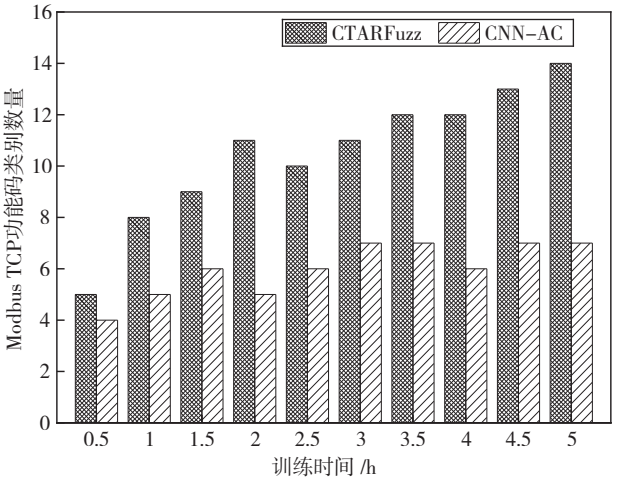

图 8　功能码类别对比

3.6 模型超参数确定

为确保 CTCA-Net 模型与 Actor-Critic 强化学习框架的协同性能最优，针对不同超参数进行对比试验，确定出适配不同工控协议的最佳参数组合，实验结果如表 3 所示。

表 3 超参数对比实验结果

CNN 卷积核	TCN 膨胀率	TCA /%	DTC
3×3	1、2、4	87.1	11
	1、4、8	88.4	13
5×5	1、4、8	88.2	13
	2、4、8	87.8	12
7×7	1、4、8	87.4	11
	3、6、12	87.7	13

结果表明，不同 CNN 卷积核与 TCN 膨胀率组合下，生成的测试用例接收率及多样性表现有差异。其中，CNN 卷积核为 3×3，TCN 膨胀率取 1、4、8 时，DTC 达到 13 种类型的数据，同时 TCA 为 88.4%，在所有组合中表现最优。该组合能精准提取报文特征并有效建模时序依赖，因此在随后的测试效率对比实验中将 CNN 的卷积核设置为 3×3，TCN 的膨胀率设置为 1、4、8。

3.7 奖励函数权重系数确定

为保证奖励函数里权重系数的客观性与模型性能，本文采用网格搜索方法确定最优权重组合，设定权重区间为 λ_1，λ_2，$\lambda_3 \in [0, 1]$ 且 $\lambda_1 + \lambda_2 + \lambda_3 = 1$，以 0.2 为步长生成 10 个组合。实验使用 10 000 条混合的 Modbus TCP、EtherNet/IP、S7Comm 协议，各协议之间的比例为 3 ∶ 4 ∶ 3，以测试用例的接收率和异常触发率为评价指标，实验结果如图 9 所示。

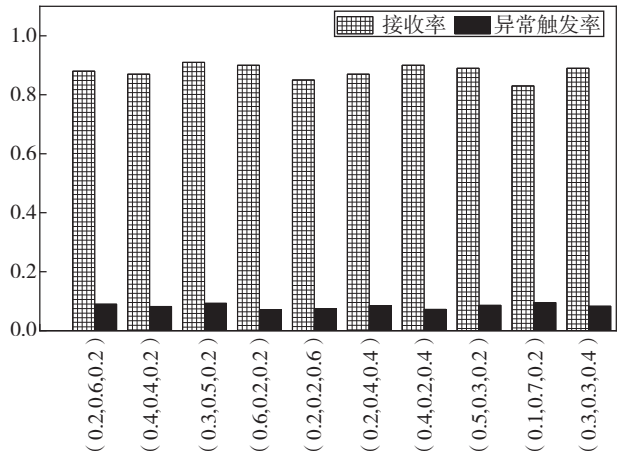

图 9 奖励函数权重系数对比

实验结果表明，当 $\lambda_1 = 0.3$，$\lambda_2 = 0.5$，$\lambda_3 = 0.2$ 时，模型拥有较高的接收率与异常触发率，因此，本文选择此参数为奖励函数的权重系数。

3.8 测试效率对比分析

本文将 Peach、MTAFuzz、CTFuzz 与 CTARFuzz 进行对比实验。其中，Peach 作为当前工业安全领域中广泛应用的模糊测试工具之一，通过创建 XML 文件，执行模糊测试；MTAFuzz 是基于 Transformer 构建的模糊测试框架；CTFuzz 是依据强化学习进行的模糊测试。通过对比实验，各模型的性能指标如下。

（1）TCA：使用 Peach、MTAFuzz、CTFuzz 和 CTARFuzz 分别对不同的测试设备进行 5 h 的模糊测试实验，结果如图 10 所示。可以看出，Peach 作为传统模糊测试工具，缺乏协议感知能力，在结构复杂的 S7Comm 与 EtherNet/IP 协议中，测试用例常因语义错误被目标设备拒绝，TCA 明显偏低。

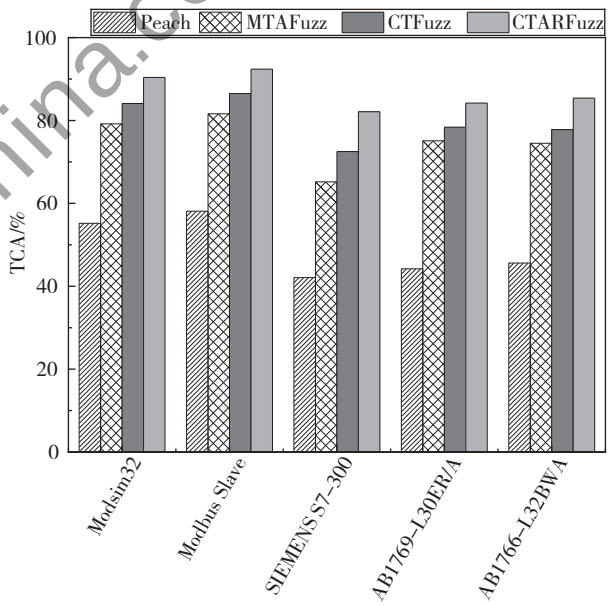

图 10 TCA 对比实验结果

MTAFuzz 在 Modbus TCP 中表现较好，但由于缺乏时序建模与语义提取，其生成的用例在复杂协议下适配性差，导致接收率较低。CTFuzz 虽引入强化学习策略，在一定程度上具备输入自适应能力，较 Peach 和 MTAFuzz 有一定提升，但其强化学习策略对协议状态理解仍不充分，泛化能力有限。

相比之下，CTARFuzz 在所有协议场景中均实现了最高的测试用例接收率，在 Modbus TCP 上的 TCA 达到 90%，在 S7Comm 与 EtherNet/IP 复杂协议中也拥有较高的接收率。这得益于 CTARFuzz 利用 CTCA-Net 模

型提取深层语义特征，并且能够根据目标设备反馈动态优化输入生成策略，从而显著提升测试用例的接收率。

（2）ATR：对不同的测试目标发送 30 000 个测试用例，记录测试过程中出现的异常情况，实验结果如表 4 所示。结果表明，CTARFuzz 在 Modbus TCP、S7Comm、Ethernet/IP 三种协议上的 ATR 指标分别为 0.400%、0.027%、0.039%，高于其他几种方法，其中在对 Modsim32 仿真软件测试时，CTARFuzz 的性能最好，触发的异常数量最多，ATR 达到了 0.513%，相较于 Peach、MTAFuzz、CTFuzz 提高了 105.2%、23.02%、16.59%。

表 4　ATR 实验结果对比

测试方法	测试时间/h	用例数量	测试设备	异常数量	ATR%	触发异常种类
Peach	—	30 000	Modsim32	75	0.250	3
			Modbus Slave	46	0.153	2
			SIEMENS S7-300	4	0.013	1
			AB1769-L30ER/A	3	0.010	1
			AB1766-L32BWA	6	0.020	2
MTAFuzz	15.97	30 000	Modsim32	125	0.417	4
			Modbus Slave	58	0.193	5
			SIEMENS S7-300	6	0.020	2
			AB1769-L30ER/A	8	0.027	4
			AB1766-L32BWA	9	0.030	3
CTFuzz	14.21	30 000	Modsim32	132	0.440	4
			Modbus Slave	62	0.207	6
			SIEMENS S7-300	4	0.013	3
			AB1769-L30ER/A	6	0.020	3
			AB1766-L32BWA	8	0.027	4
CTARFuzz	12.43	30 000	Modsim32	154	0.513	6
			Modbus Slave	86	0.287	6
			SIEMENS S7-300	8	0.027	3
			AB1769-L30ER/A	12	0.040	5
			AB1766-L32BWA	11	0.037	4

在对 SIEMENS S7-300、AB1769-L30ER/A 和 AB1766-L32BWA 测试时，由于 S7Comm 和 Ethernet/IP 协议采用分层架构，并通过校验、序列号、心跳机制等实现数据可靠传输，导致几种方法的 ATR 都较低，但 CTARFuzz 拥有反馈机制，能很好地根据设备反馈调整变异策略，其在拥有较高的异常触发能力的同时所用的时间远低于 MTAFuzz 和 CTFuzz，显著提升了模糊测试的效率。

（3）DTC：对比几种模型生成的测试用例，如图 11所示。可以发现，在训练时间相同的实验条件下，CTARFuzz 始终维持在较高的多样性水平，且波动幅度较小，体现出卓越的稳定性，能持续稳定地构造不同的测试用例，而其他模型都存在一定程度的波动。稳定且多样的测试用例，可使模糊测试更全面地覆盖输入空间，大幅减少测试盲区，从而显著提升测试效率。

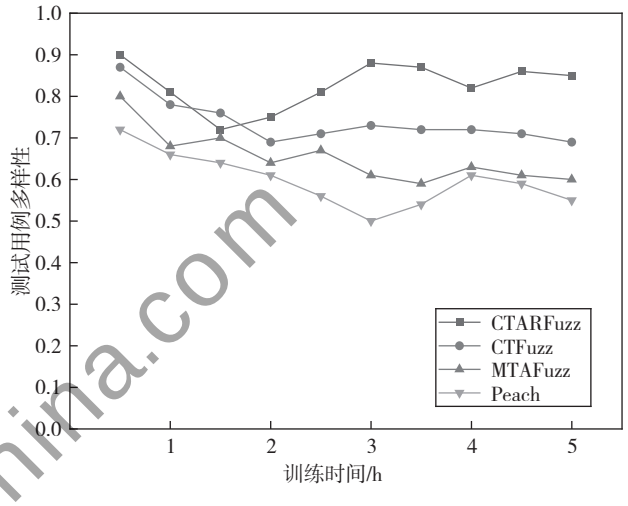

图 11　DTC 对比实验结果

3.9　异常结果分析

CTARFuzz 框架在对实际的 PLC 设备和仿真通信工具进行实际测试中，发现了一系列影响设备正常运行的异常情况，如表 5 所示。

表 5　异常结果

异常类型	Modsim32	Modbus Slave	SIEMENS S7-300	AB1769-L30ER/A	AB1766-L32BWA
无法连接	√		√	√	
异常功能码	√	√			
缓冲区溢出	√	√	√		
数据类型不匹配		√		√	√
从站无响应	√	√			
连接断开	√		√	√	√
异常功能码	√	√			
异常设备地址		√			

针对 EtherNet/IP 协议进行测试时，在 AB1769-L30ER/A 设备中发现异常。具体来说，利用模型构建协议中 Session Handle 部分数据，并通过指定端口发送测试用例给 PLC 设备，导致设备与上位机断开连接，

再次请求连接时显示连接超时，不能建立连接。此外，在 AB1766-L32BWA 设备中，通过发送大量的测试用例会导致会话资源池泛滥，从而断开上位机与 PLC 的合法连接。

对 Modbus TCP 协议进行测试时，在 Modbus Slave 和 Modsim32 中发现多个异常。异常一：在请求验证机制上存在异常，响应数据格式不匹配请求格式。具体而言，构造的测试用例的请求 PDU 长度为 06，响应的 PDU 长度为 08，表明响应数据长度和请求数据长度不匹配，影响软件正常的数据交互功能。异常二：对于未定义的功能码，仍能正确响应并返回部分数据，未能正确拒绝非法功能码请求。

在对 S7Comm 协议测试过程中，构造的测试用例导致 SIEMENS S7-300 的 CPU 停止运行。具体来说，是由于协议本身缺乏身份认证机制，构造异常请求报文，修改 PLC 内存区域的数据，使设备执行异常操作，导致 CPU 断开连接。通过对实际 PLC 设备及仿真通信工具的实验证明，CTARFuzz 在异常触发方面表现优异。

4 结论

本文提出一种融合深度特征与强化学习的工控协议模糊测试方法 CTARFuzz，通过 CTCA-Net 模型提取协议报文深层语义特征，结合 Actor-Critic 强化学习框架动态优化变异策略。实验结果表明，CTARFuzz 在 Modbus TCP、EtherNet/IP、S7Comm 协议上其测试用例接收率、异常触发率均优于 Peach、MTAFuzz、CT-Fuzz，能触发多类设备异常，因 Modbus TCP 无分层与复杂校验，CTARFuzz 特征提取与策略优化效率高。EtherNet/IP、S7Comm 为分层协议，依赖校验与会话管理，异常报文易被过滤，ATR 虽高于对比方法，但低于 Modbus TCP。

当前方法还存在一定局限性，强化学习对私有协议泛化性弱，适配中小规模工控场景及常见的工控协议，在大规模分布式场景中，模型对硬件资源依赖高，容易造成设备延迟导致奖励值出现偏差。

未来将结合协议语法树与语义知识库增强语义理解，引入迁移学习优化强化学习泛化性，通过模型轻量化降低硬件消耗，进一步提升测试性能。

参考文献

[1] 宗学军，隋一凡，王国刚，等. 基于生成对抗网络的工控协议模糊测试研究 [J]. 网络安全与数据治理，2024，43 (7)：13 -20.

[2] STOUFFER K，FALCO J，SCARFONE K. Guide to industrial control systems (ICS) security [R]. NIST Special Publication (SP) 800 -82，2011.

[3] KOAY A M Y，KO R K L，HETTEMA H，et al. Machine learning in industrial control system (ICS) security：current landscape，opportunities and challenges [J]. Journal of Intelligent Information Systems，2023，60：377 -405.

[4] SCHERLING L S. The future of hacking：the rise of cybercrime and the fight to keep us safe [M]. Bloomsbury Publishing USA，2025.

[5] KAYAN H，NUNES M，RANA O，et al. Cybersecurity of industrial cyber-physical systems：a review [J]. ACM Computing Surveys (CSUR)，2022，54 (11 s)：1 -35.

[6] 庄园，曹文芳，孙国凯，等. 基于生成对抗网络与变异策略结合的网络协议漏洞挖掘方法 [J]. 计算机科学，2023，50 (9)：44 -51.

[7] ZONG X，LUO W，NING B，et al. DiffusionFuzz：fuzzing framework of industrial control protocols based on denoising diffusion probabilistic model [J]. IEEE Access，2024，12：67795 -97808.

[8] JANIESCH C，ZSCHECH P，HEINRICH K. Machine learning and deep learning [J]. Electronic Markets，2021，31 (3)：685 -695.

[9] SHAKYA A K，PILLAI G，CHAKRABARTY S. Reinforcement learning algorithms：a brief survey [J]. Expert Systems with Applications，2023，231：120495.

[10] CHENG M，ZHU K，CHEN Y，et al. MSFuzz：augmenting protocol fuzzing with message syntax comprehension via large language models [J]. Electronics，2024，13 (13)：2632.

[11] YANG H，HUANG Y，ZHANG Z，et al. A novel generative adversarial network-based fuzzing cases generation method for industrial control system protocols [J]. Computers and Electrical Engineering，2024，117：109268.

[12] CHE X，GENG Y，ZHANG G，et al. Fuzzing technology based on information theory for industrial proprietary protocol [J]. Electronics，2023，12 (14).

[13] WANYAN H，LAI Y，LIU J，et al. NCMFuzzer：using non -critical field mutation and test case combination to improve the efficiency of ICS protocol fuzzing [J]. Computers & Security，2024，141：103811.

[14] ARCHANA R，JEEVARAJ P S E. Deep learning models for digital image processing：a review [J]. Artificial Intelligence Review，2024，57 (1)：33.

[15] ELIAZER M，JERRELL F C，SHRIVAS S. Using temporal convolutional networks (TCN) deep learning model for crop recommendation [J]. SSRN Electronic Journal，2024. DOI：10.2139/ssrn.4824959.

[16] 连莲，孙世明，王国刚，等. 基于多尺度潜在特征表示的工业控制协议模糊测试方法 [J]. 计算机应用研究，

2025，42（2）：545－554.
[17] ZHANG Z，ZHANG H，ZHAO J，et al. A survey on the development of network protocol fuzzing techniques [J]. Electronics，2023，12（13）：2904.
[18] 姜亚光，陈曦，李建彬，等. 基于 LSTM 的 S7 协议模糊测试用例生成方法 [J]. 计算机工程，2021，47（7）：183－188.
[19] ALDYSTY A R，MOUSTAFA N，LAKSHIKA E. A holistic review of fuzzing for vulnerability assessment in industrial network protocols [J]. IEEE Open Journal of the Communications Society，2025，6：4437－4461.
[20] QASEM A，SHIRANI P，DEBBABI M，et al. Automatic vulnerability detection in embedded devices and firmware：survey and layered taxonomies [J]. ACM Computing Surveys（CSUR），2021，54（2）：1－42.
[21] ROMERO A，SONG Y，SCARAMUZZA D. Actor－critic model predictive control [C] //2024 IEEE International Conference on Robotics and Automation（ICRA）. IEEE，2024：14777－14784.
[22] LIU Y，BAI X，WANG J，et al. Image semantic segmentation approach based on DeepLabV3 plus network with an attention mechanism [J]. Engineering Applications of Artificial Intelligence，2024，127：107260.
[23] SI X，SONG Y，SUN X，et al. MARLFuzz：industrial control protocols fuzzing based on multi-agent reinforcement learning [J]. Computing，2025，107（2）：1－26.
[24] ZHAO F. Application and performance improvement of K-Means algorithm in collaborative [C]//2025 International Conference on Intelligent Systems and Computational Networks（ICISCN）. IEEE，2025：1－6.
[25] REZVAN M R，SORKHI A G，PIRGAZI J，et al. AdvanceSplice：integrating N-gram one-hot encoding and ensemble modeling for enhanced accuracy [J]. Biomedical Signal Processing and Control，2024，92：106017.
[26] KEROMYTIS A D. Buffer overflow attacks [M] //Encyclopedia of Cryptography，Security and Privacy. Springer，Cham，2025：309－312.
[27] QASIM S，NSAIF S M. Advancements in time series-based detection systems for distributed denial-of-service（DDoS）attacks：a comprehensive review [J]. Babylonian Journal of Networking，2024，2024：9－17.
[28] PASQUINI D，KORNAROPOULOS E M，ATENIESE G. Hacking back the AI-hacker：prompt injection as a defense against LLM-driven cyberattacks [J]. arXiv preprint arXiv：2410.20911，2024.

（收稿日期：2025－10－12）

作者简介：

宗学军（1970－），通信作者，男，硕士，教授，主要研究方向：工业信息安全。E-mail：xuejun_zong@syuct.edu.cn。

孙俊辉（2000－），男，硕士研究生，主要研究方向：工业信息安全、漏洞挖掘。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部