基于深度学习的计算机通信网络 APT 攻击检测方法

李海芳[1] 路晓亚[1]
LI Haifang LU Xiaoya

摘 要 APT 攻击通过复杂手段干扰网络行为归约，不同安全检测设备会产生告警链。由于 APT 攻击具有隐蔽性，告警链中的信息难以被整合，导致在网络拓扑结构发生变化或者流量模式出现调整时，会产生大量误报现象，检测的精准率也会显著降低。为此，提出一种基于深度学习的计算机通信网络 APT 攻击检测方法。通过收集多种 APT 攻击类型的数据构建目标集，在此基础上，设计一种融合卷积神经网络与循环神经网络的深度学习模型，提取 APT 攻击的特征，利用 softmax 分类器依据提取的特征，对定位到的异常流量情况进行精准判定，实现对 APT 攻击的精准检测与分类识别。对比实验结果表明，该方法应用后，AUC 值达到了 0.913，证明其在识别网络 APT 攻击方面具有较高的准确性和鲁棒性，应用效果较好。

关键词 计算机通信网络；APT 攻击；深度学习模型；麻雀搜索算法；softmax 分类器

doi：10.3969/j.issn.1672-9528.2025.02.001

0 引言

当前，网络安全威胁日益严峻，高级持续性威胁（advanced persistent threat，APT）便是其中一种。APT 隐蔽性强、持续时间长、危害程度高，对国家安全、企业运营以及个人隐私都构成了巨大威胁。APT 攻击会运用复杂的手段对网络行为归约造成干扰。当这种情况发生时，不同的安全检测设备会相应地产生告警链。然而，由于 APT 攻击本身所具有的高度隐蔽性，使得这些告警链中的信息难以被有效整合。这种情况带来的后果是，一旦网络拓扑结构发生改变或者流量模式出现调整，就会产生大量的误报现象，进而导致检测的精准率显著下降。为有效应对 APT 攻击，众多学者开展了 APT 攻击检测方案研究。如李元诚等人[1]提出了基于溯源图和注意力机制的 APT 攻击检测模型，该模型通过构建系统行为的溯源图，结合深度神经网络和注意力机制，有效地识别了恶意攻击行为。该方法的弊端是在 APT 攻击干扰网络行为归约的环境下，由于网络行为归约受到干扰，用于构建溯源图的基础数据可能存在偏差。当网络拓扑结构发生变化或者流量模式出现调整时，被 APT 攻击混淆后的网络行为数据无法准确反映真实的系统行为关系，致使后续恶意攻击行为识别出现误判。另一方面，陈泽红[2]研究了基于自适应模糊聚类的无监督 APT 攻击检测方法，通过聚类算法分析未加类别标识的入侵检测样本数据，实现了对 APT 攻击的自动检测。在 APT 攻击对网络行为归约造成干扰的情况下，未加类别标识的入侵检测样本数据可能已被扭曲。因为网络行为归约是为了提取网络行为的关键特征和模式，受到干扰后，样本数据中的正常与异常行为模式就可能被混淆。当网络拓扑结构发生变化或者流量模式出现调整时，聚类算法不能及时适应这些变化，依旧按照旧的模型进行聚类，这就会导致 APT 攻击检测的准确性降低。

针对上述研究存在的不足，本文提出一种基于深度学习的计算机通信网络 APT 攻击检测方法。

1 构建全面 APT 攻击目标集

在计算机通信网络中，APT 攻击往往针对特定的、高价值的数字资产，且具有高度的隐蔽性和持续性，攻击路径复杂多变。为有效应对这一挑战，需要先明确攻击目标，构建一个计算机通信网络 APT 攻击目标集。该集合聚焦于目标网络中潜在的高价值数字资产，视为 APT 攻击者的主要目标[3]。为表征这种多层次的攻击路径，引入金字塔模型（GP），其中每个目标 i（i=1,2,⋯,n）映射为一个独特的金字塔实例。金字塔的顶点 P 代表最终目标，L 个侧面对应网络空间环境的不同维度或“位面”，其公式为：

$$\mathrm{GP}_i=\left(P_i,\{S_1,S_2,...,S_m\}\right) \tag{1}$$

不同目标 i 和 j 的金字塔实例之间可能存在位面共享，反映了攻击者可能利用的通用攻击路径或弱点。攻击者通过收集网络公开信息、追踪网络活动踪迹、实施网络扫描探测以及破解受限访问权限等手段，积累关于目标网络的详尽数据[4]。此外，还频繁采用社会工程学技巧，如利用社交媒体

1. 商丘工学院信息与电子工程学院 河南商丘 476000
[基金项目] 河南省科技攻关项目“智慧教育评价系统中基于深度学习的表情识别技术研究”（242102210111）

互动或通过分析公开信息辅以心理操纵，来推测或直接获取敏感信息。基于收集的信息，攻击者会开发或定制针对性攻击工具，并在模拟目标网络环境的测试平台上进行验证，尤其是针对自主研发的攻击工具进行严格的效能测试，以确保其有效性和隐蔽性。

在 APT 攻击的传输阶段，恶意内容（如代码、链接）的传递方式可区分为直接传输与间接传输两大类别 [5]。依据实施手段的不同，攻击可细分为鱼叉式网络钓鱼、水坑攻击及中间人攻击 [6-7]。一系列事件集构成位面，其公式为：

$$A=\left\{\frac{e}{e_1},\frac{e}{e_2},...,\frac{e}{e_n}\right\} \tag{2}$$

式中：对于每个事件$\frac{e}{e_n}$可表示为 $(U_i, T, \mathrm{id}, \Lambda)$，其中，$U_i$ 表示事件$\frac{e}{e_i}$在位面 A 中的索引；T 表示事件发生的时间戳；id 表示事件 ID；Λ 表示位面中所有事件的属性序列。通过上述步骤，可系统地建立计算机通信网络 APT 攻击目标集，为后续的安全防护和应急响应提供有力支持。

2 基于深度学习与麻雀搜索算法的 APT 攻击特征提取

计算机通信网络 APT 攻击目标集的构建可为模型训练奠定基础。由于 APT 攻击往往具有复杂性和隐蔽性，其流量数据往往呈现出多变且难以预测的特性，使得传统的检测策略往往难以适应，造成 APT 攻击的遗漏或误报。为解决这一问题，本次引入融合了 CNN 和 BiGRU 架构的深度学习模型，并基于麻雀搜索算法（SSA）算法完成模型训练，以实现海量数据中 APT 攻击特征的提取。该模块流程分为 5 步。

（1）预处理：输入数据为向量化并融合的数据集，通过时间序列分组，每组调整为 50 维（若原始流量数据维度不能被 50 整除，则采用零填充（Zero-padding）和边缘滤波处理，确保输入为 50×70 的矩阵）。

（2）CNN 层：利用卷积神经网络（CNN）初步提取空间特征，CNN 的输入即为上述处理后的矩阵。

（3）注意力机制：在 CNN 层后引入注意力机制，动态调整特征权重，增强模型对关键特征的关注度 [8]。注意力权重的计算可简化为对 CNN 输出特征的加权处理，其中注意力概率的计算公式为：

$$\alpha_i=\frac{\exp(e_i)}{\sum_{j=1}^{N}\exp(e_j)} \tag{3}$$

式中：α_i 表示每个特征元素的权重；e_i 表示第 i 个元素的原始得分；N 表示特征的总数；$\exp(\cdot)$ 表示指数函数，用于将得分转换为非负值；$\sum_{j=1}^{N}\exp(e_j)$ 表示所有元素得分的指数和，用于归一化。

（4）BiGRU 层：将注意力机制调整后的特征向量作为双向门控循环单元（BiGRU）的输入，捕捉序列的前后文信息。

（5）SSA 优化：在 BiGRU 层与全连接层之间，采用 SSA 算法优化网络权重，通过迭代搜索最优解，提升模型性能 [9]。SSA 算法的核心在于模拟麻雀的觅食和反捕食行为，通过不断更新麻雀位置（即网络权重）来逼近最优解。在该算法中，先构造麻雀种群，每只麻雀代表一组网络权重，初始位置随机或基于某种启发式方法设定 [10]。然后计算每只麻雀（即每组权重）对应的适应度值，通常基于模型在验证集上的表现。根据适应度与安全值的关系，调整麻雀的位置（即更新网络权重），旨在寻找食物源（即优化目标）并避免被捕食。最后重复上述过程，直至满足收敛条件（如多数麻雀收敛于同一解或达到最大迭代次数），此时得到的权重即为最优解 [11]。通过上述流程，该模型能从复杂的网络流量数据中提取出用于 APT 攻击检测的关键特征，提高检测的准确性和效率。

3 基于 softmax 分类器的 APT 攻击检测设计

完成 APT 攻击特征提取后，可以进一步利用这些特征进行网络中的 APT 攻击检测。接下来，本次实验通过 softmax 分类器实现二分类，以区分正常网络流量与 APT 攻击流量，从而实现网络中的 APT 攻击检测。在实时环境中，为了实现对网络 APT 攻击的快速识别，执行了以下步骤。首先，构建了一个数据集，该数据集包含了来自不同 APT 家族的 Taidoor、Scieron 等 13 类攻击流量，这些流量被统一标记为 1（代表异常）。同时，为了形成对比，从 CICIDS2017 数据集中选取了同一的常规网络流量，并将其标记为 0（代表正常）[12-13]。随后，对预处理后的数据进行特征矢量化。在神经网络的权重优化过程中，传统的优化算法可能会因为初始值的选择或者搜索策略的局限性，导致找到的权重只能使模型达到局部最优的性能，而 SSA 的全局搜索特性能够在更广泛的范围内探索可能的权重值，其通过模拟麻雀觅食和反捕食的行为，不断调整搜索策略，从而在整个权重空间中找到更有利于准确分类正常流量和 APT 攻击流量的权重组合，进而提高模型在实时监测中的准确性和响应速度 [14]。与深度学习模型相结合，充分利用这两者的优势进行高级特征提取与模式学习。最终，采用 softmax 分类器来实现二分类任务，即将网络流量判定为正常或 APT 攻击 [15]。softmax 分类器位于模型的输出层，根据输入特征矢量计算出模型最后一层的输出，并给出输入属于各个类别的概率。在 APT 检测任务中，特别关注流量被判定为 APT 攻击的概率。二分类问题的输出结果表示为：

$$P(y=j|x)=\frac{e^{zj}}{\sum_{k=1}^{K}e^{zk}} \tag{4}$$

式中：x 表示输入特征矢量；z 表示模型最后一层的输出，对于二分类问题，K=2；$P(y=j|x)$ 表示输入 x 属于类别 j 的概率。在 APT 检测中，始终关注 $P(y=1|x)$，即流量为 APT 攻击的概率。综上，完成计算机通信网络 APT 攻击检测设计。

4 对比实验

4.1 实验环境配置

在本实验环境中，采用 Windows 10 操作系统平台上的 PyCharm 集成开发环境（IDE）作为代码编写与调试的主要工具。为构建和执行深度学习模型，选择 TensorFlow 这一广泛应用的框架。在硬件方面，本实验所依赖的 CPU 为 Intel 的 i7 系列处理器，其详细配置信息如表 1 所示。

表 1 实验环境配置表

序号	配置	参数
（1）	操作系统	Windows 10 64 位操作系统
（2）	内存	24 GB
（3）	深度学习框架	Tensorflow
（4）	CPU	NVIDIA GeForce GTX 960M（4095 MB）
（5）	编程语言	Python3.8.9

该配置确保了在实验过程中数据处理与模型训练能够高效进行。在分析本文设计深度学习的计算机通信网络 APT 攻击检测方法性能的过程中，构建包含 9 个节点的交互网络环境，其中，各节点之间的联通关系设置情况如图 1 所示。

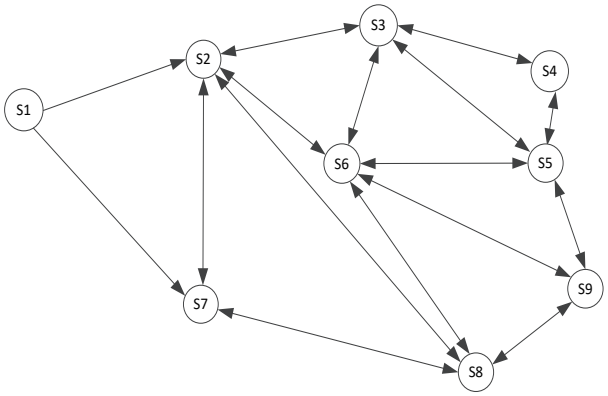

图 1 交互网络环境设置图

为体现实验的客观性，在实验中将本文上述提出的检测方法设定为实验组，将基于溯源图和注意力机制的检测方法、基于自适应模糊聚类的检测方法作为对照条件，分别设置为对照 A 组和对照 B 组。利用这 3 种检测方法，在上述配置的相同环境中对计算机通信网络 APT 攻击进行检测。

4.2 实验参数设置

在实验过程中，对于参数的配置，主要针对本文方法中的深度学习参数设置，其具体内容如表 2 所示。

表 2 实验参数设置表

序号	参数名称	数值
（1）	BatchSize	50
（2）	Dropout 率	0.5
（3）	优化器	Adam
（4）	学习率	0.000 1

这些参数直接关系到本文所提出方法的性能表现。通过上述参数设置，旨在优化模型的训练效果与泛化能力。

4.3 实验评价指标

为了量化评估 3 种检测方法的效能，设定 APT 流量作为正面实例，并运用混淆矩阵这一工具，对方法的应用性能进行客观评价。混淆矩阵的具体内容如表 3 所示。

表 3 实验结果评价指标混淆矩阵

序号	样本类型	预测为负面例样本	预测为正面例样本
（1）	负面例样本	TN 真负例	FP 假正例
（2）	正面例样本	FN 假负例	TP 真正例

表 3 中，TN 表示正确地将负类预测为负类；FP 表示错误地将负类预测为正类；FN 表示错误地将正类预测为负类；TP 表示正确地将正类预测为正类。

依据表 3 数据，本次选取 AUC 值作为衡量检测效果的核心指标。实验结束后，直接依据结果执行基础运算，即可得出上述各指标的具体数值。为验证 3 种检测方法的有效性，对数据集进行了随机划分，比例设定为 7:2:1，分别用于训练、测试和预测，以确保方法的全面评估。

4.4 实验结果分析

先对所提方法进行功能检测，测试随着迭代次数的增加，所提 softmax 分类器对 APT 攻击的检测精度变化，得到结果如图 2 所示。

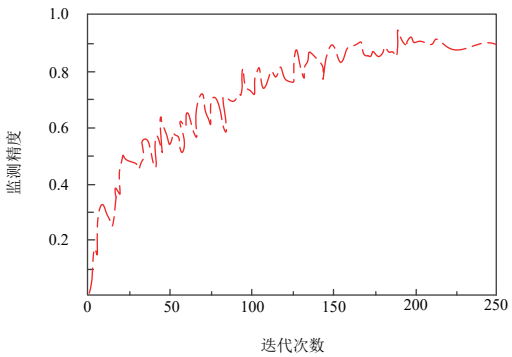

图 2 所用 softmax 分类器性能

如图 2 所示，随着迭代次数的增加，其监测精度逐渐提高，当迭代次数达到 190 次可呈现稳定状态，检测精度可达到 0.91，由此可证明，所提的 softmax 分类器在经过足够的

迭代训练后，能够高效地识别并检测 APP 攻击，展现出良好的稳定性和准确性。

完成所提方法功能性测定后，应用 3 种方法开展网络 APT 攻击检测，得到实验结果如图 3 所示。

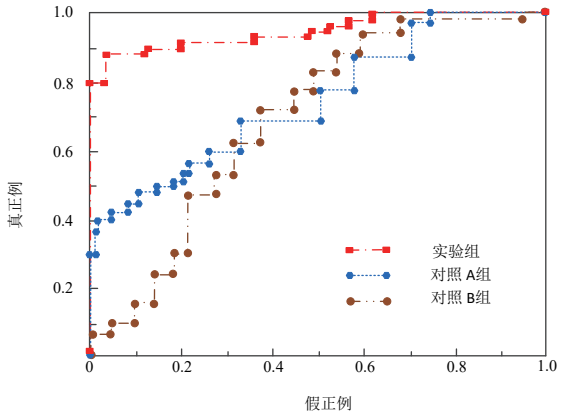

图 3 三种检测方法检测性能对比图

如图 3 所示，应用实验组得到的 AUC 可达到 0.913，而应用对照 A 组和对照 B 组，其 AUC 仅为 0.753 和 0.694，由此可证明，应用所提方法进行网络 APT 攻击检测时具有较高的准确性和鲁棒性。造成这一现象的原因在于所提方法先构建多种 APT 攻击类型的目标集，使得模型能够学习到更全面的 APT 攻击特征，从而提高对不同类型 APT 攻击的识别能力。然后融合卷积神经网络（CNN）和循环神经网络（RNN）构建深度学习模型，有效提取 APT 特征；最终，通过 softmax 分类器实现二分类，实现对 APT 攻击的精准检测，这使得实验组方法在应对计算机通信网络中的 APT 攻击时具有显著优势。

5 结语

为保障计算机通信网络稳定运行，本文提出了一种创新的计算机通信网络 APT 攻击检测方法，该方法深度融合了深度学习技术，实现了对 APT 攻击的高效且精准地检测。实验结果显示，相较于现有方法，该方法在检测精度与可靠性上均展现出显著优势，为提升网络通信安全提供了坚实的支撑。然而，鉴于 APT 攻击手段的持续演进与复杂化，未来的研究应当聚焦于深入探索并开发更为高效的特征提取与整合技术、与其他先进安全技术协同整合、共同构建更加坚固的网络安全防御体系等方向。

参考文献：

[1] 李元诚，罗昊，王欣煜，等．基于溯源图和注意力机制的 APT 攻击检测模型构建 [J]. 通信学报，2024, 45(3): 117-130.

[2] 陈泽红．基于自适应模糊聚类的无监督 APT 攻击检测方法研究 [J]. 网络安全技术与应用，2023 (7): 45-47.

[3] 田小芳．基于路径追踪的轨道交通系统 APT 攻击检测研究 [J]. 网络空间安全，2023, 14(3): 85-90.

[4] 丁顺莺．APT 攻击检测在大数据分析中的应用 [J]. 网络安全技术与应用，2023(6): 57-59.

[5] 董程昱，吕明琪，陈铁明，等．基于异构溯源图学习的 APT 攻击检测方法 [J]. 计算机科学，2023, 50 (4): 359-368.

[6] 谢丽霞，李雪鸥，杨宏宇，等．基于样本特征强化的 APT 攻击多阶段检测方法 [J]. 通信学报，2022, 43 (12): 66-76.

[7] 林玉坤，于新会，李元诚，等．基于改进 CNN 的新型电力系统 APT 攻击检测 [J]. 电力信息与通信技术，2023, 21 (6): 1-7.

[8] 张威武，朱江，马峥巍，等．基于沙箱的恶意文件与 APT 攻击检测方法改进研究 [J]. 智能物联技术，2022, 5(5): 23-31.

[9] 杨桂山，安庆．受 APT 攻击的网络实验室域名智能检测仿真 [J]. 计算机仿真，2022, 39(9): 441-445.

[10] 梁若舟，高跃，赵曦滨．基于序列特征提取的溯源图上 APT 攻击检测方法 [J]. 中国科学：信息科学，2022,52 (8): 1463-1480.

[11] 时林，时绍森，文伟平．基于 LSTM 的 Linux 系统下 APT 攻击检测研究 [J]. 信息安全研究，2022, 8 (8): 736-750.

[12] 梁鹤，李鑫，尹南南，等．结合动态行为和静态特征的 APT 攻击检测方法 [J]. 计算机工程与应用，2023, 59 (18): 249-259.

[13] 韩翔，李超，李鑫，等．基于 AlexNet 卷积神经网络的 APT 攻击检测技术研究 [J]. 保密科学技术，2022(6):42-48.

[14] 魏峰，张驯．基于 LSTM 模型的 APT 攻击信道检测方法 [J]. 微型电脑应用，2022, 38(3): 134-137.

[15] 庞九凤，张亚昊，胡威，等．高级持续性威胁（APT）攻击检测与防御的深度学习方法研究 [J]. 办公自动化，2024, 29(14): 73-76.

【作者简介】

李海芳（1991—），女，河南周口人，硕士，讲师，研究方向：信息安全、深度学习。

路晓亚（1983—），女，河南商丘人，硕士，副教授，研究方向：软件设计、深度学习。

（收稿日期：2024-10-16）