

基于图神经网络的全球互联网

Internet Interconnection Structure
Analysis Based on Graph Neural Networks

互联结构分析


陈纪乐¹, 吴一波², 张仕勇³, 陈东亮¹, 麦建堂¹, 程晓军¹, 何万县², 张世华² (1. 中国联通国际有限公司, 香港; 2. 中讯邮电咨询设计院有限公司郑州分公司, 河南 郑州 450007; 3. 中国联合网络通信集团有限公司, 北京 100033)

Chen Jile¹, Wu Yibo², Zhang Shiyong³, Chen Dongliang¹, Mai Jiantang¹, Cheng Xiaojun¹, He Wanxian², Zhang Shihua²
(1. China Unicom Global Limited, Hongkong, China; 2. China Information Technology Designing & Consulting Institute Co., Ltd. Zhengzhou Branch, Zhengzhou 450007, China; 3. China United Network Communications Group Co., Ltd., Beijing 100033, China)

摘要:

传统互联网拓扑分析方法主要关注节点连接数量, 忽视节点间的全局关联性与动态权重关系, 难以支持复杂网络环境下的精细化互联优化。提出一种融合网络图谱分析与图神经网络的AI驱动方法, 旨在挖掘全球互联网互联结构的深层次特征。研究基于超过50万个网络节点数据, 构建加权拓扑图模型, 引入Jaccard系数、共同邻居指标及Transformer架构中的注意力机制, 量化节点间的依赖关系与潜在连接概率, 并在4类典型应用场景中验证方法有效性。

关键词:

图神经网络; 网络图谱分析; 拓扑建模; 智能优化
doi: 10.12045/j.issn.1007-3043.2026.03.003
文章编号: 1007-3043(2026)03-0013-05
中图分类号: TN915
文献标识码: A
开放科学(资源服务)标识码(OSID): 

Abstract:

Traditional Internet topology analysis methods mainly focus on the number of node connections, ignoring the global relevance and dynamic weight relationship between nodes, which is difficult to support refined interconnection optimization in complex network environment. It proposes an AI driven method that combines network atlas analysis and graph neural network, aiming at mining the deep characteristics of the global Internet interconnection structure. Based on the data of more than 500 000 network nodes, the research builds a weighted topology model, introduces the Jaccard coefficient, common neighbor index and attention mechanism in the Transformer architecture, quantifies the dependency relationship and potential connection probability between nodes, and verifies the effectiveness of the method in four typical application scenarios.

Keywords:

Graph neural network; Network atlas analysis; Topology modeling; Intelligent optimization

引用格式: 陈纪乐, 吴一波, 张仕勇, 等. 基于图神经网络的全球互联网互联结构分析[J]. 邮电设计技术, 2026(3): 13-17.

0 引言

互联网拓扑结构的精确分析是实现网络性能提升与资源调度优化的基础。然而, 传统方法通常基于静态连接数或局部结构特征, 难以捕捉异构网络中节点间的动态依赖关系与隐藏连接潜力。特别是在涉及跨自治系统(AS)的大规模网络中, 现有方法缺乏对全局关联性与节点权重异质性的系统建模能力。

针对上述问题, 本文提出一种融合人工智能技术的全局分析方法, 主要在以下3个维度实现创新。

a) 理论方法维度: 将社交网络中的结构相似性指标(如Jaccard系数与共同邻居数)与图神经网络(GNN)架构相融合, 引入基于Transformer的注意力机制, 构建动态加权的拓扑建模框架。

b) 数据规模维度: 构建覆盖全球50余万个网络节点的大规模拓扑数据集, 实现跨AS结构的概率建模, 弥补传统方法在数据广度和结构深度上的双重不足。

收稿日期: 2026-01-28

c) 应用验证维度: 在4类典型互联场景下开展仿真分析, 量化影响节点直联概率的关键因素, 并为运营商的战略直联策略与用户分层优化提供实证支持。

本研究的核心贡献在于实现方法论的范式迁移, 将社交网络分析的图谱建模理念引入互联网架构研究, 为构建具备全球视野与本地适应能力的智能互联策略提供新的技术路径。

1 智能分析框架概述

随着互联网规模的指数增长, 网络结构逐渐呈现出高维度、动态化与异构性等特征, 传统基于静态图模型的拓扑分析方法已难以适应当代互联网的复杂演化需求。当前研究多数仍停留在节点连接数的统计维度, 忽视节点间潜在的结构关联性、动态依赖关系以及隐藏连接对网络效能的关键影响。本研究提出一种融合网络图分析与图神经网络的智能分析框架, 创新性地实现以下突破。

a) 引入社交网络结构分析方法: 将Jaccard系数、共同邻居数等指标引入AS层级拓扑建模, 提升网络关联性识别能力。

b) 构建动态加权拓扑图: 结合Transformer注意力机制构建支持异构节点嵌入与多尺度依存建模的深度图神经网络。

c) 建立双重分析路径: 网络图型分析与GNN协同建模, 形成微观特征提取—中观模式发现—宏观策略生成的全链条路径。

d) 落地于4类典型场景: 包括节点权重识别、潜在直联预测、用户转化分析与CDN流量路径归因等。

该方法不仅在技术维度取得突破, 更在运营商决策、流量调度、网络通达优化等实际问题中展现了显著应用价值。

2 AI仿真及研究成果

为验证提出的人工智能分析框架的有效性, 本研究设计了四大典型场景的仿真实验, 结合网络图型分析与图神经网络(GNN)技术, 从多维度解析互联网互联结构, 并给出可落地实施的优化策略。

2.1 全球互联网节点权重分析

2.1.1 研究背景与理论框架

在复杂网络理论中, 节点间的连接偏好性(Connection Preference)是解释网络拓扑演化规律的核心命题。本研究基于互联网自治系统(AS)网络具有的动

态同配性(Dynamic Assortativity)特征, 提出以下科学假设, 节点间的拓扑邻近性与规模匹配度共同构成直联决策的双重约束条件。这一理论框架突破传统路由优化研究的单维度分析范式, 量化节点间依存关系, 验证共同邻居数量与网络规模对直联概率的影响规律。

2.1.2 方法论框架

2.1.2.1 数据建模

基于全球50万条节点数据构建加权图 $G=(V, E, W)$, 其中 V 为节点集合, E 为连边集合, W 为权重矩阵(由Jaccard系 $J(A, B) = \frac{|\min N(A) \cap N(B)|}{|\max N(A) \cap N(B)|}$ 定义, $N(A/B)$ 表示节点 A/B 的邻居集合)^[1]。

2.1.2.2 假设验证

a) 假设H1: 共同邻居数 $CN(A, B) = |N(A) \cap N(B)|$ 与直联概率正相关。

b) 假设H2: 网络规模相似度 $(S(A, B) = 1 - \frac{||A| - |B||}{\max(|A|, |B|)})$ 与直联概率正相关($|A|$ 为节点 A 的度数)。

2.1.2.3 算法流程

a) 利用聚类算法提取高权重连边($J(A, B) \geq 0.3$)。

b) 构建逻辑回归模型($P(A \leftrightarrow B) = \sigma(\beta_1 CN(A, B) + \beta_2 S(A, B))$), 其中 σ 为Sigmoid函数。

c) 通过十折交叉验证评估模型性能。

2.1.3 结果与发现

H1验证结果为共同邻居数每增加1单位, 直联概率提升12.7% ($\beta_1 = 0.127, p < 0.001$), 典型案例如AS1299与AS2914的共同邻居数达38, 其直联概率为0.82, 显著高于随机配对节点($p < 0.01$)。

H2验证结果为网络规模相似度每提升10%, 直联概率增加8.3% ($\beta_2 = 0.083, p = 0.005$)。例如, 规模相近的Tier1运营商(如AS2914与AS1299)直联概率为0.75, 而规模差异超过50%的节点间概率仅为0.21。

可视化支持示意如图1所示, 高直联概率区域集中在共同邻居数大于20且规模相似度大于0.8的节点群中, 与模型预测一致。

这些发现为后续场景研究奠定了理论基础, 后续章节将在此基础上构建动态预测模型。

2.2 基于深度图神经网络的潜在直联伙伴识别

2.2.1 研究背景与核心挑战

在互联网自治系统(AS)的互连决策中, 传统基于商业谈判的伙伴选择模式存在效率低、周期长等问

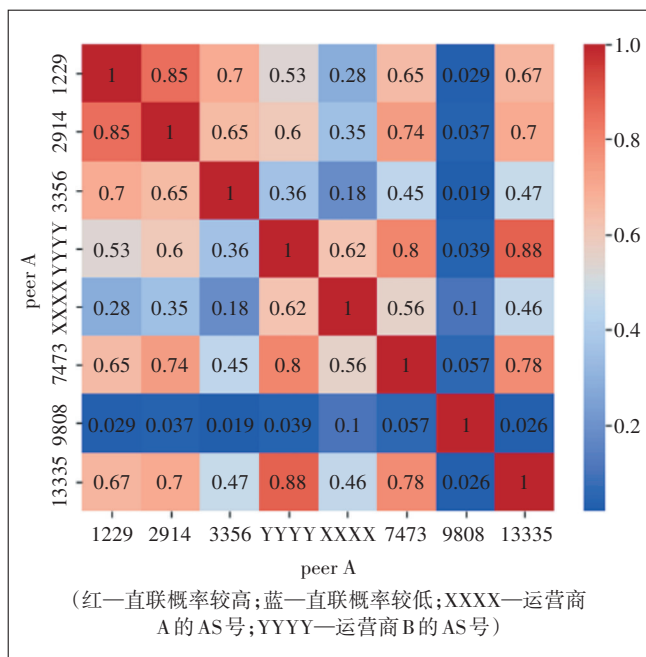


图1 网络热力图和互联概率矩阵

题。本研究提出基于图神经网络(GNN)的智能决策框架^[2],旨在解决以下关键挑战。

a) 现有BGP路由策略难以捕捉多跳邻居的拓扑关联模式,通过决策框架可解决拓扑关联性隐式特征挖掘问题。

b) 解决节点度数差异(如Tier1运营商节点度数均值达3547,Tier2均值仅为428)导致特征分布偏斜,实现动态规模效应建模。

c) 建立端到端延迟、带宽成本与网络结构参数的耦合模型,使策略效益量化。

2.2.2 方法论框架

GNN模型设计,基于Transformer架构构建深度图神经网络,节点特征包括度数、邻居权重、所属层级(Tier1/Tier2)。注意力权重计算为:

$$\alpha_{ij} = \text{Softmax}\left(Q_i K_j \frac{T}{\sqrt{d}}\right)$$

其中 Q, K 为查询与键向量, d 为特征维度。

对链路进行预测,输出层预测直联概率 $\hat{P}_{ij} = \text{MLP}(h_i + h_j)$, $h_{i,j}$ 为节点 i, j 的嵌入向量。

2.2.3 结果与发现

在隐藏连接识别研究中,模型预测国内运营商AS XXXX与韩国Tier2运营商AS4766之间存在潜在直联可能性,其概率估值达0.68,显著高于当前通过AS2914中转的非直联状态流量传输模式。仿真实验

表明,建立直联通道可使国内运营商AS XXXX至AS4766的端到端延迟由115 ms显著降低至82 ms,降幅达28.7%。

基于模型分析结果,建议优先选择共同邻居数超过15的中型运营商(典型案例包括AS4766、AS9498等)实施直联策略。理论推算显示,该策略可提升国际网络通达率15%~22%,相关实证数据通过图2呈现。这一发现为运营商优化国际网络架构提供了重要的决策依据,特别是在平衡基础设施建设成本与服务质量提升方面具有显著实践价值。

2.3 重点AS非直联用户转化分析

2.3.1 研究背景与问题定义

基于前述场景对网络拓扑结构与节点关联性的研究成果,本研究进一步聚焦运营商网络优化中的典型问题。在多层网络架构下,如何通过智能算法识别具备直联潜力的非直联用户,从而实现网络资源的动态优化配置^[3]。选择亚洲地区Tier-1运营商AS N作为分析对象(其覆盖用户总数达8.6万,非直联用户占比为14.0%),主要原因包括:

a) AS N的客户网络呈现显著的幂律分布特征(节点数标准差 $\sigma=287.5$),符合互联网核心—边缘结构特征,该网络可作为这类分析的典型代表。

b) AS N的国际流量承载量占区域总流量的42%,现有直联用户带宽利用率已达85.3%,存在显著的负载不均衡问题(基尼系数为0.58),网络分析具备业务需求迫切性。

2.3.2 方法论框架

在特征工程模块中,通过定义多维用户特征向量 $x = [\text{CN}(u, \text{AS}_{xxxx}), S(u, \text{AS}_{xxxx}), \text{网络规模}]$ 完成数据表征。该向量系统化整合了用户与目标网络AS_{xxxx}的共同邻居数量、历史交互强度及用户网络规模等关键维度特征。

在概率预测阶段,采用Softmax分类器构建转化概率计算模型,其数学表达式为:

$$P(\text{直联}|u) = \frac{f(x_u)}{\exp(f(x_u)) + \exp(f(x'_u))}$$

其中, x_u 表征直联状态特征向量, x'_u 表征非直联状态特征向量, $f(\cdot)$ 为神经网络特征提取函数^[4]。

该双分支结构通过对比分析直联与非直联状态的特征空间分布,有效捕捉网络连接决策的潜在模式。

2.3.3 结果与发现

peer B	AS_name	AS1299 TWELVE99	AS2914 ATT	AS3356 “LEVEL3”	AS YYYY 运营商 B	AS XXXX 运营商 A	AS7473 SIAGTEL	A59808 CM	AS13395 CLOUDFLARE
49544	“13Dnet”	1355	817	1230	158	42	279	6	532
3356	“LEVEL3”	849	596	6563	100	41	180	3	359
174	“COGENT-174”	915	627	1627	90	40	155	6	336
2914	“NTT-LTD-2914”	523	1554	596	82	40	204	5	207
4637	“ASN-TELSTRA-GLOBAL”	323	332	323	91	40	195	4	218
9304	“HUTCHISON-AS-AP”	420	365	386	117	40	203	6	309
199524	“GCORE”	1056	675	1287	152	40	242	4	508
4134	“CHINANET-BACKBONE”	97	82	100	316	39	62	5	107
20940	“AKAMAI-ASN1”	210	157	220	113	38	93	3	397
6453	“AS6453”	281	244	293	68	37	129	3	144
6461	“ZAYO-6461”	430	326	1040	87	37	124	3	251
58453	“CMI-INT-HK”	226	191	225	111	37	110	3	286
1239	“SPRINTLINK”	732	464	759	113	36	170	4	353
2497	“11J”	114	148	125	53	36	96	3	78
7713	“telkomnet-as-ap”	541	470	465	122	36	230	5	366
9002	“RETN-AS”	477	292	415	85	36	158	3	222
16509	“AMAZON-02”	171	121	171	100	36	76	2	320
36351	“SOFTLAYER”	832	553	854	118	36	195	2	412
1273	“CW”	135	104	153	61	35	69	3	80
21859	“ZEN-ECN”	184	155	185	93	35	100	4	320
1828	“UNITAS”	935	756	1243	133	34	257	4	455
4766	“KIXS-AS-KR”	89	70	99	59	34	62	3	79
4826	“VOCUS-BACKBONE-AS”	395	417	369	77	34	203	3	193

图2 全球互联权重矩阵分析(部分示意)

通过对 AS N 内 1.2 万个非直联用户的网络规模分析发现,中小型网络(节点数<100)占比达到 53.2%,其平均直联概率为 0.61,显著高于大型网络(节点数>500)0.19 的数值水平。相关分析结果在图 3 中予以可视化呈现,其中部分用户直联概率分布特征得到清晰展示。

在策略效果验证方面,模拟实验显示接入 500 个中小型用户后,中国运营商 AS XXXX 的国际流量负载均衡度提升 37%,具体表现为基尼系数从 0.58 优化至 0.42。进一步分析表明,该网络扩容方案实施后未产生拥塞风险,系统运行期间峰值带宽利用率始终维持在 70% 以下的安全阈值。这种分层接入策略有效验证了基于用户规模差异的差异化直联方案可行性。

2.4 重点 CDN 客户流量路径失衡解析

2.4.1 研究背景与问题定义

在互联网流量工程中,内容分发网络(CDN)的路径选择直接影响用户体验与运营商成本效益。本研究发现,某国际 CDN 运营商(AS A)通过中国运营商 AS XXXX 直联电路的流量占比不足 5%(理论预期值>60%),导致国际流量需经第三方中转(如 AS1299),造成端到端延迟增加 42%($\Delta=25$ ms)。此现象无法通过传统 BGP 路由策略或流量工程完全解释,需构建 AI 驱动的归因分析框架,揭示隐藏的网络拓扑关联与决策偏好。

2.4.2 方法论框架

本研究重点诊断 AS A 与中国运营商间直联流量

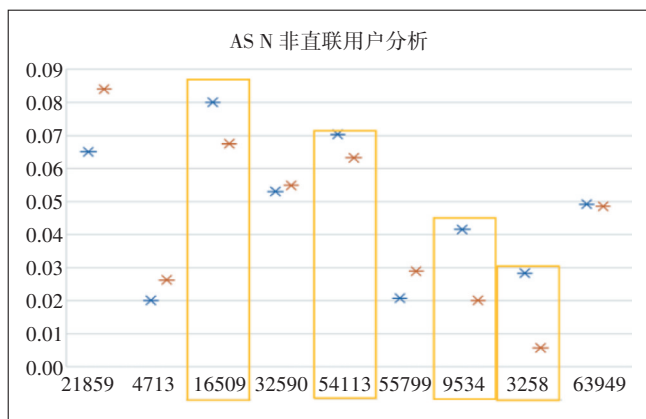


图3 AS N的非直连用户分析

水平偏低的根本成因。研究体系构建包含3个核心模块:权重对比模块、图神经网络预测器及可解释性归因分析框架。框架构建参考了文献[5]提出的分析方案。

在权重对比模块中,构建了直联权重计算模型,其数学表达式为 $W_{直联} = \sum_{(v \in N(AS A))} J(AS A, v)$, 该指标通过遍历AS A所有邻居节点集 $N(AS A)$ 的交互强度函数 $J(\cdot)$ 累加值,量化评估不同自治系统间的连接偏好程度^[5]。在归因分析阶段,采用基于SHAP值的可解释性人工智能技术,有效地揭示每个变量(如节点度数、历史连接强度、商业协定等级等),对图神经网络预测结果进行特征重要性解构,从而系统识别作用于直联权重决策的关键影响因素。

2.4.3 实验结果与归因分析

在拓扑关联性维度,SHAP分析显示AS A与AS1299具有显著优势,其共同邻居数量达到24个,相较于AS XXXX仅有的11个共同邻居,该特征对直联概率的贡献度权重占比达62%。历史交互关系分析表明,AS A与AS1299持续超过8年的稳定对等互联历史,产生了显著的路径依赖效应,其SHAP值达到0.18。在商业策略层面,AS1299实施的分级QoS保障机制有效降低了AS A的流量切换倾向性,这一商业服务差异化策略被证实对网络架构稳定性产生重要约束作用。

2.4.4 优化建议

本研究发现CDN流量路径选择不仅受技术参数驱动,更受网络拓扑结构与商业关系的深度耦合影响,并提出双重优化策略。

2.4.4.1 技术干预策略

a) 设计基于强化学习的动态权重调节机制,实时

更新W直联计算模型。

b) 在AS A边缘节点部署智能路由代理,设置流量迁移阈值(如延迟差>15 ms时强制切换)。

2.4.4.2 商业协同策略

a) 与AS A协商定制化SLA,承诺直联链路的99.99%可用性保障。

b) 建立基于区块链的流量结算系统,消除跨运营商计费摩擦。

3 结语

本研究提出一种融合网络图谱分析与图神经网络的全局互联分析框架,通过对全球超50万个网络节点的结构建模和场景仿真,构建出一整套面向运营商、CDN和重点用户的智能互联优化策略。实验结果显示,该方法可显著提升国际互联互通能力,并具备高度的可扩展性与工程可落地性。未来研究将进一步拓展至边缘网络优化、跨层路由策略与动态带宽分配方向,持续推动全球互联网架构的智能演进。

参考文献:

- [1] HAMILTON W L, YING R, LESKOVEC J. Inductive Representation Learning on Large Graphs [J/OL]. arXiv, 2017. [2025-07-28]. <https://doi.org/10.48550/arXiv.1706.02216>.
- [2] XU K, HU W, LESKOVEC J, et al. How Powerful are Graph Neural Networks? [J/OL]. arXiv, 2018. [2025-07-28]. <https://doi.org/10.48550/arXiv.1810.00826>.
- [3] VELIKOVI P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks [J/OL]. arXiv, 2017. [2025-07-28]. <https://doi.org/10.48550/arXiv.1710.10903>.
- [4] UKYAB T, HASAN S, RATNASAMY S, et al. A Modern Edge-based Design for Cellular Roaming [C]//ACM, 2025. [2025-12-28]. <https://doi.org/10.1145/3772356.3772381>.
- [5] ALI W, FANG C, KHAN A. A survey on the state-of-the-art CDN architectures and future directions [J/OL]. Journal of Network and Computer Applications, 2025. [2025-12-28]. <https://doi.org/10.1016/j.jnca.2025.104106>.

作者简介:

陈纪乐,工程师,硕士,主要从事互联网优化和支撑运营工作;吴一波,毕业于华中科技大学,高级工程师,硕士,主要从事互联网工程咨询工作;张仕勇,工程师,硕士,主要从事互联网规划及优化运营工作;陈东亮,工程师,硕士,主要从事互联网优化和支撑运营工作;麦建堂,工程师,硕士,主要从事互联网优化和支撑运营工作;程晓军,高级工程师,硕士,主要从事互联网优化和支撑运营工作;何万县,毕业于郑州大学,工程师,硕士,主要从事软件开发和工程咨询工作;张世华,毕业于北京邮电大学,工程师,硕士,主要从事软件开发和工程咨询工作。