一种融合多源异构数据的图神经网络联合框架*

胡开明，陈建华
（广东松山职业技术学院，广东　韶关　512126）

摘　要：针对网络空间攻防对抗呈现出多步骤、隐蔽化、异构化复杂特征，传统依赖规则匹配和统计分析的方法已难以满足精准溯源与实时态势感知需求的问题，提出一种融合多源异构数据的图神经网络联合框架，实现网络攻击的自动化溯源与动态态势感知。首先，通过构建网络实体－攻击行为异构信息网络，整合流量日志、漏洞库、告警信息等多源数据；其次，设计基于注意力机制的时空图卷积网络（ST-GAT），捕捉攻击行为的时序依赖与节点关联特征；最后，通过攻击路径推理与风险等级量化，形成从攻击溯源到态势评估的闭环。实验基于 CTU-13 和 CSE-CIC-IDS2018 数据集验证，结果表明该框架在攻击溯源准确率（92.7%）、态势评估响应时间（≤0.3 s）等指标上显著优于传统方法、近年主流时序 GNN 变体及网络安全领域专用模型，为网络安全应急响应提供技术支撑。

关键词：图神经网络；网络攻击溯源；态势感知；异构信息网络；时空图卷积
中图分类号：TP393.08　　**文献标志码**：A　　**DOI**：10.19358/j.issn.2097－1788.2026.04.007

中文引用格式：胡开明，陈建华. 一种融合多源异构数据的图神经网络联合框架[J]. 网络安全与数据治理，2026，45(4)：51－58.
英文引用格式：Hu Kaiming，Chen Jianhua. A joint framework of graph neural networks integrating multi-source heterogeneous data[J]. Cyber Security and Data Governance，2026，45(4)：51－58.

A joint framework of graph neural networks integrating multi-source heterogeneous data

Hu Kaiming，Chen Jianhua
（Guangdong Songshan Polytechnic，Shaoguan 512126，China）

Abstract：Aiming at the complex characteristics of cyberspace attack-defense confrontation，such as multi-step，concealed，and heterogeneous，traditional methods relying on rule matching and statistical analysis can hardly meet the needs of accurate traceability and real-time situation awareness. This paper proposes a joint framework of graph neural networks（GNNs）integrating multi-source heterogeneous data to realize automatic traceability of network attacks and dynamic situation awareness. Firstly，a heterogeneous information network（HIN）of network entities-attack behaviors is constructed to integrate multi-source data such as traffic logs，vulnerability databases，and alarm information. Secondly，a spatiotemporal graph attention network（ST-GAT）based on the attention mechanism is designed to capture the temporal dependence of attack behaviors and the correlation characteristics of nodes. Finally，through attack path reasoning and risk level quantification，a closed loop from attack traceability to situation assessment is formed. Experiments are verified based on the CTU-13 and CSE-CIC-IDS2018 datasets. The results show that the framework is significantly superior to traditional methods，the mainstream temporal GNN variants and dedicated models in the field of network security in indicators such as attack traceability accuracy（92.7%）and situation assessment response time（≤0.3 s），providing technical support for network security emergency response.

Key words：Graph Neural Network（GNN）；network attack traceability；situation awareness；Heterogeneous Information Network（HIN）；Spatiotemporal Graph Convolution（SGC）

0　引言

网络攻击溯源与态势感知作为网络安全防御的核心技术，在威胁检测、应急响应、风险管控等关键场景中具有不可替代的应用价值，是保障关键信息基础设施安全、提升主动防御能力的重要支撑。然而，传统攻击溯源方法主要依赖规则匹配（如 Snort 规

*基金项目：广东省普通高校特色创新项目（2023KTSCX269）

则)[1]、日志关联分析和攻击图推理，存在显著技术局限：规则匹配难以应对未知攻击变体，日志分析受数据质量与完整性影响较大，攻击图推理则面临状态空间爆炸问题[2]；态势感知技术多基于统计学习方法(如 SVM、随机森林)[3]，难以捕捉网络实体间的复杂关联关系与攻击行为的时序动态演化特性。随着网络数据规模的指数级增长，多源异构安全数据的融合建模与深层特征提取已成为突破现有技术瓶颈的核心关键。

针对上述问题，现有改进方案多聚焦于规则库迭代、日志预处理优化或攻击图剪枝技术，但这些方法仍未脱离对结构化数据的强依赖，无法有效处理动态变化的攻击场景，且存在智能化程度低、依赖人工干预等固有缺陷。在复杂攻击场景（如高级持续性威胁 APT）中，由于攻击路径隐蔽性强、跨节点协同性高，传统方法难以实现攻击行为的完整溯源与态势的实时评估，导致安全响应延迟，无法满足主动防御需求[4]。

近年来，图神经网络（Graph Neural Network，GNN）作为深度学习领域的重要分支，通过将图结构数据与神经网络深度融合，能够自动学习节点与边的高维特征表示，在节点分类、链路预测、路径推理等任务中展现出优异性能[5]，为解决网络安全领域的复杂关联建模问题提供了新路径。GNN 具备的端到端学习能力与结构表征优势，能够有效突破传统方法的技术局限，实现多源异构数据的有机融合与攻击模式的智能挖掘。

综合上述研究，本文将 GNN 技术引入网络攻击溯源与态势感知领域，构建端到端的联合框架，实现从数据建模、特征学习到路径推理与态势评估的全流程优化。本文的贡献总结如下：

（1）提出一种网络实体 – 攻击行为异构信息网络（EHIN）建模方法，将设备、漏洞、攻击行为等多类型实体抽象为图节点，实体间的关联关系抽象为边，通过规范化公式实现多源异构安全数据的统一表征，解决数据融合难题。

（2）设计融合时序注意力与空间注意力的 ST-GAT 模型，完成时序依赖捕捉、空间关联提取、特征融合与路径推理的数学推导，显著提升复杂攻击路径的识别精度与效率。

（3）构建“溯源 – 评估 – 预测”三位一体的态势感知体系，基于 GNN 推理得到的攻击路径计算攻击成功率、影响范围等核心指标，结合网络资产价值量化安全态势，实现动态风险评估与主动防御支撑。

1 相关工作

1.1 网络攻击溯源技术研究

网络攻击溯源技术可分为基于主机、基于网络和基于混合策略三类。基于主机的方法通过分析主机日志、进程行为等本地数据定位攻击源，如基于系统调用序列的异常检测[4]，但该类方法易受单点攻击规避影响。基于网络的方法依赖流量分析、IP 追踪等技术，如基于数据包标记的溯源方案[6]，但该类方法在复杂网络拓扑中准确率较低。混合策略结合主机与网络数据，如基于贝叶斯网络的攻击路径推理[7]，但该类方法依赖人工构建概率模型，扩展性较差。

近年来，机器学习方法开始应用于攻击溯源，如基于决策树的攻击路径分类[8]、基于 LSTM 的攻击序列识别[9]。但这些方法多将溯源任务转化为分类问题，忽略了攻击行为的关联性与时序性。GNN 的出现为关联型数据建模提供了新思路，如 Wang 等人提出基于 GCN 的攻击图推理方法[10]，提升了路径识别准确率，但未考虑多源数据融合与时序动态特征。

随着时序图神经网络（Temporal GNN，TGNN）的快速发展，诸多适配网络安全场景的时序 GNN 变体相继涌现，为攻击溯源提供了更优的技术路径。PINT[11]作为具有理论保证的时序图网络，扩展了 1-WL 测试到时序图，通过引入相对位置特征和单调时间因果树（TCT）捕获时序依赖，在多个基准数据集上表现优于传统时序模型。网络安全领域专用溯源模型的研究也不断深入。TCG-IDS[12]作为首个基于自监督时序对比图神经网络的入侵检测系统，其核心技术架构同样适用于网络攻击溯源任务，该模型能够在标记数据稀缺、网络数据存在噪声或不完整的场景下，精准学习节点行为表示，进而实现攻击路径的有效识别与溯源。这些专用模型虽能针对性解决网络安全数据的时序性、异构性问题，但在多源数据融合与攻击路径推理的协同性方面仍有提升空间，这也凸显了本文 ST-GAT 模型的创新价值。

1.2 网络态势感知技术研究

网络态势感知（NSA）的概念由 Endsley 于 1988 年提出[13]，核心是“感知 – 理解 – 预测”的闭环过程。现有研究主要集中在态势要素提取、态势评估与态势预测三个环节。态势要素提取方面，传统方法依赖特征工程（如基于 TF-IDF 的日志特征提取）[14]，深度学习方法如 CNN、AutoEncoder 用于流量特征自动提取[15]。态势评估多采用层次化分析（AHP）、模糊综合评价等方法[16]，但权重设定依赖专家经验。态势预测多基于时

序模型[17]，但难以捕捉网络实体间的关联影响。

GNN 在态势感知中的应用尚处于起步阶段，如 Li 等人[18]提出基于 GraphSAGE 的网络态势评估模型，通过节点嵌入表示网络安全状态，但未结合攻击行为的动态变化；Zhao 等人[19]设计时序图卷积网络用于攻击趋势预测，但缺乏与溯源任务的协同。

PINT 通过引入相对位置特征和单调时间因果树（TCT）精准捕捉时序依赖，可有效适配态势感知中攻击行为的动态演化特性，为态势要素的时序提取与趋势预测提供了可靠技术支撑[20]。TCG-IDS[12]利用 TGN 编码器捕捉网络实体交互的时序演变与空间依赖，其核心的时空特征提取能力也可有效适配态势感知中的威胁量化与状态评估。这些研究表明，时空融合的图神经网络或网络安全专用模型虽然能够有效提升态势感知的精度与实时性，但现有模型多未实现溯源与态势感知的深度协同，这也构成了本文 ST-GAT 联合框架的差异化优势所在。

1.3 GNN 在网络安全中的应用

GNN 已在网络安全领域展现出广泛应用前景，除攻击溯源与态势感知外，还包括恶意代码检测、漏洞挖掘、欺诈检测等场景。恶意代码检测中，将函数调用关系建模为图，通过 GCN 实现恶意代码分类[21]；漏洞挖掘中，基于代码抽象语法树（AST）的 GNN 模型能够识别漏洞模式[22]；欺诈检测中，GraphSAGE 用于捕捉用户行为关联，提升欺诈识别率[23]。

现有研究表明，GNN 在处理关联型安全数据方面具有天然优势，但在攻击溯源与态势感知的联合建模、多源异构数据融合、动态时序特征捕捉等方面仍存在不足，本文针对这些问题展开深入研究。

2 系统设计

2.1 系统整体架构

本文提出的基于 GNN 的网络攻击溯源与态势感知系统，整体架构分为数据层、图建模层、模型层、应用层四个部分，如图 1 所示。数据层采集多源安全数据并完成预处理；图建模层构建异构信息网络，将多类型实体与关联关系映射为图结构；模型层基于 ST-GAT 模型实现攻击特征学习、攻击路径推理与态势评估；应用层提供可视化界面与决策支持，形成端到端的技术闭环。

该架构的核心优势在于“模块化设计”与“数据－模型－应用的深度耦合”：模块化设计便于后续功能扩展与技术迭代，数据与模型的耦合确保了特征提取的针对性，模型与应用的耦合则保障了技术输出的实用性，有效解决传统系统“数据处理与建模脱节、建模结果与实际应用脱节”的问题。

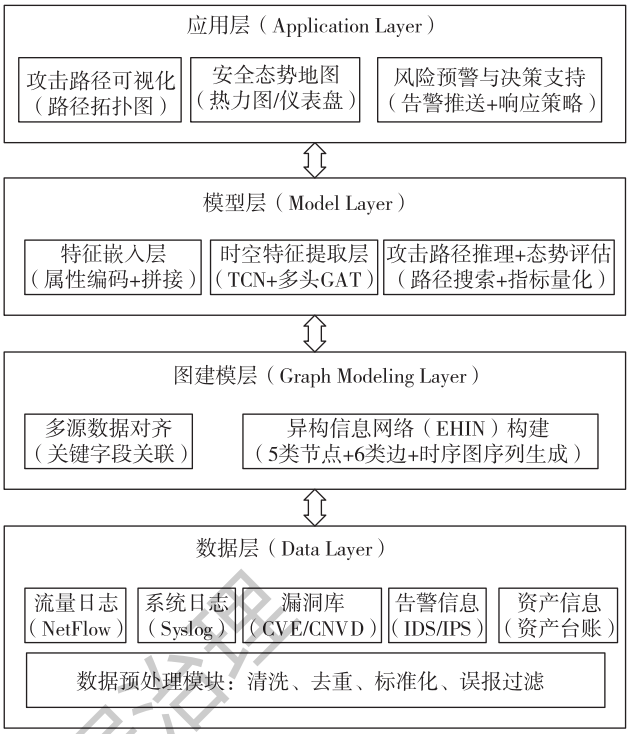

图 1 系统架构

2.2 多源数据融合与图建模

2.2.1 数据类型与预处理

系统采集的多源数据包括网络流量日志、主机系统日志、漏洞库数据（CVE/CNVD）、安全设备告警信息、资产信息等。预处理阶段需完成数据清洗、标准化与关联对齐，其中连续型特征采用 Min-Max 归一化处理，如式（1）所示：

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

其中，x 为原始特征值，$x_{\min}$ 和 $x_{\max}$ 分别为特征的最小值和最大值，x_{norm} 为归一化后的特征值。离散型特征（如攻击类型、设备类型）采用独热编码转化为向量形式，对于具有 k 个取值的离散特征，编码后生成维度为 k 的二进制向量。

2.2.2 异构信息网络构建

异构信息网络（EHIN）定义为 $G = (V, E, T_V, T_E)$，其中 V 为节点集合，E 为边集合，T_V 为节点类型集合，T_E 为边类型集合。本文定义 5 类节点（设备 D、漏洞 V、攻击行为 A、账户 U、进程 P）和 6 类边（D-V、D-U、D-P、A-V、A-D、A-U）。边权重计算采用归一化后的属性值加权方式，以 D-V 边（设备－漏洞关联）为例，权重计算如式（2）所示：

$$wD-V(i,\ j)=\frac{\mathrm{CVSS}\ (j)}{\max_j\in \mathrm{CVSS}\ (j)} \tag{2}$$

其中，CVSS(j)为漏洞 j 的 CVSS 评分，取值范围为 0～10，$wD-V(i,\ j)$ 为设备 i 与漏洞 j 之间的边权重，取值范围为 0～1。为捕捉攻击行为的时序动态性，基于时间窗口划分生成时序图序列 $G=\{G_1,\ G_2,\ \cdots,\ G_T\}$，其中 T 为时间窗口数量，每个时间窗口的长度 Δt 经交叉验证设为 5 min，满足实时性要求。

2.3 基于 ST-GAT 的核心模型

2.3.1 模型整体结构

ST-GAT 模型分为特征嵌入层、时空特征提取层、攻击路径推理层、态势评估层四个模块。ST-GAT 模型的设计目标为精准捕捉攻击行为的时空特征，实现攻击路径的自动推理与态势的动态量化，因此采用“特征嵌入—时空特征提取—路径推理—态势评估”的四级串联结构，各模块层层递进、环环相扣。

2.3.2 特征嵌入层

异构节点的初始特征包含离散属性与连续属性，需通过嵌入层转化为统一维度的低维向量。对于离散节点类型 $t\in T_V$，其嵌入向量通过可学习的嵌入矩阵生成，如式（3）所示：

$$\boldsymbol{h}_v^t=\boldsymbol{W}_t\cdot\boldsymbol{X}_v^t+\boldsymbol{b}_t \tag{3}$$

其中，$\boldsymbol{X}_v^t$ 为节点 v 的离散属性编码向量，$\boldsymbol{W}_t\in\mathbf{R}^{d\times k_t}$ 为类型 t 的嵌入矩阵（d 为嵌入维度，k_t 为类型 t 的属性维度），$\boldsymbol{b}_t\in\mathbf{R}^d$ 为偏置项，$\boldsymbol{h}_v^t$ 为离散属性嵌入向量。

连续属性经归一化后直接拼接至离散属性嵌入向量，形成节点初始特征向量，如式（4）所示：

$$\boldsymbol{h}_v^0=\mathrm{Concat}\ (\boldsymbol{h}_v^t,\ \boldsymbol{X}_v^c) \tag{4}$$

其中，$\boldsymbol{X}_v^c$ 为归一化后的连续属性向量，Concat(·) 表示向量拼接操作，$\boldsymbol{h}_v^0\in\mathbf{R}^{d+m}$（$m$ 为连续属性维度）为节点初始特征向量。

2.3.3 时序特征提取（TCN 模块）

采用因果膨胀卷积捕捉攻击行为的时序依赖，卷积核大小为 k，膨胀率为 d_l（第 l 层），则第 l 层的输出特征如式（5）所示：

$$\boldsymbol{h}_v^T(l)=\sigma\ \left(\sum_{i=0}^{k-1}\boldsymbol{W}_T(l)\cdot\boldsymbol{h}_v^T(l-1)[t-i\cdot d_l]+\boldsymbol{b}_T(l)\right) \tag{5}$$

其中，$\boldsymbol{W}_T(l)\in\mathbf{R}^{d\times d}$ 为第 l 层时序卷积权重矩阵，$\boldsymbol{b}_T(l)\in\mathbf{R}^d$ 为偏置项，$\sigma(\cdot)$ 为 LeakyReLU 激活函数，$[t-i\cdot d_l]$ 表示时序索引偏移，$\boldsymbol{h}_v^T(l)$ 为第 l 层时序特征向量。

TCN 的感受野 R 随网络层数 L 的增长呈指数级扩大，计算公式如式（6）所示：

$$R=1+2\cdot\sum_{l=1}^{L}(k-1)\cdot d_l \tag{6}$$

本文设置 $k=3$，$d_l=2^{l-1}$，$L=3$，则感受野 $R=13$，可覆盖 65 min 内的时序依赖关系，满足复杂攻击行为的时序捕捉需求。

2.3.4 空间特征提取（GAT 模块）

图注意力机制通过计算邻居节点的注意力权重，实现关键关联特征的强化。对于节点 v 的邻居节点 u，注意力系数计算如式（7）所示：

$$e_{vu}=\mathrm{LeakyReLU}(\boldsymbol{a}^T\cdot\mathrm{Concat}(\boldsymbol{W}_S\cdot\boldsymbol{h}_v^T,\ \boldsymbol{W}_S\cdot\boldsymbol{h}_u^T)) \tag{7}$$

其中，$\boldsymbol{W}_S\in\mathbf{R}^{d'\times d}$ 为空间特征转换矩阵，$\boldsymbol{a}\in\mathbf{R}^{2d'}$ 为注意力参数向量，$\boldsymbol{h}_v^T$ 和 $\boldsymbol{h}_u^T$ 分别为节点 v 和 u 的时序特征向量，e_{vu} 为原始注意力系数。

通过 Softmax 函数对注意力系数进行归一化，如式（8）所示：

$$\alpha_{vu}=\frac{\exp(e_{vu})}{\sum_{u'\in N_v}\exp(e_{vu'})} \tag{8}$$

其中，N_v 为节点 v 的邻居节点集合，α_{vu} 为归一化后的注意力系数，满足 $\sum_{u\in N_v}\alpha_{vu}=1$。

采用多头注意力机制提升特征表达能力，设头数为 K，则节点 v 的空间特征向量如式（9）所示：

$$\boldsymbol{h}_v^S=\mathrm{Concat}_{k=1}^{K}\ (\sigma\ (\sum_{u\in N_v}\alpha_{vu}^k\cdot\boldsymbol{W}_S^K\cdot\boldsymbol{h}_u^T)) \tag{9}$$

其中，α_{vu}^k 和 $\boldsymbol{W}_S^K$ 分别为第 k 个头的注意力系数和权重矩阵，$\boldsymbol{h}_v^S\in\mathbf{R}^{K\cdot d'}$ 为空间特征向量。

2.3.5 时空特征融合

采用自适应加权融合策略，根据时序特征与空间特征的方差动态调整融合权重，如式（10）、（11）所示：

$$\lambda=\frac{\mathrm{Var}(\boldsymbol{h}_v^T)}{\mathrm{Var}(\boldsymbol{h}_v^T)+\mathrm{Var}(\boldsymbol{h}_v^S)} \tag{10}$$

$$\boldsymbol{h}_v^{\mathrm{final}}=\lambda\cdot\boldsymbol{h}_v^T+(1-\lambda)\cdot\boldsymbol{h}_v^S \tag{11}$$

其中，Var(·) 表示方差计算，λ 为时序特征权重（取值范围 0～1），$\boldsymbol{h}_v^{\mathrm{final}}\in\mathbf{R}^{d_{\mathrm{final}}}$ 为节点最终增强特征向量（$d_{\mathrm{final}}=\max(d,\ K\cdot d')$）。

2.3.6 攻击路径推理

攻击路径推理本质是图中的最优路径搜索，基于节点最终特征计算边的可信度得分：

$$s_{vu}=\sigma(\boldsymbol{h}_v^{\mathrm{final}}\cdot\boldsymbol{h}_u^{\mathrm{final}^{\mathrm{T}}})+b_s \tag{12}$$

其中，b_s 为偏置项，$s_{vu}\in[0,\ 1]$ 为边（v，u）的可信度得分，得分越高表示该边为攻击路径组成部分的概率越大。

采用改进的 Dijkstra 算法搜索攻击源到攻击目标的最优路径，路径总可信度计算如式（13）所示：

$$S(P)=\prod_{(v,u)\in P}s_{vu} \tag{13}$$

其中，$P=\{v_0, v_1, \cdots, v_n\}$ 为攻击路径（v_0 为攻击源，v_n 为攻击目标），$S(P)$ 为路径总可信度，选择 $S(P)$ 最大的路径作为最终溯源结果。

2.3.7 态势评估指标计算

态势评估的目标是“量化网络安全状态，为防护决策提供依据”，因此指标设计需兼顾攻击威胁的严重性与资产的重要性，形成多维度、可解释的评估体系。

攻击威胁度 T（式（14））从攻击本身的危害性出发，选择攻击类型威胁系数、漏洞威胁系数、攻击成功率三个核心维度：攻击类型威胁系数根据攻击的危害程度设定（如勒索病毒攻击为 5，端口扫描攻击为 1）；漏洞威胁系数由 CVSS 评分归一化得到，直接反映漏洞的危险程度；攻击成功率基于历史攻击数据统计（如利用 CVE-2021-44228 漏洞的攻击成功率为 85%）。三者加权求和能够全面反映攻击的威胁水平。

$$T=w_1\cdot T_A+w_2\cdot T_V+w_3\cdot T_S \tag{14}$$

其中，T_A 为攻击类型威胁系数（根据攻击危害程度设定为 1～5），T_V 为漏洞威胁系数（$T_V=\mathrm{CVSS}/10$），T_S 为攻击成功率（基于历史数据统计），$w_1=0.4$，$w_2=0.3$，$w_3=0.3$ 为权重（通过层次分析法确定）。

资产风险度 R（式（15））聚焦受攻击资产的重要性，资产价值系数根据设备的核心程度设定（如数据库服务器为 5，普通办公电脑为 1）；受攻击设备占比反映攻击的影响范围。二者乘积能够体现攻击对网络整体资产的风险大小。

$$R=V\cdot\frac{N_{\text{affected}}}{N_{\text{total}}} \tag{15}$$

其中，V 为资产价值系数（1～5），N_{affected} 为受攻击设备数量，N_{total} 为网络设备总数。

态势等级 L 通过模糊综合评价法确定，定义 5 个态势等级（安全、基本安全、一般风险、较高风险、严重风险），其隶属度函数如式（16）所示：

$$\mu_L(x)=\begin{cases}1 & x\in[0,\ 0.2)\\ \dfrac{0.4-x}{0.2} & x\in[0.2,\ 0.4)\\ 0 & x\in[0.4,\ 1.0]\end{cases} \tag{16}$$

3 实验验证与结果分析

3.1 实验环境与数据集

3.1.1 实验环境

为确保实验结果的可靠性与可复现性，本次实验采用标准化的硬件与软件配置，硬件选型充分考虑深度学习模型的计算需求：Intel Core i9-12900K CPU，32 GB DDR5 内存，NVIDIA RTX 3090 GPU（24 GB 显存）；软件环境选择当前深度学习与网络安全领域的主流工具栈：Python 3.8，PyTorch 1.12，PyTorch Geometric 2.1，Scikit-learn 1.2，TensorFlow 2.9。

3.1.2 数据集

采用公开网络安全数据集 CTU-13 和 CSE-CIC-IDS2018，同时补充模拟生成的 APT 攻击数据集。CTU-13 包含 13 个恶意软件攻击场景的流量数据，数据量约 10 GB，记录攻击源、攻击路径、攻击目标等信息；CSE-CIC-IDS2018 包含 DDoS、暴力破解、SQL 注入等 14 种攻击类型的流量与日志数据，数据量约 100 GB；模拟 APT 数据集，基于 NS-3 网络仿真工具搭建企业网络拓扑，模拟多步骤 APT 攻击，生成包含漏洞利用、权限提升、数据窃取等行为的日志与流量数据，数据量约 5 GB。数据集按照 7∶2∶1 的比例划分为训练集、验证集、测试集，训练集用于模型训练，验证集用于超参数调优，测试集用于性能评估。

3.2 评价指标

3.2.1 攻击溯源评价指标

攻击溯源任务的核心挑战在于“攻击数据稀缺性”与“路径识别的准确性”，因此评价指标需兼顾“识别精准度”与“召回完整性”，同时考虑实时性要求。

（1）准确率（Precision）：聚焦识别结果的可靠性，避免模型将正常网络链路误判为攻击路径，这对安全运维人员的决策至关重要——误判的溯源结果会导致运维资源浪费，甚至误导应急响应方向。

（2）召回率（Recall）：聚焦攻击路径的完整性，确保模型不遗漏真实攻击路径，尤其是复杂攻击中的中间跳板节点，遗漏关键节点会导致溯源中断，无法定位攻击源。

（3）F1-score：由于攻击数据（正样本）在整体数据中占比极低（通常不足 5%），准确率与召回率存在此消彼长的关系（如提高准确率可能导致召回率下降），因此采用 F1-score 作为综合评价指标，平衡二者的性能。

（4）溯源耗时：网络攻击的应急响应具有极强的时效性（如勒索病毒攻击需在数据加密前阻断），因此溯源耗时直接决定模型的实际应用价值，要求模型在保证精度的同时，满足实时响应需求（通常≤1 s）。

3.2.2 态势感知评价指标

态势感知任务的核心是“量化网络安全状态的准

确性”与“响应的实时性”，因此评价指标设计围绕“评估精度”与“实时性”展开。

（1）态势评估准确率：态势等级评估结果与真实等级的匹配率。

（2）平均绝对误差（MAE）：威胁度、风险度预测值与真实值的平均绝对偏差。

（3）响应时间：从攻击发生到态势评估结果输出的时间。

3.3 对比实验设计

对比实验的设计目标是全面验证 ST-GAT 模型的优越性，通过选择不同技术路线的代表性方法作为基线，从“传统方法→机器学习方法→基础 GNN 方法”的梯度对比，层层递进验证本文模型的技术创新价值，设置以下对比方法验证模型优越性，所有对比方法均在相同实验环境下运行，超参数通过网格搜索优化。

（1）传统方法：基于规则匹配的攻击溯源（Snort 规则）+层次化分析（AHP）的态势评估。

（2）机器学习方法：LSTM（时序特征）+随机森林（溯源与态势评估）。

（3）基础 GNN 方法：GCN（图卷积网络）、GAT（图注意力网络）、GraphSAGE（归纳式 GNN）。

（4）时序 GNN 变体：PINT[11]，具有理论保证的时序图网络，时序依赖捕捉能力优异。

（5）网络安全专用模型：TCG-IDS[12]，自监督时序对比图神经网络，入侵检测精度高。

3.4 实验结果与分析

3.4.1 攻击溯源性能对比

从表 1 的攻击溯源性能对比结果可以看出，ST-GAT 模型在各项溯源指标上均优于对比方法，且优势具有显著性。

（1）准确率、召回率、F1-score 分别比传统方法提升 20.4%、22.8%、21.7%，得益于 GNN 对关联特征的有效捕捉。

（2）相较于基础 GNN 方法，ST-GAT 融合时序特征与注意力机制，F1-score 提升 3.8% ~7.6%，能够更好地捕捉攻击行为的动态变化与关键关联。

（3）相较于新增的时序 GNN 变体，ST-GAT 在 F1-score 上比 PINT 提升 3.5%；溯源耗时方面，比 PINT（0.4 s）体现出更优的实时性。

（4）相较于网络安全专用模型，ST-GAT 的 F1-score 比 TCG-IDS 提升 1.6%，溯源耗时与专用模型相比处于较优水平。

表 1 攻击溯源性能对比结果

方法	准确率/%	召回率/%	F1-score/%	溯源耗时/s
传统方法	72.3	68.5	70.3	1.8
LSTM + 随机森林	81.5	79.2	80.3	0.9
GCN	85.7	83.1	84.4	0.6
GAT	88.9	86.5	87.7	0.5
GraphSAGE	89.3	87.2	88.2	0.4
PINT	89.2	87.8	88.5	0.4
TCG-IDS	91.1	89.7	90.4	0.35
ST-GAT	92.7	91.3	92	0.3

3.4.2 态势感知性能对比

表 2 的态势感知性能对比结果表明，ST-GAT 模型在态势评估准确率、MAE、响应时间三个指标上均表现最优：

（1）态势评估准确率达到 94.3%，比传统方法提升 18.7%，比基础 GNN 方法提升 4.2% ~7.5%。

（2）MAE（威胁度、风险度）均不高于 0.05，说明态势评估结果与真实值偏差较小。

（3）相较于新增的时序 GNN 变体和网络安全专用模型，评估准确率比 PINT 提升 3.6%、比 TCG-IDS 提升 1.6%。

（4）响应时间仅为 0.2 s，满足实时态势感知需求。

表 2 态势感知性能对比结果

方法	评估准确率/%	MAE（威胁度）	MAE（风险度）	响应时间/s
传统方法	75.6	0.18	0.21	2.3
LSTM + 随机森林	83.2	0.12	0.15	1.1
GCN	86.8	0.09	0.11	0.7
GAT	89.5	0.07	0.09	0.6
GraphSAGE	90.1	0.06	0.08	0.5
PINT	90.7	0.07	0.08	0.6
TCG-IDS	92.7	0.05	0.06	0.3
ST-GAT	94.3	0.04	0.05	0.2

3.4.3 消融实验分析

消融实验通过逐一移除 ST-GAT 模型的核心模块（时序卷积、注意力机制、异构建模），验证每个模块的独立贡献，消融实验结果（如表 3 所示）表明：

（1）去除时序卷积后，F1-score 下降 4.3%，评估准确率下降 4.8%，说明时序特征对捕捉攻击行为的动态变化至关重要。

（2）去除注意力机制后，性能下降更为明显（F1-score 下降 6.8%，评估准确率下降 7.2%），验证了注

意力机制能够有效突出关键节点与关联的作用。

（3）去除异构建模后，性能有所下降，说明多源数据的异构表征能够为模型提供更丰富的信息。

表 3　消融实验结果（F1-score/评估准确率）

模型配置	攻击溯源 F1-score/%	态势评估准确率/%
完整模型（ST-GAT）	92	94.3
去除时序卷积（GAT）	87.7	89.5
去除注意力机制（ST-CNN）	85.2	87.1
去除异构建模（同构图 ST-GAT）	88.5	90.2

4　结束语

本文针对网络攻击溯源与态势感知面临的多源数据异构、攻击行为关联复杂、实时性要求高等挑战，提出一种基于时空图注意力网络（ST-GAT）的联合框架。通过构建异构信息网络实现多源安全数据的统一表征，利用 ST-GAT 模型完成时序特征捕捉、空间关联提取、路径推理与态势量化的完整推导，形成攻击溯源与态势感知的有机融合。实验结果表明，该框架在攻击溯源准确率（92.7%）、态势评估准确率（94.3%）、响应时间（≤0.3 s）等指标上显著优于传统方法、基础 GNN 方法、近年主流的时序 GNN 变体和网络安全领域的专用模型，能够为网络安全应急响应提供精准、实时的技术支撑。但是模型训练依赖大量标注数据，而真实网络环境中攻击数据的标注成本较高，小样本学习能力有待提升，下一步的改进是结合元学习、对比学习等技术，提升模型在少量标注数据下的性能。

参考文献

[1] ROESCH M. Snort：lightweight intrusion detection for networks[C]//Proceedings of the 13th Conference on Systems Administration（LISA－99），1999.

[2] SHEYNER O，HAINES J，JHA S，et al. Automated generation and analysis of attack graphs[C]//2002 IEEE Symposium on Security and Privacy. IEEE，2002：273－284.

[3] ZHOU X，HU X，PEI D. Network security situation awareness based on SVM and random forest[C]//2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics（IHMSC）. IEEE，2018，1：309－312.

[4] SCHULTZ M G，ESKIN E，ZADOK F，et al. Data mining methods for detection of new malicious executables[C]//2001 IEEE Symposium on Security and Privacy. IEEE，2001：38－49.

[5] VELICKOVIC P，CUCURULL G，CASANOVA A，et al. Graph attention networks[J]. arXiv preprint arXiv：1710.10903，2017.

[6] SAVAGE S，WETHERALL D，KARLIN A，et al. Practical network support for IP traceback：AU3813401A[P]. 2001－08－20.

[7] PEARL J. Probabilistic reasoning in intelligent systems：networks of plausible inference[M]. San Francisco：Morgan Kaufmann Publishers，1988.

[8] LIU Y，CHEN X，ZHANG H，et al. Attack path classification based on improved decision tree algorithm for network forensics[J]. IEEE Access，2019，7：123645－123654.

[9] ZHANG S，WANG Y，LI X，et al. Attack sequence recognition using LSTM neural network for cyber security situational awareness[J]. Neural Computing and Applications，2020，32（15）：11607－11618.

[10] WANG L，LI J，WANG X，et al. Attack graph reasoning based on graph convolutional networks for cyber threat intelligence[J]. Computer Networks，2021，197：108285.

[11] SOUZA A H，MESQUITA D，KASKI S，et al. Provably expressive temporal graph networks[C]//Advances in Neural Information Processing Systems（NeurIPS），2022.

[12] WU C，SUN J，CHEN J，et al. TCG-IDS：robust network intrusion detection via temporal contrastive graph learning[J]. IEEE Transactions on Information Forensics and Security，2025，20：1475－1486.

[13] ENDSLEY M R. Design and evaluation for situation awareness enhancement[J]. Human Factors，1988，30（1）：97－110.

[14] GOODFELLOW I，BENGIO Y，COURVILLE A. Deep learning[M]. Cambridge：MIT Press，2016.

[15] LI H，ZHANG W，LIU Z，et al. Traffic feature extraction based on CNN-AutoEncoder for network anomaly detection[J]. Journal of Network and Computer Applications，2022，201：103287.

[16] CHEN L，WANG Z，LI Y，et al. Fuzzy comprehensive evaluation for network security situational assessment based on AHP[C]//2020 Chinese Control and Decision Conference（CCDC）. IEEE，2020：3840－3845.

[17] BOX G E P，JENKINS G M，REINSEL G C. Time series analysis：forecasting and control[M]. Hoboken：Wiley，2015.

[18] LI S，ZHAO H，LIU J，et al. Network situational assessment model based on GraphSAGE and attention mechanism[J]. IEEE Transactions on Network and Service Management，2023，20（2）：1890－1903.

[19] ZHAO J，LIU C，LI S，et al. Temporal graph convolutional network for cyber attack trend prediction[C]//2022 International Conference on Cyber Security and Protection of Digital Services（Cyber Security）. IEEE，2022：1－6.

[20] ZHANG Y，LIU Z，CHEN F，et al. TodyNet：temporal dynamic graph neural network for multivariate time series classification[J].

IEEE Transactions on Neural Networks and Learning Systems, 2023, 34 (7): 4215 - 4227.

[21] XU J, LIU Y, CHEN X, et al. Malware classification with graph convolutional networks [C]//2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019: 7284 - 7288.

[22] KIM J, PARK S, LEE H. Heterogeneous graph neural networks for multi-source security data fusion[C]//2022 IEEE INFOCOM IEEE Conference on Computer Communications. IEEE, 2022: 1 - 10.

[23] DIJKSTRA E W. A note on two problems in connexion with graphs [J]. Numerische Mathematik, 1959, 1 (1): 269 - 271.

（收稿日期：2026 - 01 - 14）

作者简介：

胡开明（1974 - ），男，硕士，副教授，主要研究方向：信息安全与大数据处理。

陈建华（1979 - ），通信作者，女，硕士，讲师，主要研究方向：大数据挖掘与分析。E-mail：52952497@ qq. com。

“密码前沿技术与应用”主题专栏征稿启事

随着量子计算技术的快速发展、人工智能与大模型应用的广泛渗透，以及数据要素市场化对隐私保护的迫切需求，密码学作为网络空间安全的核心支撑，正迎来前所未有的机遇与挑战。一方面，能够抵抗量子计算攻击的后量子密码算法正加速从理论走向标准化与工程化部署，另一方面，数字经济的深化催生了隐私计算技术的爆发，全同态加密、安全多方计算等展现出巨大应用潜力，与此同时，人工智能与大模型的崛起，既为密码分析提供了新的范式，也对密码技术保护 AI 安全提出了新的需求。为了集中展示我国在密码前沿领域的创新成果，促进学术交流与产学研深度融合，《网络安全与数据治理》拟在 2026 年第 8 期推出“密码前沿技术与应用”主题专栏。现诚邀相关领域的专家学者、科研人员踊跃投稿。

一、征文主题：密码前沿技术与应用

包括但不限于以下学术方向：

1. 抗量子计算密码：后量子密码算法的设计、分析、侧信道防护及标准化迁移应用。

2. AI 与密码学融合：利用 AI 进行密码算法辅助设计/自动化分析；密码技术在 AI 模型安全中的应用。

3. 前沿隐私计算技术：全同态加密、安全多方计算、零知识证明、差分隐私等。

4. 轻量级密码与物联网安全：面向资源受限设备的密码算法及物联网安全协议。

5. 数字身份与零信任安全：抗量子身份认证、分布式数字身份、零信任架构。

6. 可信数据空间密码技术：面向可信数据空间的密码技术与应用。

7. 新型密码分析与评估：量子环境下的密码分析、自动化安全评估平台。

二、投稿要求

1. 稿件请用 word 格式录入，并套用本刊模板（http：//files. chinaaet. com/files/Periodical/pcachina_Templates. doc）。

2. 投稿文章不存在一稿多投现象。

3. 无抄袭、剽窃、侵权、虚假引用等不良学术行为，且不违反相关法律法规，不涉及国家、企业秘密，稿件文责自负。

4. 论文要求观点鲜明、逻辑严谨、论据充分、方法合理，字数在 5000 ~ 8000 字。

5. 请在官方投稿网站（http：//www. pcachina. com）注册、投稿。投稿请选“主题专栏”栏目，“稿件标题”填写“密码技术 + 文章题目”。稿件经评审合格录用后，在《网络安全与数据治理》2026 年第 8 期（正刊）以主题专栏形式发表。

三、专栏主编

祝烈煌，北京理工大学网络空间安全学院党委书记，特聘教授，博士生导师，入选国家高层次领军人才，国家重点研发计划首席科学家。长期从事网络与信息安全方向的教学与研究工作，主持国家重点研发计划项目、国家自然科学基金重点项目等国家级、省部级科研项目 50 余项。近年来，出版外文专著 5 本，发表高水平学术论文 400 余篇。

姜伟，哈尔滨工业大学网络空间安全学院院长兼计算学部副主任，二级教授，博士生导师，国家高层次人才。长期从事网络安全、数据安全、人工智能安全、认知安全等研究。主持国家自然基金重点项目、国家重点研发计划课题、国家社科基金重大项目、中宣部、中央网信办、中联部、教育部等省部级以上课题 30 余项。发表学术论文 50 余篇，参与国家标准编制 10 余项。

杨建，三未信安总经理助理兼市场总监，高级工程师，中国通信学会网络与数据安全委员会委员，黑龙江计算机学会网络空间安全委员会执行委员，江苏省大数据交易和流通工程实验室数据安全专委会专家委员，抗量子密码技术与应用北京市重点实验室学术带头人，沈阳航空航天大学兼职硕士生导师。参与省部级科技进步一等奖 3 项，发表论文 5 篇，多次参与国家重点研发计划和试点示范工程。

四、时间安排

截稿日期：2026 年 6 月 30 日

审稿反馈日期：2026 年 7 月 15 日

出版日期：2026 年 8 月 15 日

《网络安全与数据治理》编辑部

2026 年 2 月

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