

专题：低空智联网

基于GCN-Transformer的空管系统网络威胁实体识别方法

刘超[1]，田继伟[2]，贾楠[1]，张欢乐[1]，王布宏[2]，刘洲洲[1]
（1. 西安航空学院计算机与人工智能学院，陕西 西安 710077；
2. 中国人民解放军空军工程大学，陕西 西安 710051）

摘　要：随着空中交通管理（air traffic management，ATM）系统数字化转型持续推进，其网络安全风险不断加剧，威胁实体识别已成为构建网络安全知识图谱的核心环节。针对传统序列标注模型难以有效捕捉文本语法结构与语义依赖的不足，提出一种融合图卷积网络（graph convolutional network，GCN）与Transformer的ATM系统网络威胁实体识别方法。该方法依托语法解析与词共现关系构建文本图，利用节点重要性排序与广度优先搜索生成语义子图并完成标准化处理；设计基于层次正则化的深度图卷积网络（hierarchical regularized deep graph convolutional neural network，HR-DGCN），引入父子标签相似性约束，增强模型对层次化威胁实体的识别能力；借助Transformer编码器提取全局上下文语义特征，结合条件随机场（conditional random field，CRF）实现最优实体序列解码。基于自建ATM威胁语料库开展对比实验，结果表明，所提方法的F1分数、精确率、召回率均优于BiLSTM-CRF、BERT-Tagging等基线模型。该方法为ATM系统网络威胁实体识别提供了新的解决方案，具备理论与应用价值。

关键词：空中交通管理；网络安全；威胁实体识别；图卷积网络；Transformer；知识图谱
中图分类号：TP393.08；V355.1
文献标志码：A
doi: 10.11959/j.issn.1000−0801.DXKX260210

A GCN-Transformer method for network threat entity recognition in air traffic management systems

Liu Chao[1], Tian Jiwei[2], Jia Nan[1], Zhang Huanle[1], Wang Buhong[2], Liu Zhouzhou[1]
1. School of Computer Science, Xi'an Aeronautical University, Xi'an 710077, China
2. Air Force Engineering University, Xi'an 710051, China

Abstract: With the continuous advancement of digital transformation in air traffic management (ATM) systems, their cybersecurity risks have been increasingly intensifying, and threat entity recognition has become a core component in constructing cybersecurity knowledge graphs. To address the deficiency that traditional sequence labeling models struggled to effectively capture textual syntactic structures and semantic dependencies, a hybrid method integrating

收稿日期：2026−04−01；修回日期：2026−05−26
通信作者：田继伟，tianjiwei2016@163.com
基金项目：国家自然科学基金资助项目（No.62472437）
Foundation Item: The National Natural Science Foundation of China (No. 62472437)

graph convolutional network (GCN) and Transformer was proposed for ATM network threat entity recognition. Specifically, text graphs were constructed based on syntactic parsing and word co-occurrence relationships, and semantic subgraphs were generated and normalized through node importance ranking and breadth-first search. A hierarchical regularized deep graph convolutional neural network (HR-DGCNN) was designed, in which parent-child label similarity constraints were introduced to enhance the model's capability in identifying hierarchical threat entities. Meanwhile, global contextual semantic features were extracted by the Transformer encoder, and optimal entity sequence decoding was accomplished in combination with conditional random field (CRF). Comparative experiments were conducted on a self-constructed ATM threat corpus, and the results demonstrated that the proposed method outperformed baseline models such as BiLSTM-CRF and BERT-Tagging in terms of F1 score, precision, and recall. This method provides a novel solution for ATM system cyber threat entity recognition and possesses both theoretical and practical value.

Key words: air traffic management, cybersecurity, threat entity recognition, graph convolutional network, Transformer, knowledge graph

0　引言

空中交通管理（air traffic management，ATM）系统作为国家关键信息基础设施的重要组成部分，承担着保障航空运输安全与高效运行的核心使命[1-3]。近年来，随着新一代通信、导航、监视（communications, navigation and surveillance，CNS）技术与ATM系统的持续深度融合，ATM系统正逐步朝着以系统广域信息管理（system wide information management，SWIM）、基于性能的导航（performance-based navigation，PBN）、自动相关监视–广播（automatic dependent surveillance-broadcast，ADS-B）等为代表的网络化、协同化、智能化发展方向[4]。CNS技术与ATM系统融合示意图如图1所示。然而，这一技术演进在提升空管（空中交通管理的简称）效能的同时，也显著扩大了系统网络的攻击面，使其面临日益严峻的网络安全威胁[5]。据中国民用航空局统计，2023年航空业遭遇的网络攻击事件同比增长53%，频繁的攻击行为导致大量航空数据和旅客信息泄露[6]。此外，全球法律顾问公司HFW与航运网络安全公司CyberOwl联合发布的报告指出，2023年航运业关联式网络攻击及相关勒索支出显著上升，平均网络攻击成本由2022年的182 000美元上升至550 000美元，平均勒索支付金额达到320万美元[7]。上述事实充分表明，网络安全已成为制约空管系统可靠运行的关键因素，传统基于边界防护的安全架构已难以应对当前高度复杂与动态演化的威胁环境。

在此背景下，构建空管系统网络威胁知识图谱成为实现网络态势感知、威胁预警与主动防御的有效路径[8]。知识图谱可将碎片化的威胁情报（如漏洞信息、攻击指标、威胁组织等）系统化为结构化知识，支持语义检索、关联分析与逻辑推理。其中，威胁实体识别作为知识图谱构建的关键基础环节，旨在从海量、多源、异构数据（如安全通告、漏洞报告、网络日志等）中自动抽取恶意软件、攻击技术、漏洞编号和威胁组织等安全相关实体[9]。实现高精度、高鲁棒性的威胁实体识别，对于提升空管系统网络安全防护的智能化水平具有重要意义。

早期的威胁实体识别方法主要基于规则与词典[10]，依托人工构建的匹配规则完成威胁实体的抽取。此类方法虽在特定场景下可取得较高识别精度，但高度依赖人工领域先验知识，面对大规模数据及新型威胁时，普遍存在召回率低、模型

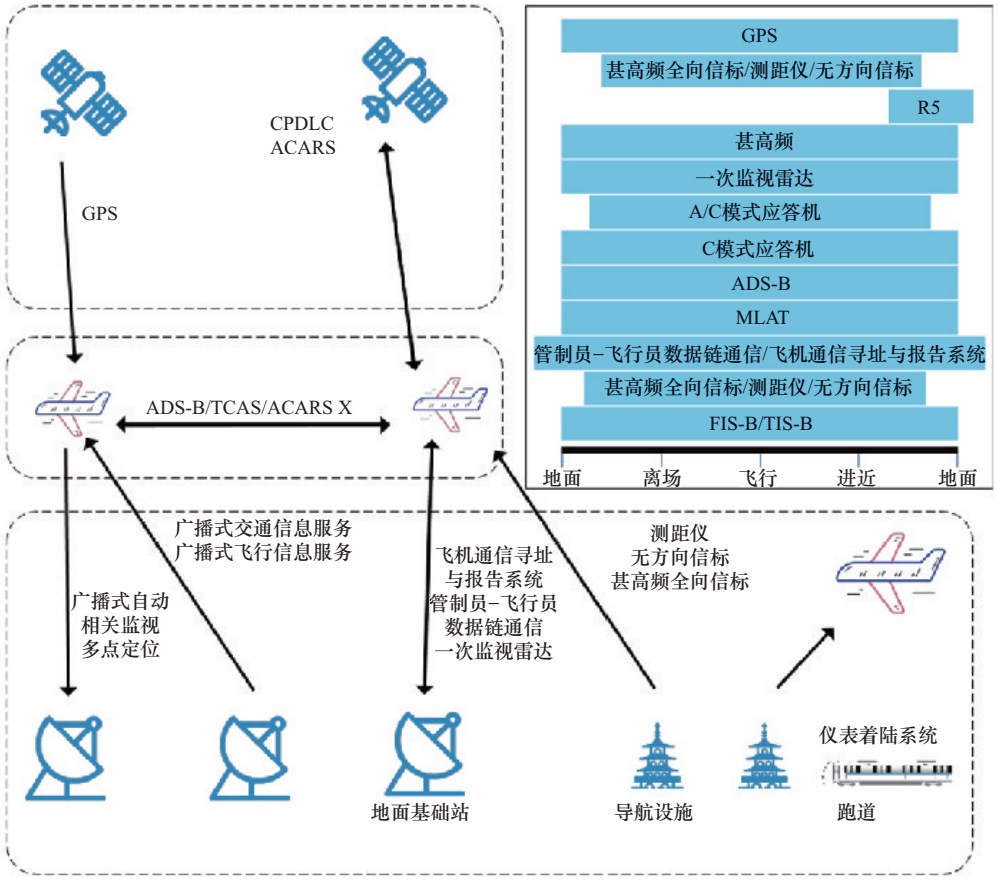

图1　CNS技术与ATM系统融合示意图

泛化能力差等局限。随着统计学习方法的发展，条件随机场（conditional random field，CRF）[11]、支持向量机（support vector machine，SVM）[12]等机器学习模型借助特征工程在一定程度上提升了泛化性能，但仍高度依赖特征设计的质量与全面性。近年来，深度学习的序列标注方法在威胁实体识别任务中取得显著突破，尤其是BiLSTM-CRF模型[13]，凭借双向长短期记忆（bidirectional long short-term memory，BiLSTM）网络在上下文建模中的优势与CRF对标签依赖约束的建模能力，已成为威胁实体识别的主流基线架构。预训练语言模型（如BERT[14]）的引入，使模型能够基于大规模语料获得通用语言表示，从而显著提升威胁实体识别的泛化能力。

尽管深度学习方法已在多种实体识别任务中取得较好效果，但在面向空管场景的威胁实体识别中仍面临两方面核心挑战。一方面，语法结构建模不足。当前主流方法大多将文本处理为线性序列，难以有效利用语法依赖关系这类结构化语言信息。例如，在“攻击者利用CVE-2023-1234漏洞入侵雷达系统”这一描述中，“利用”与“漏洞”之间的动宾结构是识别“CVE-2023-1234”为漏洞实体的重要线索，但线性序列模型难以显式捕捉此类结构化语义。另一方面，长程语义依赖建模能力有限。空管威胁文本中往往存在跨句乃至跨段的复杂语义关联，如威胁组织与其攻击行为之间的关系往往分布在不同上下文中。传统循环神经网络（recurrent neural network，RNN）模型在长距离依赖建模中易受梯度消失或爆炸问题影响，卷积神经网络（convolutional neural network，CNN）模型则受限于固定感受野，均难以有效捕捉此类全局语义联系。

为应对上述挑战，本文提出了一种融合图卷积网络（graph convolutional network，GCN）与Transformer[15]架构的空管威胁实体识别方法。GCN[16]具备对图结构数据的建模能力，可通过消息传递机制聚合语法邻接信息，从而实现对依存结构的有效建模。GCN将传统卷积操作推广至图空间，为文本的结构化表示提供了新范式。Transformer模型[17]基于自注意力机制，能够捕捉输入序列中任意位置之间的依赖，在长程语义建模方面具有显著优势。具体而言，本文的主要贡献如下。

（1）提出一种面向空管威胁文本的图结构构建与子图归一化方法。针对威胁文本的语法特性，设计一种融合词共现关系与句法依赖关系的异构图构建方法，并提出基于节点重要性的子图生成算法和基于广度优先搜索（breadth-first search，BFS）的归一化策略，实现非结构化文本向规范化图张量的有效转换。

（2）设计一种基于层次正则化的深度图卷积网络（hierarchical regularized deep graph convolutional neural network，HR-DGCN）。针对威胁实体固有的层次化分类体系（如“恶意软件–勒索软件–WannaCry”），在损失函数中引入层次正则化项，通过对父子标签分类器参数的相似性进行约束，提升对稀有实体和层次化实体的识别能力。

（3）构建GCN-Transformer混合架构。融合GCN在局部语法结构建模中的优势与Transformer对全局语义的建模能力，实现语法与语义信息的协同学习。利用GCN提取基于图结构的节点表示，再由Transformer捕获长距离依赖关系，最终结合CRF实现最优实体序列的解码。

1　相关工作

1.1　威胁实体识别技术

威胁实体识别作为网络安全知识图谱构建的关键技术，其发展经历了从规则驱动到数据驱动的显著演变。早期研究普遍采用基于规则和词典方法[18]，依靠人为设计匹配规则完成各类威胁实体的抽取工作。这类方法在小规模、特定领域数据上表现出较高准确率，但也存在明显局限性：一方面，匹配规则的设计高度依赖领域专家知识，规则搭建与维护成本高昂；另一方面，模型泛化能力有限，难以适应新型威胁和多样化的表达方式。

随着统计机器学习技术的发展，CRF、SVM等方法被广泛应用于实体识别任务。这些方法借助精心设计的特征工程（如词形特征、上下文特征、词典特征等），有效提升了模型性能。然而，特征工程的质量仍然严重依赖领域知识，且难以捕捉深层次的语义信息。

近年来，基于深度学习的实体识别方法取得了突破性进展。与传统方法相比，深度学习模型具备以下优势：一是，借助深度神经网络强大的非线性建模能力，深度学习模型能够自动学习实体特征的分布式表示，避免了烦琐的人工特征工程；二是，采用端到端的学习方式，大幅降低了对领域知识的依赖和人工干预成本；三是，深度学习模型具备处理深层语义组合的能力，能够识别复杂的实体表达模式[18]。

在深度学习模型架构方面，研究者已提出多种创新方案。Kong等[19]开创性地将CNN应用于序列标注任务，为神经网络在自然语言处理中的应用奠定了基础。与此同时，曹依依等[20]提出CNN-CRF混合模型，采用CNN提取局部特征，并结合CRF的序列优化能力，显著提升了实体识别的准确率与处理效率。杨红梅等[13]提出的BiLSTM-CRF模型进一步推动了领域发展，该模型结合了BiLSTM的上下文特征提取能力和CRF的标签转移约束，成为实体识别领域的基准模型。

预训练语言模型的兴起标志着实体识别技术进入新阶段。Meng等[14]提出的BERT模型通过

在大规模语料上进行掩码语言建模和下一句预测预训练，获得了丰富的语言表示知识，在下游实体识别任务上微调后表现出色。这种“预训练–微调”范式显著提升了模型的泛化能力和性能上限。

特别值得关注的是，基于Transformer的实体识别方法近年来成为研究热点。Transformer模型凭借其自注意力机制，能够直接捕捉序列中任意位置间的依赖关系，克服了传统RNN在长程依赖建模方面的局限性。在代表性研究中，李博等[21]首次将Transformer与CRF相结合，提出Transformer-CRF模型。其中，Transformer层负责学习深层次的上下文表示，CRF层则建模标签间的转移约束，两者协同，显著提升了模型的实体识别性能。大量实证研究表明[22]，基于自注意力机制的模型不仅能有效提高识别精度，还可大幅缩短训练时间，成为实体识别技术的重要发展方向。

然而，现有方法在应用于空管系统网络安全领域时仍面临显著挑战。一方面，大多数方法将文本视为线性序列进行处理，未能充分利用文本中丰富的图结构信息（如语法依赖关系、语义关联网络等）；另一方面，现有研究普遍忽略了威胁实体本身具有的层次化分类特性（如“恶意软件–勒索软件–WannaCry”的多层结构），导致对复杂实体关系的建模能力不足。这些局限性促使本研究探索GCN与Transformer的融合架构，以更好地适应空管系统威胁实体识别的特殊需求。

1.2 图神经网络与Transformer模型

图神经网络（graph neural network，GNN）[23]作为一种专门处理图结构数据的深度学习框架，通过消息传递机制实现节点间信息的有效聚合。这一特性使其特别适合处理具有丰富结构关系的自然语言处理任务，如语法依赖分析、语义角色标注等。Zhang等[24]提出的GCN采用基于谱图理论的方法，通过图傅里叶变换在谱域定义卷积操作，为处理非欧几里得数据提供了理论依据。Wang等[25]进一步提出了动态图卷积网络，通过设计可学习的卷积滤波器实现了对图数据的高效特征提取。

与此同时，Transformer模型凭借其完全基于自注意力机制的架构，在序列建模任务中展现出卓越性能。自注意力机制的核心优势在于能够直接计算序列中任意两个元素间的关联强度，从而有效捕捉长程依赖关系。Shaw等[26]提出的Transformer架构正是基于这种自注意力机制，其采用的缩放点积注意力（scaled dot-product attention）通过引入缩放因子$\frac{1}{\sqrt{d_k}}$对点积结果进行调节，有效避免了内积过大导致的梯度趋于零问题，从而保障了长程依赖关系的稳定建模。

近年来，研究者开始探索GNN与Transformer的融合架构，以期结合两者优势。现有融合方法主要可分为以下三类：（1）序列到图的转换方法。这类方法先将文本序列转换为图结构，再应用GNN进行处理。例如，Jiang等[27]提出的TextGCN模型，先构建文本图（单词作为节点，词共现关系作为边），再使用GCN进行文档分类。然而，这类方法通常需要预先定义图结构，且难以处理动态变化的语义关系。（2）图增强的Transformer。该类方法尝试在Transformer中引入图结构信息，例如，Jin等[28]提出了GraphTransformer，通过在注意力计算中融入图结构偏置，使模型能够同时考虑序列顺序和图拓扑关系。这类方法的优势在于保持了Transformer的并行计算效率，同时引入了结构化归纳偏置。（3）多模态融合架构。这是一种更先进的思路，即设计并行或串行的多模态架构，使GNN和Transformer能够协同工作。例如，Zhu等[29]提出的Multi-Modal Fusion Network通过图卷积层提取结构特征，再与Transformer的序列特征进行融合，在关系抽取任务上取得了当前最优的性能。

尽管已有研究证明了GNN与Transformer融合的潜力，但在空管系统威胁实体识别这一特定领域仍存在明显不足：（1）现有方法大多针对通用领域设计，未能充分考虑空管系统网络安全文本的特殊性，如大量的专业术语、缩写词和特定的表达习惯；（2）缺乏对威胁实体层次化特性的专门建模机制；（3）针对空管系统这一特定关键基础设施的优化研究尚属空白。

针对上述问题，本文提出了一种面向空管系统网络安全的GCN-Transformer混合架构。与现有方法相比，本文的创新主要体现在3个方面：第一，设计了专门针对威胁文本的图构建方法，同时考虑词共现关系和语法依赖关系；第二，引入了层次正则化机制，显式建模威胁实体的分类层次结构；第三，优化了GCN与Transformer的融合策略，实现了局部结构特征与全局语义上下文的高效协同。通过这些创新，本文的方法能够更好地适应空管系统威胁实体识别的特殊需求，为关键信息基础设施安全防护提供技术支撑。

2　基于GCN-Transformer的威胁实体识别方法

空管系统网络安全文本信息通常来源于安全通告、漏洞报告、威胁情报、网络日志和告警描述等，原始形态多表现为自然语言文本或半结构化记录，本质上属于非结构化数据。此类文本虽然缺乏显式的数据表结构，但其内部蕴含着较为稳定的实体关联、语法依赖和语义共现关系，具备转化为图结构表示的显著特征。具体而言，空管网络安全文本中通常包含漏洞编号、攻击技术、恶意软件、IP地址、统一资源定位符（uniform resource locator，URL）、哈希值、受影响系统和攻击行为等关键实体，这些实体之间往往通过“利用”“攻击”“连接”“下载”“影响”等谓词或上下文共现形成明确的关系链。例如，在“攻击者利用通用漏洞披露（common vulnerabilities and exposures，CVE）漏洞入侵雷达系统”这一类描述中，“攻击者”“漏洞”“入侵行为”“雷达系统”之间并非孤立存在，而是通过句法依存关系和语义作用关系构成可建模的关联结构。因此，可将词语或实体视为图中的节点，将词共现关系、句法依赖关系，以及潜在语义关联视为边，从而将非结构化文本转化为具有拓扑结构的文本图。该转化方式能够显式保留空管网络威胁文本中的局部语法结构和实体间依赖关系，为后续GCN进行结构特征聚合提供基础。

融合GCN与Transformer的空管系统网络威胁实体识别方法的整体框架如图2所示。该方法首先对输入的空管系统网络安全文本进行预处理和解析，构建包含语法结构与语义关系的文本图；随后，采用基于节点重要性的子图生成算法提取关键语义子图，并经过归一化处理形成规范化的图张量输入；接着，采用HR-DGCN进行局部结构特征提取，并引入父子标签参数相似性约束机制，有效提升对层次化威胁实体的识别能力；最后，利用Transformer编码器捕获全局语义依赖关系，并通过CRF解码输出最优实体序列。该方法实现了从局部结构特征到全局语义表示的多层次信息融合，为空管系统网络威胁实体识别提供了端到端的解决方案。

2.1　文本图构建

文本图构建的核心目标是将非结构化的威胁文本数据转化为结构化图表示。假设给定文档集 $\boldsymbol{D}=\{d_s,\boldsymbol{\mathcal{T}}_s\}_{s=1}^{M}$，其中 M 为文档总数，s 为文档索引，$\boldsymbol{\mathcal{T}}_s$ 为文档 s 对应的标签集，且 $\boldsymbol{\mathcal{T}}_s\subseteq\boldsymbol{\mathcal{L}}=\{l_i|i=1,\cdots,K\}$，$K$ 为标签类别总数，i 为标签索引。标签集具有父子层关系（子标签记为 $l_i^{(\varphi)}$，其中 φ 表示子标签在其父标签下的子类序号）。首先，进行语义规范化处理，采用Stanford CoreNLP工具[30-31]进行文本预处理，包括分词与词干化（如“attacking”→“attack”）、去除停用词，以及依

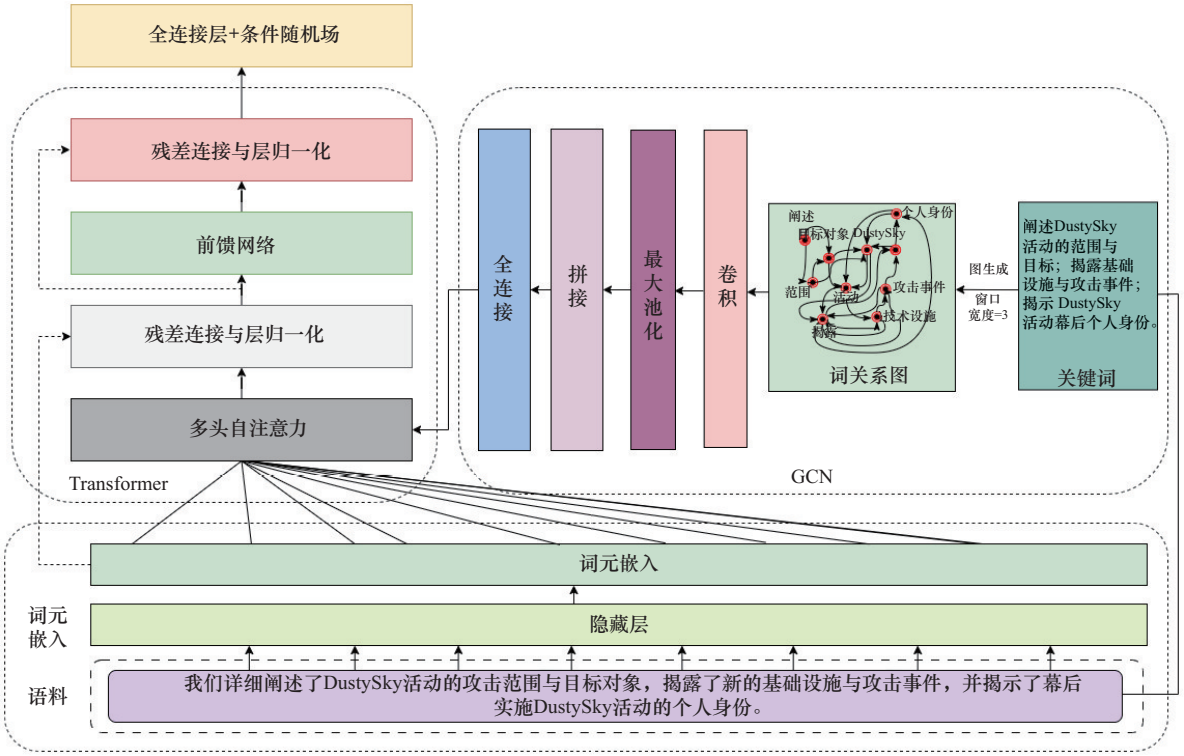

图2　融合GCN与Transformer的空管系统网络威胁实体识别方法的整体框架

赖解析，移除RCV1-v2标准[32]停用词表定义的120个无实义词（如冠词/介词），保留具有实体指示功能的关键词（名词/动词），减少噪声干扰。其次，进行图拓扑构建，将每个词干化后的单词作为独立节点，基于此构建文本图$G=(\boldsymbol{V},\boldsymbol{E})$。其中，节点集$\boldsymbol{V}$由词干化后的词汇组成，每个节点$V_i$的特征向量由GloVe词嵌入[33]（300维）与词性标注（one-hot编码）拼接而成；边集$\boldsymbol{E}$包含共现边（窗口大小设置为3）和语法依赖边两种类型。边权重w_{ij}分别反映共现频率和依赖强度，其表示式如式（1）所示。

$$w_{ij}=c_{ij} \tag{1}$$

其中，i和j为图中的节点索引，c_{ij}表示节点i与节点j在预设窗口内共现的次数，k为遍历节点i的所有共现邻居节点的索引，分母$\sum_k \mathrm{count}(i,k)$为节点$i$与所有邻居节点的共现次数之和。

式（1）将边权重w_{ij}归一化至$[0,1]$，用以反映节点j在节点i所有共现关系中的相对强度。

例如，在句子“APT32 launches spear phishing attacks”中，“APT32”（实体）为节点，“launch”为动作，“spear_phishing”为威胁类型。边“APT32”→“launch”的权重=0.6，“launch”→“spear_phishing”边的权重=0.8。

2.2　子图处理

为从全局图中提取具有判别性的局部结构，采用基于节点重要性的子图生成算法。节点重要性得分的表达式如式（2）所示。

$$\mathrm{Scores}_i(v_i)=\alpha\cdot\deg(v_i)d_i+\beta\cdot t_i+\gamma\cdot\sum_{j\in N(i)}w_{ij} \tag{2}$$

其中，α、β、γ为权重系数，经网格搜索分别设为0.4、0.3、0.3；s_i为节点v_i的重要性得分；d_i表示节点v_i的度；t_i表示节点v_i的TF-IDF值（term frequency-inverse document frequency）；w_{ij}表示边权重；$\boldsymbol{N}(i)$表示节点v_i的邻居节点集合；j为$\boldsymbol{N}(i)$中的邻居节点索引。

式（2）通过加权融合节点的度、TF-IDF值

及与邻居节点的边权重，用以衡量节点在文本图中的重要性。

选择重要性最高的N个节点作为根节点，对每个根节点执行BFS，直至子图包含的节点数达到预设阈值g（默认g=5）。

为解决子图尺寸不一致导致的卷积操作难题，需要对子图进行归一化处理：节点按到根节点的BFS深度[34]、节点度（降序）和与父节点的共现权重（降序）排序；对节点数多于g的子图截断后g个节点，对节点数少于g的子图添加全零填充节点。该归一化过程的时间复杂度为$O(Ng^2)$，其中N为根节点个数（即生成的子图数量）。

2.3　层次正则化深度图卷积网络

HR-DGCN是本文所提方法中的核心模块，其结构如图3所示。网络输入为规范化后的子图张量$X\in\mathbf{R}^{NgD}$，其中$\mathbf{R}$为实数集，N为根节点个数（即子图数量），g为子图规模（每个子图的节点数），D=50为嵌入维度。经过多层卷积与池化操作提取层次化特征。第一卷积层使用尺寸为50×1的卷积核，沿嵌入维度进行卷积，生成$N\times k_1$维特征图（k_1=128）；随后经过最大池化层将特征图尺寸降为$N/2\times k_1$；高层卷积层使用5×1卷积核进行跨子图特征组合，输出$k_1\times k_2$维矩阵（k_2=64）。所有卷积层后使用ReLU激活函数，全连接层中使用丢弃法（dropout=0.5）防止过拟合。

为利用实体标签的层次结构，在损失函数中引入层次正则化项，如式（3）所示。

$$\lambda(\boldsymbol{W})=\sum_{l_i\in\mathcal{L}}\sum_{j=1}^{K_i}\frac{1}{2}\|\boldsymbol{w}_{\mathrm{li}}-\boldsymbol{w}_{\mathrm{li}}^{(\varphi)}\|^2 \tag{3}$$

其中，$\mathcal{L}$为标签集合，$l_i\in\mathcal{L}$为第i个父标签，K_i为父标签l_i所包含的子标签数量，φ为子标签索引（$\varphi=1,2,\cdots,K_i$），$\boldsymbol{W}$为所有标签分类器参数的集合，$\lambda(\boldsymbol{W})$为层次正则化项，$\boldsymbol{w}_{l_i}$和$\boldsymbol{w}_{l_i}^{(\varphi)}$分别为父标签$l_i$和子标签$l_i^{(\varphi)}$对应的分类器参数向量。

该正则项通过强制父子标签的参数向量相近，实现参数共享，提升对稀有子类实体的识别能力。模型优化的最终目标函数如式（4）所示。

$$J(\boldsymbol{\theta})=H+C\cdot\lambda(\boldsymbol{W}) \tag{4}$$

其中，$J(\boldsymbol{\theta})$为模型优化的整体目标函数，$\boldsymbol{\theta}$为模型中所有可训练参数的集合，H为二元交叉熵损失，C为正则化强度系数（通过5折交叉验证设为0.7），$\lambda(\boldsymbol{W})$为层次正则化项。

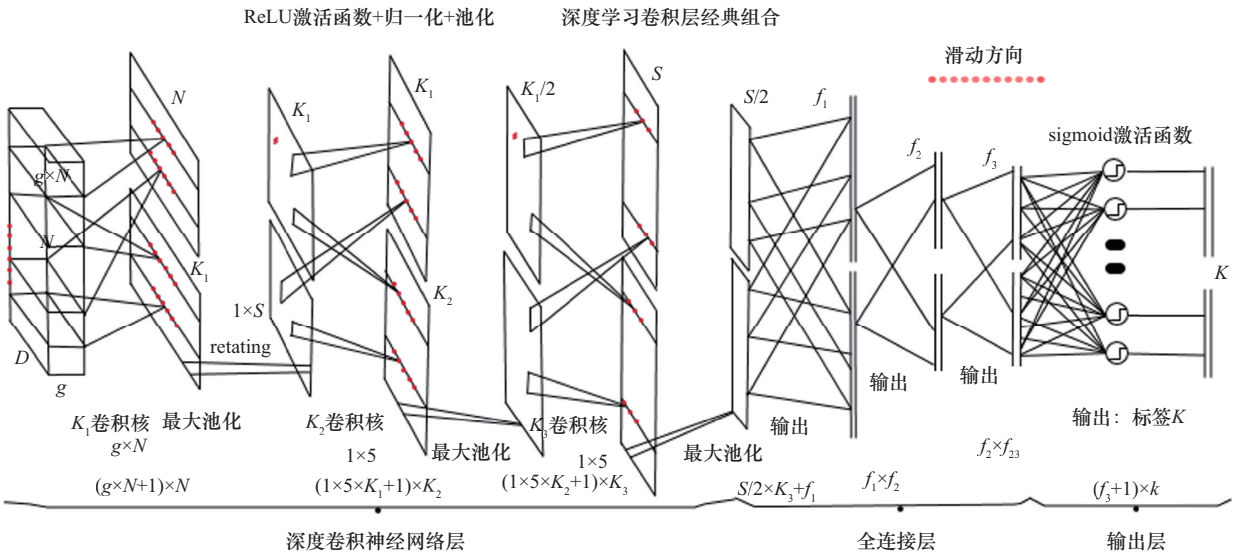

图3　HR-DGCN结构

2.4 GCN-Transformer集成架构

GCN-Transformer集成架构实现了局部结构特征与全局语义信息的有效融合，如图4所示。该架构首先通过HR-DGCN提取节点级特征表示 $H^{(G)} \in \mathbf{R}^{N \times d}$（$d$为特征维度）。随后，输入Transformer编码器捕获全局依赖关系。编码器由4层组成，每层包含8个注意力头，隐藏层维度为512，通过多头自注意力机制建模序列中任意两个元素之间的关系，其中自注意力机制的表达式如式（5）所示。

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{Z}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{Z} \quad (5)$$

其中，$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{Z}$分别为自注意力机制中的查询矩阵、键矩阵和值矩阵，$\boldsymbol{Q}\boldsymbol{K}^{\top}$为查询矩阵与键矩阵的转置相乘，$d_k$为键向量的维度。

最后，使用CRF解码最优标签序列，CRF层学习标签转移矩阵 $\boldsymbol{T} \in R^{K \times K}$（$K$为标签数量），通过考虑标签间的转移约束提升序列标注的准确性，从而实现端到端的联合优化。

3 实验与结果分析

3.1 实验设置

3.1.1 数据集

为验证本文所提方法的有效性，本文首先构建了空管系统威胁语料库（ATM-ThreatCorpus）。ATM-ThreatCorpus语料库统计信息见表1。

表1 ATM-ThreatCorpus语料库统计信息

数据类别	句子数/句	实体数/个	平均长度（词·句$^{-1}$）
训练集	10 246	28 735	23.4
开发集	1 281	3 592	22.8
测试集	1 280	3 581	24.1
合计	12 807	35 908	23.5

语料库数据主要来源于计算机应急响应小组（computer emergency response team，CERT）安全通告、空管系统漏洞报告、网络安全预警信息等。

为保证语料库标注质量，本文制定了系统化的标注流程，主要包括数据清洗、标注规范制定和标注者间一致性评估3个环节。

（1）数据清洗

① 去除无关文本内容。对收集到的各类原始文本，如CERT安全通告、空管系统漏洞报告、网络安全预警信息等，进行全面审查，剔除其中与威胁实体识别无关的部分，以减少噪声干扰，提升标注质量。

② 处理特殊字符与格式。统一文本中的特殊字符编码，将不同编码格式的特殊符号转换为标准格式，避免字符编码问题导致后续处理错误。同时，规范文本格式，如统一段落格式、标点符号使用等。

③ 去除重复数据。利用数据去重算法，如基于文本哈希值的方法，检测并去除语料库中的重复文本。重复数据不仅浪费存储资源，还可能对标注结果的统计和分析产生偏差。在去除重复

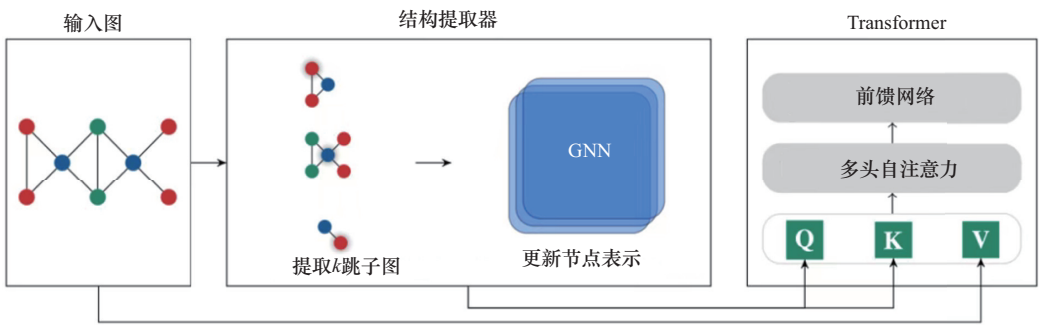

图4 局部结构特征与全局语义信息的有效融合

数据时，还需要保留一份完整的、具有代表性的文本。

（2）标注规范制定

根据空管系统网络安全领域的特点和威胁实体识别任务的需求，明确语料库中需要标注的实体类别。本语料库定义了 6 大类实体，分别为恶意软件（Malware）、IP 地址（IP）、URL 链接（URL）、哈希值（Hash）、CVE 漏洞（CVE）、攻击技术（Attack）。对于每个类别，给出详细的定义和示例，以便标注人员准确理解和区分。

（3）标注者间一致性评估

选择具有相关语言知识和标注经验的人员作为标注者。在正式标注前，对标注者进行全面培训，使其熟悉标注规范、标注工具的使用方法，以及常见问题的处理方式。培训过程中通过对实际案例的讲解和练习，确保标注者对标注任务有准确的理解。将语料库划分为若干个子集，随机分配给不同的标注者进行标注。确保每个子集的大小适中，既能保证标注效率，又能使标注结果具有代表性。同时，为避免标注者之间的相互影响，在标注过程中要求标注者独立完成任务，不得交流讨论。

3.1.2　评价指标与基线模型

采用精确率（Precision）、召回率（Recall）和 F1 分数（F1-score）作为评价指标，各指标的表达式如下。

$$\text{Precision}=\frac{\text{TP}}{\text{TP}+\text{FP}}\times 100\% \tag{6}$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP}+\text{FN}}\times 100\% \tag{7}$$

$$\text{F1}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\times 100\% \tag{8}$$

其中，TP（true positive）为正确识别的正例数，FP（false positive）为误识别的正例数，FN（false negative）为漏识别的正例数。

为验证本文所提方法（GCN-Transformer 模型）在网络安全领域的有效性，选择目前具有代表性的命名实体识别方法作为基线模型与之进行对比分析。基线模型包括：CRF[11]、SVM[12]、CNN[35]、CNN-CRF[20]、BiLSTM-CRF[36]、BERT-Tagging[14]、Transformer-CRF[23]、GCN[37]以及 GAT[38]。其中，（1）CRF 是一种经典的序列标注模型，能够有效利用上下文特征并建模标签间的依赖关系。（2）SVM 基于特征工程开展实体识别，依赖人工构造的特征模板，在数据稀缺场景下仍具有一定竞争力。（3）CNN 通过卷积操作捕捉文本中的局部特征，适用于实体边界的检测，但难以建模长距离依赖。（4）CNN-CRF 模型结合了 CNN 的特征提取能力与 CRF 的序列标注优化能力，利用 CNN 捕获局部上下文特征，并通过 CRF 约束输出标签的合理性。（5）BiLSTM-CRF 作为一种经典的序列标注模型，融合了 BiLSTM 和 CRF 的优势，能够有效捕获上下文双向依赖并对标签序列进行全局优化。实验中使用随机初始化的字符嵌入向量作为输入。（6）BERT-Tagging 基于预训练语言模型 BERT 的序列标注方法，凭借在大规模语料上预训练获得的深层语言表示，在小样本场景下仍表现出色。本研究采用 Google 发布的 BERT-base-Chinese 预训练模型，并利用 CRF 进行解码优化。（7）Transformer-CRF 模型将 Transformer 的自注意力机制与 CRF 相结合，既能有效捕捉长距离上下文依赖，又能引入标签间的结构约束。（8）GCN 将文本构建为图（节点为词，边为共现与依存关系），通过图卷积层聚合邻居节点信息并更新节点表示，最后接入全连接分类器进行实体识别。（9）GAT 在图结构基础上引入注意力机制，为每个邻居节点分配不同权重，动态聚合信息，能够更灵活地捕获语法依赖中不同邻居的重要性。这些模型涵盖了从传统机器学习方法到深层神经网络，以及结合预训练技术和图结构建模的不同技术路线，能够系统反映当前命名实体识别任务的主流发展脉络，并为本文所提方法的性能评估提供多维度、多层次的对比依据。

3.1.3 实验环境

实验在以下环境中进行：Intel Xeon Gold 6248R CPU、NVIDIA Tesla V100 GPU（32 GB）、256 GB RAM。使用 Python 3.8 和 PyTorch 1.9 框架实现所有模型。

3.1.4 实验参数和优化

训练过程中采用 Adam 优化器，初始学习率为 0.001，批量大小为 32，训练轮数为 100。为抑制过拟合，使用 L2 正则化，强度设置为 1.0×10^{-8}。词向量与词性标注（POS）向量维度分别设置为 200 维和 50 维。字符嵌入采用预训练的 200 维字符向量进行初始化。BiLSTM 隐藏层维度设为 100，层数为 2，最后一层隐藏状态应用 Dropout 机制，丢弃率设置为 0.5。融合后的混合嵌入层总维度为 350。模型结构中使用了 4 层 Star-Transformer 和 2 层 GCN，自注意力头数设为 3。Star-Transformer 隐藏层的 Dropout 丢弃率设置为 0.1，并引入门控线性单元（gated linear unit，GLU），以缓解梯度消失或梯度爆炸问题。

3.2 结果与分析

3.2.1 总体性能对比

为全面评估本文所提 GCN-Transformer 模型在空管系统威胁实体识别任务上的性能，在自建语料库 ATM-ThreatCorpus 上进行了多组实验，并与基线模型进行了对比。各模型在 ATM-ThreatCorpus 测试集上的整体性能对比见表 2。

从表 2 可以看出，GCN-Transformer 模型在精确率、召回率和 F1 分数 3 项指标上均显著优于所有对比基线模型，其 F1 分数达到了 88.9%，较最优基线模型 Transformer-CRF 提高了 3.6%。具体而言，与传统机器学习模型（CRF、SVM）相比，神经网络模型（CNN、BiLSTM-CRF）性能有明显提升，说明深度学习模型在特征自动提取与上下文建模方面具有显著优势。BERT-Tagging 与 Transformer-CRF 依托预训练语言模型强大的语义表征能力，性能进一步提升，但仍未充分利用文本的图结构信息与标签间的层次关系。GCN-Transformer 模型通过引入图结构建模、层次化正则约束，以及全局–局部特征融合机制，最终实现了最优的识别效果，尤其在召回率上提升更为明显，表明 GCN-Transformer 模型能够更全面地识别出文本中的威胁实体。

表 2 各模型在 ATM-ThreatCorpus 测试集上的整体性能对比

模型	精确率	召回率	F1 分数
CRF	76.2%	72.8%	74.5%
SVM	74.8%	70.3%	72.5%
CNN	78.5%	75.1%	76.8%
CNN-CRF	80.3%	77.6%	78.9%
BiLSTM-CRF	82.1%	78.3%	80.2%
BERT-Tagging	85.7%	83.2%	84.4%
Transformer-CRF	86.5%	84.1%	85.3%
GCN	79.1%	78.3%	77.6%
GAT	81.5%	82.4%	83.6%
GCN-Transformer	89.3%	88.6%	88.9%

3.2.2 消融实验

为分析本文所提模型各模块的贡献，本文开展了消融实验，结果见表 3。由表 3 结果可知，各模块对模型性能均有正向贡献。其中，层次正则化对模型性能的提升贡献最大（+3.8%F1 分数），证明利用标签层次关系约束分类学的重要性。GCN 模块贡献了 5.2% 的 F1 分数提升，说明引入图结构信息进行局部语法建模的有效性；Transformer 和 CRF 也分别带来了 2.6% 和 1.7% 的性能提升。各模块移除后模型性能均出现不同程度的下降，证明了本文所提架构中各组件设计的必要性与有效性。

表 3　消融实验结果

模型变体	精确率	召回率	F1 分数	ΔF1
完整模型	89.3%	88.6%	88.9%	—
无层次正则化	86.2%	84.0%	85.1%	−3.8%
无 GCN	84.5%	82.9%	83.7%	−5.2%
无 Transformer	87.1%	85.5%	86.3%	−2.6%
无 CRF 解码	87.8%	86.6%	87.2%	−1.7%

3.2.3　不同实体类型识别效果

为进一步分析模型性能，对比了 BiLSTM-CRF、BERT-Tagging 与 GCN-Transforme 模型在不同实体类型上的 F1 分数，结果见表 4。

表 4　不同实体类型上的 F1 分数对比

实体类型	BiLSTM-CRF	BERT-Tagging	GCN-Transforme
Malware	81.5%	85.2%	89.7%
IP	83.2%	86.7%	90.3%
URL	79.8%	82.4%	87.6%
Hash	78.3%	81.5%	86.9%
CVE	82.7%	86.3%	90.1%
Attack	76.4%	80.2%	85.7%

表 4 结果显示，GCN-Transforme 模型在所有实体类型上均取得最佳性能，特别是在 Attack 和 Hash 等难识别实体上提升显著，相较于 BERT-Tagging，其 F1 分数分别提高了 5.5% 和 5.4%。

3.2.4　词云图

实体识别词云图如图 5 所示。从图 5 可以看出，在 6 类空管网络威胁实体中，Malware 实体占比最多，例如，DustySky、VirusTotal、ClearSky、ScarCruft、DarKComet 等恶意软件。其次是 Hash 实体，如 db1156b072d58 acdac1aeab9af2160a2、331eca9c7d9fd9cbe7cd192af09880a3、5689448b4b6260ec9c35f129 df8b8f2622c66 a45 等，URL 和 IP 等其他类威胁实体显示相对较少。

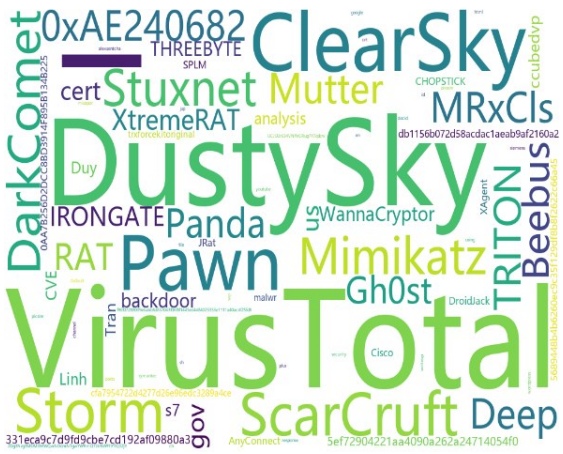

图 5　实体识别词云图

3.2.5　推理时延性能分析

为突出 GCN-Transformer 模型在推理时延方面的性能，选择了部分具有代表性的基线模型进行对比分析，包括 BiLSTM-CRF、BERT-Tagging 和 Transformer-CRF。这些模型在之前的实验中已经展示了其在识别准确率方面的性能，通过对比它们的推理时延，可以更全面地评估 GCN-Transformer 模型的综合性能。

各模型在测试集上的平均推理时延见表 5。

表 5　各模型在测试集上的平均推理时延

模型	平均推理时延/ms
BiLSTM-CRF	12.5
BERT-Tagging	45.2
Transformer-CRF	28.7
GCN-Transformer	22.3

从表 5 可以看出，GCN-Transformer 模型的平均推理时延为 22.3 ms，虽然略高于 BiLSTM-CRF 模型，但显著低于 BERT-Tagging 和 Transformer-CRF 模型。具体而言，BiLSTM -CRF 模型结构相对简单，主要基于 RNN 进行特征提取和序列标注，其推理过程相对直接，因此推理时延较低。GCN-Transformer 模型引入了 GCN 和更复杂的层次正则化机制，增加了模型的计算复杂度，导致

推理时延略有增加。BERT-Tagging模型基于大规模预训练语言模型BERT，虽然能够获得强大的语义表示能力，但BERT模型本身参数量巨大，在推理过程中需要进行大量的矩阵运算，导致推理时延较高。相比之下，GCN-Transformer模型没有采用如此庞大的预训练模型，而是基于GCN和Ttransformer的融合架构提取特征，在保证识别准确率的同时，有效降低了推理时延。Transformer-CRF模型主要依赖Transformer的自注意力机制进行特征提取和序列建模，虽然能够捕捉长距离依赖关系，但自注意力机制的计算复杂度较高，尤其是在处理较长序列时，会导致推理时延增加。GCN-Transformer模型通过引入GCN提取局部语法结构信息，减少了对Transformer自注意力机制的依赖，从而在一定程度上降低了推理时延。

实验结果表明，GCN-Transformer模型在推理时延方面具有一定的优势，能够在保证较高识别准确率的同时，满足空管系统网络威胁实时监测场景对推理速度的要求。

3.2.6 子图规模敏感性分析

在本文子图归一化策略中，固定子图规模（g=5），主要是为了在图卷积计算效率与语义完整性之间取得平衡，同时通过前置节点筛选机制降低截断/填充带来的风险。子图生成阶段已通过节点重要性评分选取Top-N关键节点作为根节点，进入子图的节点均为高重要性词汇，尾部截断的多为低权重关联节点，核心威胁实体与语法依赖节点极少被截断，有效降低了关键语义的丢失概率。全零填充为结构化计算服务，对节点数不足g的子图采取全零填充，目的是满足图卷积统一张量输入的要求，而非引入有效语义，因此无效信息有限。

为验证上述分析，在消融实验部分补充了定量分析与敏感性实验结果，不同子图规模的敏感性实验结果见表6。

表6 不同子图规模（g）敏感性实验结果

g	精确率	召回率	F1分数
3	87.1%	86.3%	86.7%
5	89.3%	88.6%	88.9%
7	88.9%	88.2%	88.5%

表6数据表明，子图规模为5时，是最优规模，子图规模增至7时模型性能不升反降，说明本文截断策略未丢失关键语义，反而过多节点会引入噪声；填充/截断操作仅在子图规模偏离最优值时产生微弱影响，在最优子图规模下的负面影响可忽略。

3.3 讨论

实验结果表明，本文提出的GCN-Transformer混合架构在空管系统威胁实体识别任务中表现出色。层次正则化机制有效缓解了稀有实体识别难题，GCN模块成功捕获了局部语法结构信息，而Transformer编码器则很好地建模了全局语义依赖关系。值得注意的是，本文方法在计算效率方面还有优化空间。完整模型训练时间约为BERT-Tagging的1.5倍，主要开销来自图构建和子图归一化过程。未来将通过采样优化和并行计算进一步提升模型计算效率。

4 总结与展望

针对空管系统网络威胁实体识别任务，本文提出了一种融合图卷积网络（graph convolutional network，GCN）与Transformer的空管网络威胁实体识别方法。具体而言，基于面向空管系统威胁文本的图结构构建与子图归一化方法，有效捕获语法和共现关系；采用层次正则化深度图卷积网络，通过参数共享提升层次化实体识别效果；采用GCN-Transformer混合架构，实现局部结构与全局语义的有效融合。实验结果表明，本文方法在F1分数、精确率和召回率等方面均显著优于现有其他基线模型，有效提升了空管系统威胁实体的识别性能。

未来工作将从以下3个方面展开：（1）多模态数据融合，结合网络流量日志、威胁图像数据等多模态信息，提升威胁实体识别的准确性。（2）不断提供领域适应性，探索跨领域迁移学习方法，将模型推广到其他安全领域。（3）开发基于该方法的空管系统网络威胁实时监测系统，实现工程化应用。

参考文献：

[1] Chen K, Nadirsha T N M, Lilith N, et al. Tangible digital twin with shared visualization for collaborative air traffic management operations[J]. Transportation Research Part C: Emerging Technologies, 2024, 161: 104546.

[2] De Giovanni L, Lancia C, Lulli G. Data-driven optimization for air traffic flow management with trajectory preferences[J]. Transportation Science, 2024, 58(2): 540-556.

[3] Post J. The next generation air transportation system of the United States: vision, accomplishments, and future directions[J]. Engineering, 2021, 7(4): 427-430.

[4] Yu X Q, Chen C H, Yang H H. Cognitive workload quantification for air traffic controllers: an ensemble semi-supervised learning approach[J]. Advanced Engineering Informatics, 2025, 64: 103065.

[5] Bogoda L, Mo J, Bil C. A systems engineering approach to appraise cybersecurity risks of CNS/ATM and avionics systems[C]// Proceedings of the 2019 Integrated Communications, Navigation and Surveillance Conference (ICNS). Piscataway: IEEE Press, 2019: 1-15.

[6] 湖北机场集团. 湖北机场集团聚焦"三力" 构筑航空网安屏障[EB]. (2024-04-25)[2026-04-20].
Hubei Airport Group focuses on "Three Forces" to build an aviation cybersecurity barrier[EB]. (2024-04-25)[2026-04-20].

[7] HFW, CyberOwl. Maritime cyber risk report: shipping industry remains "Easy Target", pays average US$3.2m in cyberattacks[R]. (2023-10-16).

[8] 丁兆云, 刘凯, 刘斌, 等. 网络安全知识图谱研究综述[J]. 华中科技大学学报(自然科学版), 2021, 49(7): 79-91.
Ding Z Y, Liu K, Liu B, et al. Survey of cyber security knowledge graph[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2021, 49(7): 79-91.

[9] Ren Y T, Xiao Y J, Zhou Y H, et al. CSKG4APT: a cybersecurity knowledge graph for advanced persistent threat organization attribution[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(6): 5695-5709.

[10] 彭嘉毅, 方勇, 黄诚, 等. 基于深度主动学习的信息安全领域命名实体识别研究[J]. 四川大学学报(自然科学版), 2019, 56(3): 457-462.
Peng J Y, Fang Y, Huang C, et al. Cyber security named entity recognition based on deep active learning[J]. Journal of Sichuan University (Natural Science Edition), 2019, 56(3): 457-462.

[11] McCallum A, Li W. Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons[C]//Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003. Morristown, NJ, USA: ACL, 2003: 188-191.

[12] Chauhan V K, Dahiya K, Sharma A. Problem formulations and solvers in linear SVM: a review[J]. Artificial Intelligence Review, 2019, 52(2): 803-855.

[13] 杨红梅,李琳,杨日东,等.基于双向LSTM神经网络电子病历命名实体的识别模型[J].中国组织工程研究,2018,22(20):3237-3242.
Yang H M, Li L, Yang R D, et al. Named entity recognition based on bidirectional long short-term memory combined with case report form[J].Chinese Journal of Tissue Engineering Research,2018,22(20):3237-3242.

[14] Meng F Q, Yang S S, Wang J D, et al. Creating knowledge graph of electric power equipment faults based on BERT-BiLSTM-CRF model[J]. Journal of Electrical Engineering & Technology, 2022, 17(4): 2507-2516.

[15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017: 6000-6010.

[16] Fu X L, Pan Y, Zhang L M. A causal-temporal graphic convolutional network (CT-GCN) approach for TBM load prediction in tunnel excavation[J]. Expert Systems with Applications, 2024, 238: 121977.

[17] Evangelatos P, Iliou C, Mavropoulos T, et al. Named entity recognition in cyber threat intelligence using transformer-based models[C]//Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience (CSR). Piscataway: IEEE Press, 2021: 348-353.

[18] Liu X, Chen H Q, Xia W G. Overview of named entity recognition[J]. Journal of Contemporary Educational Research, 2022, 6(5): 65-68.

[19] Kong J, Zhang L X, Jiang M, et al. Incorporating multi-level CNN and attention mechanism for Chinese clinical named entity recognition[J]. Journal of Biomedical Informatics, 2021, 116: 103737.

[20] 曹依依,周应华,申发海,等.基于CNN-CRF的中文电子病历命名实体识别研究[J].重庆邮电大学学报(自然科学版),2019,31(6):869-875.
Cao Y Y, Zhou Y H, Shen F H, et al. Research on named entity recognition of chinese electronic medical record based on CNN-CRF[J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition), 2019, 31(6): 869-875.

[21] 李博,康晓东,张华丽,等.采用Transformer-CRF的中文电子病历命名实体识别[J].计算机工程与应用,2020,56(5):153-159.
Li B, Kang X D, Zhang H L, et al. Named entity recognition in Chinese electronic medical records using Transformer-CRF[J]. Computer Engineering and Applications,2020,56(5):153-159.

[22] Yan H, Deng B C, Li X N, et al. TENER: adapting transformer encoder for named entity recognition[PP]. V3. arXiv (2019-12-10)[2026-07-23]. arXiv: 1911.04474.

[23] Wu Z H, Pan S R, Chen F W, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(1): 4-24.

[24] Zhang S, Tong H H, Xu J J, et al. Graph convolutional networks: a comprehensive review[J]. Computational Social Networks, 2019, 6(1): 11.

[25] Wang Y, Sun Y B, Liu Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1-12.

[26] Shaw P, Uszkoreit J, Vaswani A. Self-attention with relative position representations[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human LanguageTechnologies, Volume 2 (Short Papers). Stroudsburg, PA, USA: ACL, 2018: 464-468.

[27] Jiang C M, Yang Z P, Yao J T. Three-way decision enhanced graph convolutional networks for text classification[J]. Neural Processing Letters, 2025, 57(2): 37.

[28] Jin W Y, Ma H F, Zhang Y Y, et al. Dual-view graph-of-graph representation learning with graph Transformer for graph-level anomaly detection[J]. Neural Networks, 2025, 187: 107291.

[29] Zhu Q H, Wang Y, Li W M, et al. Advancing multi-modal beam prediction with cross-modal feature enhancement and dynamic fusion mechanism[J]. IEEE Transactions on Communications, 2025, 73(9): 7931-7940.

[30] Kumar S, Alam M S, Khursheed Z, et al. Enhancing relational database interaction through open AI and Stanford Core NLP-based on natural language interface[C]//Proceedings of the 2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST). Piscataway: IEEE Press, 2024: 589-602.

[31] Manning C, Surdeanu M, Bauer J, et al. The Stanford CoreNLP natural language processing toolkit[C]//Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Stroudsburg, PA, USA: ACL, 2014: 55-60.

[32] Lewis D D, Yang Y, Rose T G, et al. RCV1: A new benchmark collection for text categorization research[J]. Journal of Machine Learning Research, 2004, 5(4): 361-397.

[33] Pennington J, Socher R, Manning C. GloVe: global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA, USA: ACL, 2014: 1532-1543.

[34] Chand P K, Kumar M, Molla A R. Agent-driven BFS tree in anonymous graphs with applications[M]//Networked Systems. Cham: Springer Nature Switzerland2024: 67-82.

[35] Collobert R, Weston J, Bottou L, et al. Natural language processing (almost) from scratch[J]. The Journal of Machine Learning Research, 2011, 12: 2493-2537.

[36] Huang Z H, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging[PP]. arXiv (2015-08-09)[2026-07-23]. arXiv: 1508.01991.

[37] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Stroudsburg, PA, USA: ACL, 2019: 4171-4186.

[38] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[PP]. V3. arXiv (2018-02-04) [2026-07-23]. arXiv: 1710.10903.

[作者简介]



刘超（1989–），男，博士，西安航空学院计算机与人工智能学院研究生导师、高级工程师，主要研究方向为无线传感器网络、物联网、信号与信息处理以及群智能优化算法等。



田继伟（1993–），男，中国人民解放军空军工程大学副教授、研究生导师，主要研究方向为物联网/信息物理系统安全、人工智能安全以及低空智联网。



贾楠（1992– ），女，博士，西安航空学院计算机与人工智能学院讲师，主要研究方向为人工智能、图数据挖掘与图神经网络、遥感场景分类、图像处理等。



张欢乐（2003– ），女，西安航空学院计算机与人工智能学院硕士生，主要研究方向为程序设计、软件开发与计算机应用。



王布宏（1975– ），男，博士，中国人民解放军空军工程大学教授，主要研究方向为网络空间安全、空中交通管理系统网络安全以及人工智能安全等。



刘洲洲（1981– ），男，博士，西安航空学院计算机与人工智能学院教授、硕士生导师、副院长，主要研究方向为无线传感器网络、物联网、信号与信息处理和群智能优化算法等。