

大语言模型文献挖掘驱动的网络指标体系与场景差异化分析

徐祺坤^① 刘娅汐^{*①} 韩淑娴^{①②} 张慧峰^② 皇甫伟^①

^①(北京科技大学计算机与通信工程学院 北京 100083)

^②(之江实验室 杭州 311121)

摘要: 网络指标是网络设计、运维与优化的基础。本文提出一种基于大语言模型(Large Language Model, LLM)的文献挖掘方案,对《IEEE/ACM Transactions on Networking》2023-2024年刊载的583篇文献进行指标提取、术语规范化、场景标注与指标关联量化,并通过双LLM交叉验证和人工抽样复核检验挖掘结果。结果表明,网络指标呈现“少量共性指标+大量长尾专有指标”的结构,不同应用场景所关注的核心指标集及指标关联均具有场景依赖性。本文提出服务质量(Service, S)、资源复用成本(Multiplexing, M)以及场景适配重构能力(Versatility, V)三类相互独立且可组合的分析维度,为网络架构的设计、部署、运维和优化提供指导。

关键词: 网络指标; 大语言模型; 指标体系; 多模态网络

中图分类号: TN915.0

文献标识码: A

文章编号: 1009-5896(2026)00-0001-11

DOI: 10.11999/JEIT251120

CSTR: 32379.14.JEIT251120

1 引言

网络已成为现代社会最重要的信息基础设施之一,在社会生产生活中持续演进。近年来,在物联网(Internet of Things, IoT)、虚拟现实与增强现实、人工智能推理等新兴服务的推动下,第六代移动通信网络(6G)、算力网络、多模态智能网络等新型网络技术不断涌现^[1-3],网络制式日益复杂,网络应用场景日益多样。网络技术的发展以及业务的多样化与复杂化,对科学、精确地认知、理解和评估网络提出了更高要求。在此意义上,网络指标体系,即用于量化评估网络在性能、可靠性、可用性与资源效率等多维属性的标准参数集,构成了认知与调控网络的核心依据。网络指标体系不仅映射了多元利益主体的差异化诉求,更贯穿于网络设计、部署及运维的全生命周期,是连接学术理论与工程实践的重要桥梁^[4]。

网络指标体系随网络技术演进而发展,学术界已对其形式化描述和分类展开研究。从不同的视角,网络指标具有不同的具体指称和内涵。例如,从用户视角出发,这些指标通常被归类于业务服务质量(Quality of Service, QoS)或用户体验质量(Quality of Experience, QoE);而运营商则更侧重于关注服务等级协定(Service Level Agreements, SLA)或关键性能指标(Key Performance Indicators,

KPI)。因此,学者们从多种视角对网络指标进行了整理与分类。Andrew Campbell等人^[5]针对IP(Internet Protocol)网络提出了QoS框架,将网络指标分为丢包率、吞吐量、时延与抖动等类别,对每类又分为确定性指标、统计性指标和尽力而为(Best Effort)指标。Murat Karakus等人^[6]聚焦软件定义网络(Software Defined Network, SDN),从带宽、时延、抖动与丢失四个核心类别研究QoS体系,每类指标再分为硬约束和软约束两个子类。Neha Kaushik等人^[7]将网络微服务架构中的指标概念扩展至可部署性、可测试性、可靠性等九个维度,相对全面地涵盖了网络的属性特征。这些工作对网络指标体系进行了描述与分类,所考察的指标范围不断扩充完善。

网络指标之间并非相互孤立,而是存在复杂的耦合关系与层级结构。Andreas Hanemann等人^[8]提出了时间聚合、空间聚合与空间级联三种聚合方式,如多跳传输中逐跳时延可以级联聚合为端到端总时延,揭示了宏观整体指标与微观局部指标间的联系。Salman Al-Shehri等人^[9]考察了不同利益方对网络指标的关注差异,分析不同利益方在能耗、服务质量与运营成本等方面的关注侧重,并进一步讨论其所关注指标间的映射关联。面向特定类型网络的指标体系研究也逐步深入,如数据中心网络^[10]、边缘计算网络^[11]、5G-R(5G-Railway,基于5G技术的铁路移动通信系统)^[12]、无线传感器网络^[13]等场景。网络应用场景的边界越清晰,指标间的关联约束越显著,如文浩等人针对无线传感器网络场景,不仅提出了层次化的指标体系模型,还明确辨识出特定指标间的正、负向耦合影响^[13]。

收稿日期: 2026-xx-xx; 改回日期: 2026-06-29; 网络出版: 2026-07-08

*通信作者: 刘娅汐 yaxi.ustb@gmail.com

基金项目: 国家重点研发计划“网络模态共生演化模型与体系结构”(2023YFB2903901)

Foundation Item: 2023YFB2903901

现有网络指标体系研究多面向单一场景、单类技术或少量代表性指标展开,在跨场景、跨文献的大规模网络指标分析时,仍缺乏统一的抽取规则与可复核的分析流程^[14]。与此同时,共现分析、引文分析等传统文献计量方法能够较好揭示知识传播结构和共现规律,但难以直接表征指标对之间的语义方向及强度。因此,网络指标的系统化研究需同时兼顾规模化覆盖、术语一致性与关系可解释性^[15]。

近年来,以GPT(Generative Pre-trained Transformer)和DeepSeek^[16]为代表的大语言模型(Large Language Model, LLM)迅速发展,在语言理解、生成与推理等任务中展现出显著优势,为实现科学文献中的实体识别、关系抽取和知识图谱构建^[17]提供了可能。本文引入LLM对网络领域的研究文献进行文本挖掘,围绕网络指标研究中的术语歧义、场景差异和复杂关系量化等问题,构建可复核的领域化分析链,主要贡献包括:

(1)构建“术语规范化—场景标注—关系量化”的领域化LLM文献挖掘流程,对网络领域权威期刊《IEEE/ACM Transactions on Networking》(ToN)于2023-2024年刊载的全部583篇文献开展挖掘,实现网络指标自动提取、术语差异消解、网络场景标注及指标间关联关系量化。

(2)开展数据统计与分析,通过双LLM交叉验证与人工抽样复核相结合的方式,对自动化挖掘结果的可靠性进行验证;揭示了网络指标体系的重尾分布和场景依赖性规律,讨论了不同场景的差异化关注指标集合及其关联模式。

(3)提出服务质量(Service, S)、资源复用成本(Multiplexing, M)及场景适配重构能力(Versatility, V)三类相互独立且可组合的分析维度。其中, S用于回答“服务是否满足需求”, M用于回答“代价

是否可接受”, V用于回答“网络能否以较低重构成本适配新场景与新业务”,从指标体系视角为网络架构设计及其全生命周期管理提供了理论参考。

2 基于LLM的网络领域文献挖掘流程

本节将详细介绍基于LLM的网络领域文献挖掘方法与具体流程。LLM是自然语言处理领域的重大技术突破,它首先通过自监督学习在大规模无标注语料中构建通用语言表征,捕获词汇、语法及语义的复杂关联规则,随后通过监督式指令微调对齐人类意图,赋予模型上下文感知的任务泛化能力。LLM的泛化性表现在它能够对模糊表述、隐含逻辑及专业术语进行语义理解和深度解析,为网络领域文献挖掘提供了工具支撑。

2.1 总体流程与任务分解

本文选取网络领域权威期刊《IEEE/ACM Transactions on Networking》2023-2024年度刊载的583篇论文构建文献数据集。该期刊是电气电子工程师协会(Institute of Electrical and Electronics Engineers, IEEE)与计算机学会(Association for Computing Machinery, ACM)在网络领域的汇刊,全面覆盖网络研究前沿方向,适合作为网络指标研究的文献来源。

为高效开展网络领域文献挖掘工作,本文将整体研究流程划分为语料输入、文本预处理、智能体执行、知识融合、双LLM交叉验证、文献聚合及统计可视化六个核心环节,该流程围绕网络指标研究需求,依次完成术语规范化、场景标注与关系量化,将三者输出结果统一映射至同一套指标词表与场景标签体系中。

2.2 文本预处理与知识融合细节

文本预处理器负责完成PDF文本提取、页面噪

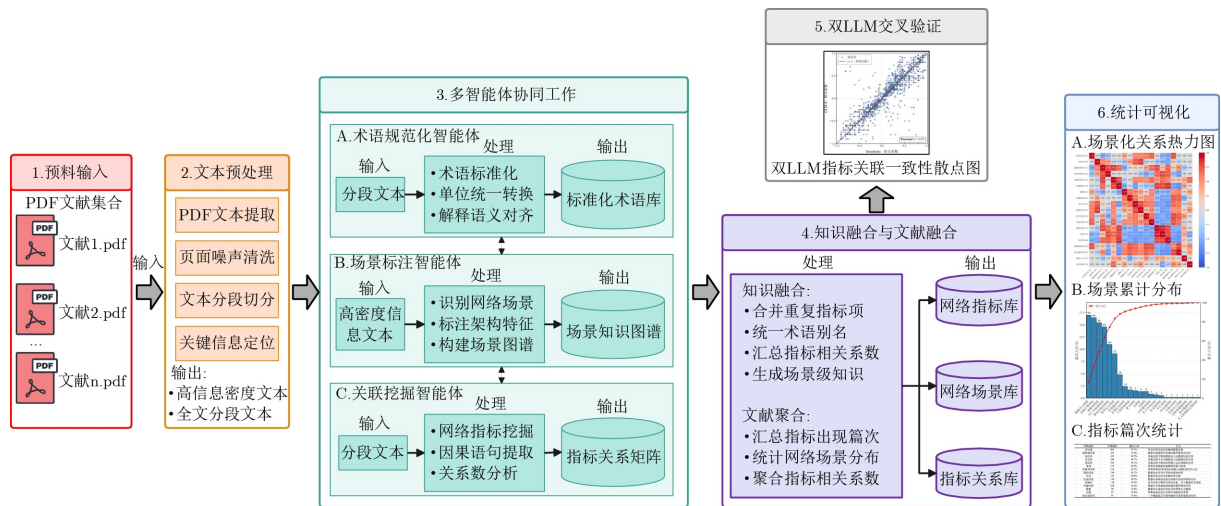


图1 LLM驱动的网络领域文献挖掘数据流图

声去除、文本分段切分及关键信息定位。具体而言，本文采用全文分段文本开展指标提取与关系挖掘工作，同时利用前若干页的摘要、引言等高信息密度文本进行场景识别，以降低无关段落对分类结果的干扰；对于公式、脚注及参考文献等低相关性文本，预处理阶段对其进行弱化或删除，从而降低模型对非正文内容的误判概率。

知识融合模块主要用于合并不同文本分段及不同文献中的重复指标项，统一指标术语的别名表述，并将多次出现的指标关系汇总为场景级统计结果；针对同一指标对在多篇文献中的多种关系表述，先对其进行统一命名，再开展聚合统计，避免因术语别名、缩写或上下文差异导致的重复计数问题。文献聚合模块以单篇论文为基本统计单元，分别汇总指标出现篇次、场景分布情况及指标关系，生成指标出现篇次统计图、场景累计分布图及场景化指标关系热力图。

2.3 三类智能体的输入、输出与协同关系

术语规范化智能体以文献分段文本作为输入，输出规范化命名的网络指标词表；其核心目标是将跨文献中出现的术语别名、缩写及近义表达映射为统一的指标名称，例如将latency与delay在对应上下文环境中统一归为同一类时延指标。场景标注智能体以摘要、引言及问题定义等高信息密度文本为主要输入，输出统一的场景标签；该任务采用预先定义的场景标签集及示例约束提示词，当多场景术语并存时，依据研究目标所聚焦的最主要应用场景赋予主标签，以提高场景标注的一致性。

关联挖掘智能体以包含指标描述的原始文本语句为输入，输出指标对间的语义方向及强度；与仅统计指标共现次数的传统方法不同，该智能体重点关注“提升某一指标是否伴随另一指标上升、下降或存在权衡关系”的语义表述，并将其量化为取值范围在-1至+1之间的带符号相关系数。三类智能体之间并非简单并行且彼此孤立，而是通过共享的规范化指标词表及场景标签实现协同工作：术语规范化结果为关联挖掘智能体的关系挖掘任务提供统一的指标名称支撑，场景标注结果为后续场景化统计分析热力图生成提供基础条件，关联挖掘结果则在场景标签的约束下汇总形成场景级知识成果。

为了验证LLM自动化挖掘的可靠性，本文建立了“微观人工抽样”和“宏观交叉验证”的双重验证机制。在微观层面，本文随机抽取若干篇格式相对规范的论文进行人工标注，并将人工标注结果作为基准，用于评估两个模型的提取结果。在宏观层面，引入不同源的LLM对相同语料进行交叉验

证，通过对比两类模型在关系量化任务上的评分分布与关联判定，评估模型间的一致性，从而验证数据挖掘结果的可靠性。

3 网络指标体系的文献挖掘结果与分析

本节将对基于LLM的文献挖掘所获得的网络指标进行系统性梳理分析。首先，通过分析双LLM挖掘结果的相关系数一致性与微观人工抽样的指标提取精确度，论证底层数据的可靠性与准确性；其次，从整体维度考察各类指标的关注度分布规律，勾勒出网络作为信息基础设施的共性能力边界；进而，在引入应用场景约束的前提下，深入剖析特定垂直网络(如数据中心、边缘计算及物联网等)的核心指标聚焦集合，并定量探究不同场景内基础指标间的异构耦合模式与内在权衡机制。

3.1 数据的可靠性分析

为评估本文提出的自动化挖掘流程的合理性和结果的可靠性，本研究在“微观人工抽样”层面随机选择了10篇基准论文，在“宏观交叉验证”层面选取了DeepSeek和GLM-4两个不同源的大语言模型开展对比验证，从两个层面共同对挖掘结果进行可靠性分析。

在微观层面，通过对[18,19]等10篇人工标注的基准文献进行核对，结果显示，大语言模型正确提取的指标数目为77个，错误或冗余提取11个，遗漏提取21个。由此计算得出，指标提取任务的精确率(Precision)达到87.5%，召回率(Recall)为78.6%，F1分数为82.8%。该结果表明，尽管在部分细粒度指标或特定领域专属称谓指标的提取中仍存在少量遗漏，但总体上所提自动化流程在网络指标抽取任务中具有较高可信度。

在宏观层面，DeepSeek与GLM-4在指标识别任务上的重合度超过85%，表明两种异源模型对网络指标边界的判定具有较高一致性。进一步地，本文提取了两模型共同识别的前200个高频网络指标对，并绘制了如图2所示的两模型指标关联评分一致性散点图。散点图展现出数据点高度聚集于45°对角线附近的分布特征，Pearson相关系数达到0.91。这一结果表明，两个大模型在强正相关、强负相关及弱相关判定上具有较高的一致性，说明关系量化结果不依赖单一模型的偶然判断，具有较好的模型间稳健性。

3.2 网络指标集及其篇次分析

依托文本预处理与LLM多智能体协同框架，本研究从583篇目标文献中自动识别并规范化了3978个独立网络指标，其中出现篇次占比(指标出现的

篇数占总文献篇数的比例)超过1%(即在6篇及以上的文献中出现)的指标共138个。若按照这些指标出现篇次从大到小排序,则指标位次与所出现篇次间的关系如图3所示。位列第1位的指标为吞吐量,共计在461篇文献中出现,篇次占比79.1%。随后依次为端到端时延和误包率,出现篇次分别为435和347,占比分别为74.6%和59.5%。丢包率、误码率、带宽、资源利用率等指标也获得了较高关注,篇次占比均超过35%。这一分布态势表明,高吞吐、低时延和高可靠传输能力是网络研究最受关注的维度,资源效率与能耗等与网络运营经济性密切相关的指标也日益成为研究热点。表1则详述了排名前15位的核心指标数据及其定义。

值得注意的是,网络指标体系呈现出极其显著的重尾分布(Heavy-tailed distribution)特征。在全部指标项中,出现篇次达到6篇及以上的指标仅占

3.5%,而出现篇次超过100篇的核心指标占比仅为0.3%。该分布表明网络指标体系呈现出“核心集中、长尾稀疏”的结构,即少量高频指标(如时延、丢包率、带宽)构成了跨文献研究的共识性核心,而绝大多数低频指标反映了网络研究在特定场景方向上的分化。

3.3 网络场景及其对应的指标集分析

高频指标反映了网络作为信息通信基础设施的核心属性,而长尾指标则映射了不同场景下特异化的性能关注。若忽略应用场景而仅从整体层面考察网络指标,将难以捕捉到特定场景下的指标子集及设计逻辑,因此有必要深入分析不同网络场景下的指标集差异并进行针对性讨论。不同场景所对应的网络技术如数据分组格式、转发策略、路由机制等各有区别,这些技术差异进一步塑造了各场景独特的指标关注重点。为了揭示场景研究热点和不同场景中所关注指标集的差异,本文对全部583篇文献进行了场景标注,统计结果如图4所示。

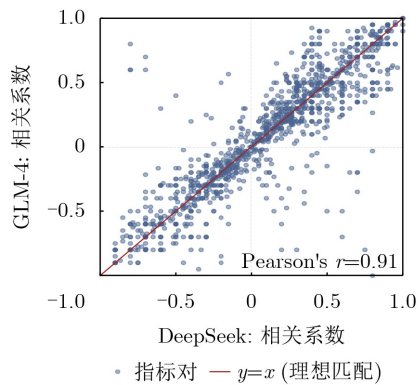


图2 DeepSeek与GLM-4指标关联评分一致性散点图

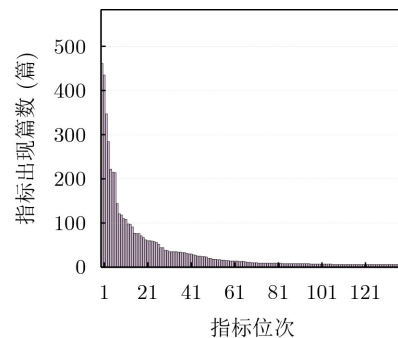


图3 网络指标出现篇次排序及其分布

表1 出现篇次前15位的核心网络指标

网络指标	出现篇次	篇次占比	定义
吞吐量	461	79.1%	单位时间内成功传输的数据总量
端到端时延	435	74.6%	数据从源端到目的端传输所需的总时间
误包率	347	59.5%	传输过程中错误数据包占总数据包的比例
丢包率	284	48.7%	传输过程中丢失数据包占总数据包的比例
误码率	222	38.1%	传输过程中错误比特数占总比特数的比例
带宽	215	36.9%	网络信道能够传输数据的最大速率
资源利用率	214	36.7%	网络资源(如带宽或处理能力)被使用的百分比
排队时延	144	24.7%	数据包在队列中等待处理的时间
抖动	121	20.8%	数据包到达时间间隔的变化量
往返时延	118	20.2%	数据从源端发送到目的端并返回所需的时间
信噪比	110	18.9%	信号功率与噪声功率的比值,用于衡量信号质量
传输时延	108	18.5%	数据从发送端到接收端传输所需的时间
跳数	98	16.8%	数据包从源到目的经过的网络节点数量
功耗	97	16.6%	网络设备在运行过程中消耗的功率
流完成时间	91	15.6%	一个数据流从开始传输到完成所需的总时间

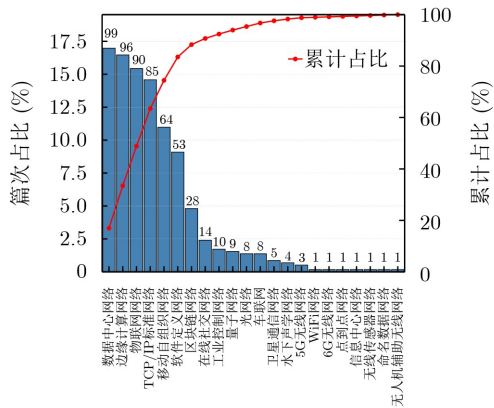


图 4 不同网络场景出现的篇次及累计占比

如图4所示，不同网络场景的关注度呈现高度差异性。出现篇次最高的为数据中心网络，共计99篇文章，占比17.0%。随后分别是边缘计算网络(96篇，16.5%)和物联网(90篇，15.4%)。这三类场景合计占据总文献数的48.9%，与当前信息服务中“云-边-端”的算力与网络协同架构高度契合。IP网络紧随在这三类网络场景之后，共计85篇文章，占比14.6%，表明IP协议作为互联网的核心技术，仍具有旺盛的研究活力。此外，移动自组织网络出现64篇(占比11.0%)，软件定义网络出现53篇(占比9.1%)，前者主要对应拓扑高度动态的无线组

网场景，后者则侧重网络控制与转发机制的可编程化。以上六类网络场景总计占比达到83.5%，是当前网络研究的主要热点。

网络场景还呈现出明显的长尾分布特征，尽管其余场景的篇次占比仅为16.5%，但类型却十分丰富，典型的区块链网络(28篇，4.8%)、工业控制网络(10篇，1.7%)、卫星网络(5篇，0.9%)等，另有一些网络场景在所选文献中仅有1篇文献涉及(如命名数据网络)。尽管这些场景的相关研究占比比较低，但其网络指标却呈现出独有的特征，如区块链网络侧重块传播时延等专有指标。因此，伴随着网络应用场景的多样化，网络指标体系也同步呈现出面向垂直场景的深度分化。

不同网络场景关注的指标集存在差异，典型结果如表2所示。数据中心网络高度关注海量数据传输能力，吞吐量、端到端时延和流完成时间分别占对应场景文献数的86.9%、79.8%和63.6%。边缘计算网络更关注资源约束下的计算卸载实时性，端到端时延和资源利用率是最受关注的指标。物联网场景除吞吐量、端到端时延等传统指标之外，还关注能耗与网络生存周期，体现出物联网终端设备资源受限的特征。类似地，移动自组织网络关注射频信号的信干噪比指标，软件定义网络关注实例扩展时

表 2 分场景的网络核心指标特征

网络类型	重点关注指标(篇次占比)	场景特性描述
数据中心网络	吞吐量(86.9%)	大规模流量调度能力基准
	端到端时延(79.8%)	任务响应时延的关键约束
	流完成时间(63.6%)	拥塞控制的效能表征
边缘计算网络	端到端时延(79.2%)	包含传输和计算的总处理时延
	资源利用率(68.8%)	跨设备和服务器的负载分配优化
	通信开销(41.7%)	参数服务器与客户端之间通信所消耗的资源
物联网	吞吐量(76.7%)	单位时间内成功传输的传感器数据量
	端到端时延(66.7%)	传感器、执行器与边缘/云端反馈控制的实时性约束
	能耗(56.7%)	网络设备在处理和传输数据时的能量消耗
	网络生存周期(31.1%)	网络在能量耗尽前的持续运行时间
IP网络	采样率(27.8%)	传感器支持的每秒采集信号的次数
	丢包率(67.1%)	传输过程中丢失的数据包占总发送数据包的比例
	抖动(47.1%)	数据包到达时间间隔的变化量
移动自组织网络	吞吐量(76.2%)	动态路径的有效传输速率
	误包率(65.1%)	高扰动环境下的传输可靠性衡量
	信干噪比(44.4%)	高扰动环境下的无线信道稳定性表征
软件定义网络	吞吐量(90.6%)	VNF实例处理网络流量的能力
	实例扩展时延(81.1%)	根据负载动态调整规模的时间开销
区块链网络	吞吐量(85.7%)	网络单位时间处理的消息数量
	块传播时延(42.9%)	分布式共识效率瓶颈

延指标, 而区块链网络关注块传播时延指标。综上, 网络指标体系与应用场景之间存在强依赖关系。

3.4 区分场景的网络指标关联分析

网络应用场景与网络指标间关联密切。本节提取不同场景中的指标集, 并基于LLM分析这些指标间的关联关系。对于给定的两个指标(如误码率和丢包率), 如果文献中出现关于这两个指标相关性的描述, 则由智能体自动给出-1到+1区间内的实值相关系数, 其中+1或-1表示严格线性正相关或负相关, 0表示无关, 绝对值介于0与1之间表示不同程度的相关性, 具体数值由LLM根据语句中的形容词、副词等语言线索体现的语义强度进行判定。本文对同一类场景内所有文献中关于同一指标对的相关表述取算术均值, 作为该场景中相应指标对的场景级相关系数。受限于篇幅, 本节主要讨论6类

网络场景, 其中4类为热点研究场景, 2类为非热点研究场景, 这些场景具有较强的代表性。

针对四类最主流的网络场景, 即数据中心网络、边缘计算网络、物联网和IP网络, 本文分别对各场景文献中出现篇次占比大于15%的指标进行关联分析, 绘制出相应场景网络指标关联关系的矩阵热力图, 如图5所示。相关系数以矩阵形式呈现, 并通过不同颜色的方格表示其取值, 红色和蓝色分别对应正相关和负相关, 颜色深浅对应相关系数的绝对值。

在数据中心网络中, 吞吐量、带宽、传输速率与链路容量之间的强正相关关系, 以及吞吐量与流完成时间、排队时延、丢包率之间的强负相关关系, 反映了该场景下围绕高效流量承载所形成的核心运行机制。更具体地, 远程直接数据访问

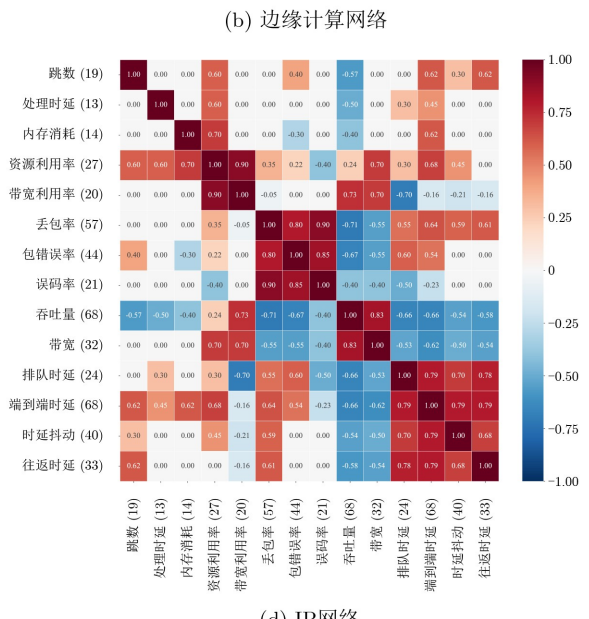
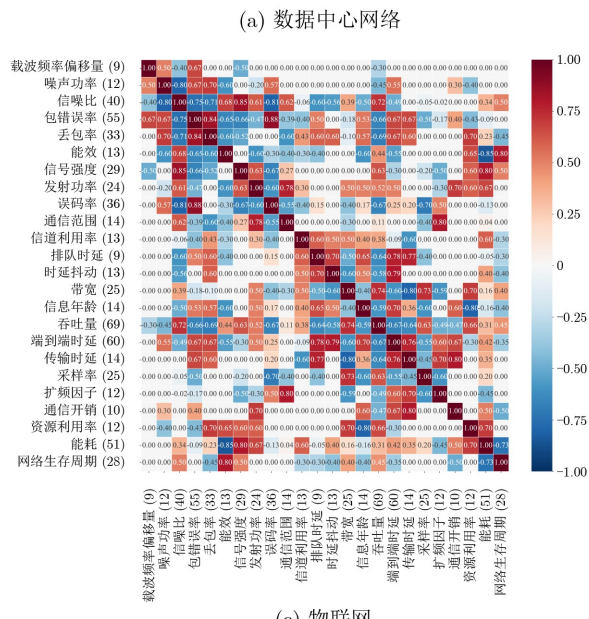
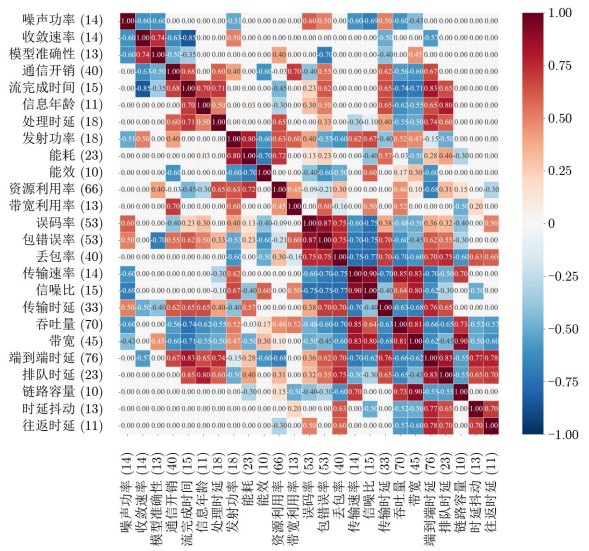
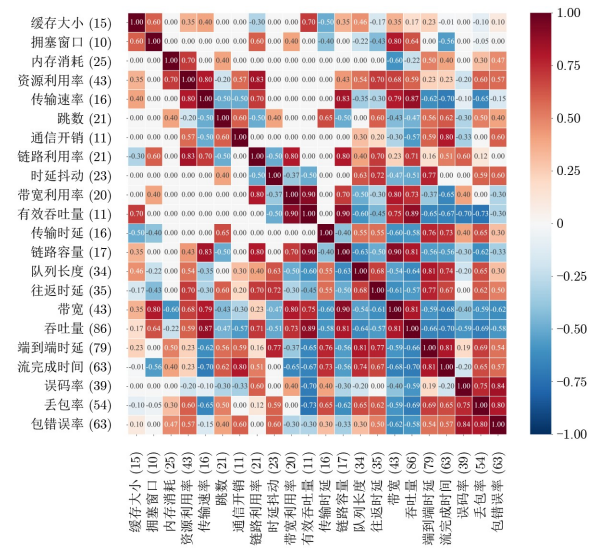


图5 热点研究网络场景指标热力图

(Remote Direct Memory Access, RDMA)传输、显式拥塞通知(Explicit Congestion Notification, ECN)拥塞反馈及队列管理机制共同决定了吞吐量、流完成时间和排队时延间的两两作用关系,例如高吞吐量能有效缩短流完成时间,但当吞吐量逼近网络容量时,排队时延可能显著增加,而排队时延的累积最终又会反向延长流完成时间。

在边缘计算场景中,端到端时延与吞吐量、资源利用率之间均呈显著负相关(ρ 分别为-0.66和-0.68),表明边缘侧计算卸载的实时性保障与系统总容量、资源效率之间存在核心权衡关系。服务迁移与任务卸载算法决定了端到端时延、资源利用率和通信开销之间的相互作用关系。通常情况下,合理的资源整合与边缘协同有助于降低端到端时延;但当资源利用率接近饱和,或系统为维持高利用率而频繁触发卸载与迁移时,通信开销会增加,并与高负载状态叠加,进而推高端到端时延。

物联网场景中,“能耗—网络生存周期”的负相关关系反映了物联网节点能量受限这一基本约束;同时,采样率的提高通常会增加数据产生速率与通信负担,因此其与能耗、网络生存周期之间存在间接耦合关系。该结果表明,物联网场景的核心矛盾并非仅在于传输性能,而是感知质量、通信负担与能量约束之间的长期平衡。

IP网络中,丢包率与吞吐量的强负相关以及抖动与时延类指标的联动,集中体现了传统分组网络中的拥塞传播效应。结合主动队列管理(Active Queue Management, AQM)与拥塞控制机制可以看出,链路竞争加剧或缓存管理不及时会导致丢包

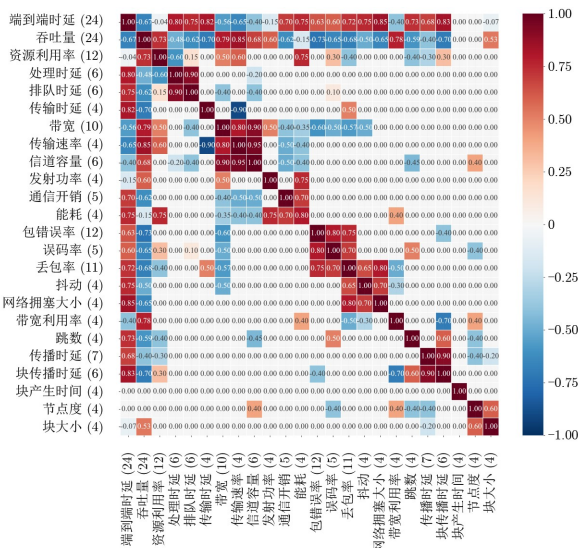
与时延抖动同时上升,并进一步降低有效吞吐量。与数据中心和边缘计算场景相比,IP网络更侧重端到端传输稳定性,其瓶颈更多来自链路竞争与缓存管理。

横向比较可知,数据中心与IP网络更侧重传输链路效率及拥塞控制效应,边缘计算与物联网则更强调资源受限条件下的时延或能耗权衡。这表明,相同指标名称在不同网络场景中可能对应不同技术机理,因此网络指标分析必须结合具体场景进行阐释。

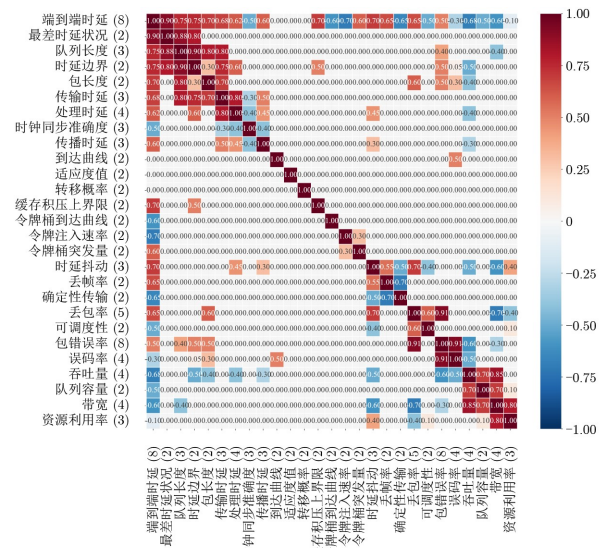
上述结论在非热点研究场景中同样适用,下文以区块链网络和工业控制网络为例开展分析,这些非热点研究场景同样具备与其技术机制相适配的专有指标组合及关联特征,如图6所示,这佐证了本文关于“场景差异化”的总体判断。

在区块链网络中,网络更侧重于共识传播路径与块确认效率。从热力图关联矩阵可以看出,除了传统的吞吐量和端到端时延外,“块传播时延”成为了关键的专有指标。数据表明,吞吐量与块传播时延呈显著负相关,而端到端时延与块传播时延呈显著正相关,这一特征凸显了区块链网络对共识传播、块确认与链上同步效率的高度关注,对传播路径和确认等待时间更为敏感,这与传统的IP网络或数据中心网络单纯以分组拥塞为主的机制存在显著差异。

在工业控制网络中,网络则更强调最差时延状况、确定性时延及控制稳定性。热力图分析揭示,该场景不仅关注丢包率、误包率等传统指标,还高度关注确定性传输和最差时延状况。具体而言,丢



(a) 区块链网络



(b) 工业控制网络

图 6 非热点研究网络场景指标热力图

包率与误包率之间呈现强正相关，且端到端时延与最差时延状况亦呈强正相关。这种强关联模式反映出工业控制网络的核心目标是维持闭环控制系统的稳定性，因此其设计理念更强调保障最坏情形下的时延边界、时钟同步精度与确定性传输，而非单纯追求平均吞吐量的最大化。

由上述统计分析结果可知：第一，网络指标呈现“少量共性指标+大量长尾专有指标”的分布特征，这表明网络研究既关注共性的基础性能指标，也保留面向特定场景的专有评价需求；第二，网络指标集合呈现出高度的场景依赖性，不同应用场景所关注的核心指标集存在显著差异；第三，不同场景中核心指标间的关联关系与相互影响各不相同，这深刻反映了不同网络在内在运行机制与设计约束上的本质区别。

4 网络指标体系的维度讨论与启示

文献挖掘结果表明，尽管网络指标数量庞大且呈现重尾分布，但其关注维度可归纳为三类：服务质量(Service, S)、资源复用成本(Multiplexing, M)以及场景适配重构能力(Versatility, V)^[20]。三者共同构成在评价语义上相互独立且可组合的三维分析框架，如图7所示。其中，S维度主要刻画吞吐量、时延、可靠性等业务保障类指标，用于回答“服务是否满足需求”这一核心问题；M维度主要刻画能耗、频谱、算力及链路资源占用等成本约束指标，用于回答“代价是否可接受”的关键问题；V维度则衡量网络在跨场景条件下的可配置、可编程、可重构能力，用于回答“网络能否以较低重构成本适配新场景与新业务”这一问题^[20]。

各网络场景中的指标体系可投影至S-M平面，反映特定网络场景下的经典工程权衡逻辑。以数据

中心网络为例，吞吐量、端到端时延更倾向于S维度，链路容量与资源利用率更倾向于M维度，而队列缓存既反映资源配置状态，又直接影响数据传输质量，兼有S与M维度诉求；在物联网场景中，吞吐量、误包率倾向于S维度，能耗、频谱利用率倾向于M维度，而数据聚合度反映了感知数据在网络内的融合处理程度，高聚合有助于减少传输次数进而降低能量消耗，但可能增加处理时延和丢包，同时反映了S与M两个维度的需求。现代网络日益注重可编程和可定制能力，从而通过同一物理基础设施支持不同逻辑协议，并适配不同网络场景，这体现了V维度的核心内涵。以软件定义网络或云原生网络为例，V维度的核心诉求在于依托网络的高可编程性与动态配置能力，精准匹配不同场景下的S-M需求。因此，S-M-V三维视图有助于识别不同网络场景下应优先关注的核心指标，并为网络工程中的系统性权衡提供理论支撑。

通过同一物理基础设施高效支撑异构业务及未来新业务，已成为网络技术发展的重要趋势。软件定义网络(SDN)^[6]、多模态网络^[21]、网络切片^[22]、可编程数据平面^[23]、网络功能虚拟化(Network Functions Virtualization, NFV)^[24]、云原生网络^[25]以及面向多网络模态的柔性报文处理机制^[26]等主流技术路线，本质上均致力于通过软硬件解耦与资源抽象，降低网络建设与运营成本，同时充分复用已有基础设施。在此背景下，如何系统性地指导网络在服务质量、资源复用与业务包容性三者之间做出合理权衡，成为亟待解决的关键问题。S-M-V三维指标体系恰好为此提供了分析框架与决策依据。

具体而言，当一个业务请求进入网络时，系统首先对业务进行服务语义解析，识别其核心目标(如低时延、高可靠、高带宽或低成本)以及对关键

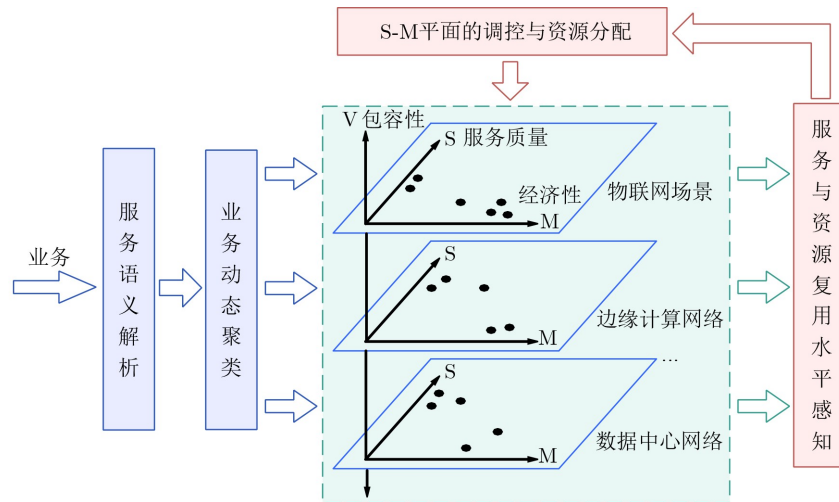


图 7 网络指标体系的S-M-V三维框架示意图

性能参数的敏感程度。基于解析结果，网络将具有相似S-M权衡关系的业务动态聚合到同一个S-M平面内。同一平面内的业务共享相近的服务与资源复用策略，因而可以采用统一的网络协议栈与资源调度机制；不同平面则对应截然不同的S-M特征，需运行差异化的协议及控制逻辑。在此基础上，由人工智能(Artificial Intelligence, AI)驱动的网络控制平面承担两类资源调控任务：平面间调控负责在不同S-M平面之间分配计算、存储与带宽资源，以匹配各平面所承载业务的总体需求；平面内调控则由特定网络协议(如IP的拥塞控制与队列调度)自发完成，负责同一平面内不同业务流之间的细粒度资源竞争与协调。

由于业务请求动态变化，网络拓扑与可用资源亦随时间波动，上述聚合、分配、调度的过程需连续迭代优化。AI引擎通过在线预测流量趋势、实时感知服务语义变化以及测量网络资源消耗，持续判断S-M平面的有效性：当业务需求与现有平面长期匹配不佳时，系统可创建新的S-M平面；当某平面承载业务大幅减少时，则自动缩小其资源配额或将其释放。这一闭环机制使S-M平面随业务演化而动态扩展、收缩、新增或消亡，从而实现网络对多样业务的包容性(V维度)，以及在服务质量与资源经济性之间的自适应权衡。上述过程表明，S-M-V框架是一种可嵌入AI原生网络控制平面的运行逻辑，它明确了从业务语义识别到平面聚合、从资源分配到平面动态演化的完整链路，为网络设计、部署、运维与优化提供指导。

5 结束语

网络指标体系是网络技术与工程实践的基础。本文基于大语言模型构建领域化文献挖掘流程，对网络领域权威期刊《IEEE/ACM Transactions on Networking》2023-2024年刊载的全部文献开展自动化指标提取、术语规范化、场景标注及指标关联量化工作，并通过双LLM交叉验证与人工抽样复检验证结果可靠性。挖掘结果表明，不同应用场景下的网络指标关注集及其关联关系均存在显著差异，即网络指标与应用场景之间存在强耦合特性。基于上述文献挖掘结果，本文进一步提出评价语义上相互独立且可组合的S-M-V三维分析框架，既考虑不同网络场景对业务质量的差异化需求，也兼顾支撑异构业务所需的物理基础设施成本，以及同一物理基础设施对未来新业务的适配重构能力。

未来工作将围绕两个方向展开：一是将当前面向静态文献的离线挖掘模式扩展为可持续更新的网络指标知识库，为网络设计与运维中的动态知识检

索提供支撑；二是在可编程网络、网络切片或多模态网络中进一步验证S-M-V框架的工程可用性，推动其从文献分析工具逐步演进为网络规划与决策支持工具。

参考文献

- [1] WANG Chengxiang, YOU Xiaolu, GAO Xiqi, *et al.* On the road to 6G: Visions, requirements, key technologies, and testbeds[J]. *IEEE Communications Surveys & Tutorials*, 2023, 25(2): 905–974. doi: [10.1109/COMST.2023.3249835](https://doi.org/10.1109/COMST.2023.3249835).
- [2] TANG Xiongyan, CAO Chang, WANG Youxiang, *et al.* Computing power network: The architecture of convergence of computing and networking towards 6G requirement[J]. *China Communications*, 2021, 18(2): 175–185. doi: [10.23919/JCC.2021.02.011](https://doi.org/10.23919/JCC.2021.02.011).
- [3] 任超, 丁思颖, 张晓奇, 等. 6G无线多模态通信技术[J]. *电子与信息学报*, 2024, 46(5): 1658–1671. doi: [10.11999/JEIT231201](https://doi.org/10.11999/JEIT231201).
REN Chao, DING Siying, ZHANG Xiaoqi, *et al.* Wireless multimodal communications for 6G[J]. *Journal of Electronics & Information Technology*, 2024, 46(5): 1658–1671. doi: [10.11999/JEIT231201](https://doi.org/10.11999/JEIT231201).
- [4] OBIODU E and SASTRY N. From ATM to MPLS and QCI: The evolution of differentiated QoS standards and implications for 5G network slicing[J]. *IEEE Communications Standards Magazine*, 2020, 4(2): 14–21. doi: [10.1109/MCOMSTD.001.1800041](https://doi.org/10.1109/MCOMSTD.001.1800041).
- [5] CAMPBELL A, COULSON G, and HUTCHISON D. A quality of service architecture[J]. *ACM SIGCOMM Computer Communication Review*, 1994, 24(2): 6–27. doi: [10.1145/185595.185648](https://doi.org/10.1145/185595.185648).
- [6] KARAKUS M and DURRESI A. Quality of service (QoS) in software defined networking (SDN): A survey[J]. *Journal of Network and Computer Applications*, 2017, 80: 200–218. doi: [10.1016/j.jnca.2016.12.019](https://doi.org/10.1016/j.jnca.2016.12.019).
- [7] KAUSHIK N, KUMAR H, and RAJ V. A systematic review of QoS enhancement techniques in microservices architecture[J]. *Computers and Electrical Engineering*, 2025, 127: 110550. doi: [10.1016/j.compeleceng.2025.110550](https://doi.org/10.1016/j.compeleceng.2025.110550).
- [8] HANEMANN A, LIAKOPOULOS A, MOLINA M, *et al.* A study on network performance metrics and their composition[J]. *Campus-Wide Information Systems*, 2006, 23(4): 268–282. doi: [10.1108/10650740610704135](https://doi.org/10.1108/10650740610704135).
- [9] AL-SHEHRI S, LOSKOT P, NUMANOGLU T, *et al.* Towards taxonomy of telecommunication network metrics[C]. *Proceedings of 2017 European Modelling Symposium (EMS)*, Manchester, UK, 2017: 227–232. doi: [10.1109/EMS.2017.46](https://doi.org/10.1109/EMS.2017.46).
- [10] PENG Cong, HE Yuanzhi, ZHAO Shanghong, *et al.* Integration of data center into the distributed satellite

- cluster networks: Challenges, techniques, and trends[J]. *IEEE Network*, 2023, 37(3): 52–58. doi: [10.1109/MNET.105.2100614](https://doi.org/10.1109/MNET.105.2100614).
- [11] XIAO Yang, SONG Yuqian, and LIU Jun. Collaborative multi-agent deep reinforcement learning for energy-efficient resource allocation in heterogeneous mobile edge computing networks[J]. *IEEE Transactions on Wireless Communications*, 2024, 23(6): 6653–6668. doi: [10.1109/TWC.2023.3335597](https://doi.org/10.1109/TWC.2023.3335597).
- [12] 张卫军, 蒋笑冰, 胡新, 等. 5G-R网络运维质量评价体系研究[J]. *铁道通信信号*, 2025, 61(3): 1–9. doi: [10.13879/j.issn.1000-7458.2025-03.24031](https://doi.org/10.13879/j.issn.1000-7458.2025-03.24031).
- ZHANG Weijun, JIANG Xiaobing, HU Xin, *et al.* Research on quality evaluation system of 5G-R network operation and maintenance[J]. *Railway Signalling & Communication*, 2025, 61(3): 1–9. doi: [10.13879/j.issn.1000-7458.2025-03.24031](https://doi.org/10.13879/j.issn.1000-7458.2025-03.24031).
- [13] 文浩, 林闯, 任丰原, 等. 无线传感器网络的QoS体系结构[J]. *计算机学报*, 2009, 32(3): 432–440. doi: [10.3724/SP.J.1016.2009.00432](https://doi.org/10.3724/SP.J.1016.2009.00432).
- WEN Hao, LIN Chuang, REN Fengyuan, *et al.* QoS Architecture in Wireless Sensor Network[J]. *Chinese Journal of Computers*, 2009, 32(3): 432–440. doi: [10.3724/SP.J.1016.2009.00432](https://doi.org/10.3724/SP.J.1016.2009.00432).
- [14] ZHAI Yunkai, XU Xing, CHEN Baozhan, *et al.* 5G-network-enabled smart ambulance: Architecture, application, and evaluation[J]. *IEEE Network*, 2021, 35(1): 190–196. doi: [10.1109/MNET.011.2000014](https://doi.org/10.1109/MNET.011.2000014).
- [15] TAO Zhenyu, XU Wei, HUANG Yongming, *et al.* Wireless network digital twin for 6G: Generative AI as a key enabler[J]. *IEEE Wireless Communications*, 2024, 31(4): 24–31. doi: [10.1109/MWC.002.2300564](https://doi.org/10.1109/MWC.002.2300564).
- [16] GUO Daya, YANG Dejian, ZHANG Haowei, *et al.* DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning[J]. *Nature*, 2025, 645(8081): 633–638. doi: [10.1038/s41586-025-09422-z](https://doi.org/10.1038/s41586-025-09422-z).
- [17] DAGDELEN J, DUNN A, LEE S, *et al.* Structured information extraction from scientific text with large language models[J]. *Nature Communications*, 2024, 15(1): 1418. doi: [10.1038/s41467-024-45563-x](https://doi.org/10.1038/s41467-024-45563-x).
- [18] CHENG Xia, SHI Junyang, SHA Mo, *et al.* Revealing smart selective jamming attacks in WirelessHART Networks[J]. *IEEE/ACM Transactions on Networking*, 2023, 31(4): 1611–1625. doi: [10.1109/TNET.2022.3224358](https://doi.org/10.1109/TNET.2022.3224358).
- [19] ZENG Yijing, CALVO-PALOMINO R, GIUSTINIANO D, *et al.* Adaptive uplink data compression in spectrum crowdsensing systems[J]. *IEEE/ACM Transactions on Networking*, 2023, 31(5): 2207–2221. doi: [10.1109/TNET.2023.3239378](https://doi.org/10.1109/TNET.2023.3239378).
- [20] 张慧峰, 胡宇翔, 朱俊, 等. 支持网络多模态共生与演化的体系结构及运行逻辑[J]. *电子与信息学报*, 2026, 48(2): 607–617. doi: [10.11999/JEIT250949](https://doi.org/10.11999/JEIT250949).
- ZHANG Huifeng, HU Yuxiang, ZHU Jun, *et al.* Architecture and operational dynamics for enabling symbiosis and evolution of network modalities[J]. *Journal of Electronics & Information Technology*, 2026, 48(2): 607–617. doi: [10.11999/JEIT250949](https://doi.org/10.11999/JEIT250949).
- [21] WU Jiangxing, LI Junfei, SUN Penghao, *et al.* Theoretical framework for a polymorphic network environment[J]. *Engineering*, 2024, 39: 222–234. doi: [10.1016/j.eng.2024.01.018](https://doi.org/10.1016/j.eng.2024.01.018).
- [22] EBRAHIMI S, BOUALI F, and HAAS O C L. Resource management from single-domain 5G to end-to-end 6G network slicing: A survey[J]. *IEEE Communications Surveys & Tutorials*, 2024, 26(4): 2836–2866. doi: [10.1109/COMST.2024.3390613](https://doi.org/10.1109/COMST.2024.3390613).
- [23] ZHANG Kaiyi, SAMAAAN N, and KARMOUCH A. A machine learning-based toolbox for P4 programmable data-planes[J]. *IEEE Transactions on Network and Service Management*, 2024, 21(4): 4450–4465. doi: [10.1109/TNSM.2024.3402074](https://doi.org/10.1109/TNSM.2024.3402074).
- [24] WU Qiang, ZHAI X B, LIU Xi, *et al.* Performance tuning via lean measurements for acceleration of network functions virtualization[J]. *IEEE/ACM Transactions on Networking*, 2023, 31(1): 366–379. doi: [10.1109/TNET.2022.3193686](https://doi.org/10.1109/TNET.2022.3193686).
- [25] MOHAMMADI A and NIKAEIN N. Athena: An intelligent multi-x cloud native network operator[J]. *IEEE Journal on Selected Areas in Communications*, 2024, 42(2): 460–472. doi: [10.1109/JSAC.2023.3336172](https://doi.org/10.1109/JSAC.2023.3336172).
- [26] 朱俊, 徐琪, 张富军, 等. 面向云网融合环境的网络模态报文处理柔性流水线构建机制[J]. *电子与信息学报*, 2026, 48(2): 522–533. doi: [10.11999/JEIT250806](https://doi.org/10.11999/JEIT250806).
- ZHU Jun, XU Qi, ZHANG Fujun, *et al.* Flexible network modal packet processing pipeline construction mechanism for cloud-network convergence environment[J]. *Journal of Electronics & Information Technology*, 2026, 48(2): 522–533. doi: [10.11999/JEIT250806](https://doi.org/10.11999/JEIT250806).
- 徐祺坤: 男, 硕士生, 研究方向为无线通信网络、多模态网络等。
刘娅汐: 女, 副教授, 研究方向为无线通信网络、6G通感一体化网络等。
韩淑娴: 女, 硕士生, 研究方向为人工智能、边缘计算等。
张慧峰: 女, 高级工程师, 研究方向为新型网络体系架构、智算中心网络等。
皇甫伟: 男, 教授, 研究方向为网络体系结构与理论模型、边缘计算等。

Network Metric System and Scenario-Differentiated Analysis Driven by LLM Literature Mining

XU Qikun^① LIU Yaxi^① HAN Shuxian^{①②} ZHANG Huifeng^② HUANGFU Wei^①

^①(School of Computer & Communication Engineering, University of Science and Technology Beijing,
Beijing 100083, China)

^②(Zhejiang Lab, Hangzhou 311121, China)

Abstract:

Objective Network metrics constitute the foundational parameters for network design, operation, and optimization. Existing research, however, predominantly focuses on single scenarios or representative parameters, lacking unified extraction rules and reproducible processes for large-scale, cross-scenario metric analysis. To address terminology ambiguity, scenario heterogeneity, and complex relationship quantification, a reproducible domain-specific literature mining chain driven by large language models (LLMs) is proposed. The objective is to automatically extract and standardize metrics, annotate application scenarios, quantify inter-metric dependencies, and establish an independent, combinable Service-Multiplexing-Versatility (SMV) theoretical framework. Rather than acting as a finalized suite of specific metric calculations, the SMV framework serves as a theoretical conceptual lens to guide multi-objective tradeoffs in architecture design and lifecycle management.

Methods An automated literature mining framework is constructed utilizing a multi-agent LLM architecture (Fig. 1). A comprehensive dataset comprising 583 articles published in the *IEEE/ACM Transactions on Networking* (2023–2024) is analyzed. The methodology relies on three specialized agents: a terminology normalization agent mapping aliases and synonymous expressions into unified metric names; a scenario annotation agent assigning primary application tags based on high-density information sections; and a correlation mining agent identifying the semantic direction and intensity between metric pairs, quantifying these relationships into signed correlation coefficients ranging from -1 to +1. Mining reliability is verified through dual-LLM cross-validation and manual sampling review (Fig. 2).

Results and Discussions The automated extraction identifies 3,978 independent network metrics, with 138 appearing in over 1% of the sampled literature (Fig. 3). The metric frequency distribution exhibits a pronounced heavy-tailed pattern, dominated by foundational parameters such as throughput (79.1%), end-to-end delay (74.6%), and packet error rate (Table 1). Focal metric sets demonstrate strong scenario dependence (Fig. 4). For instance, data center networks prioritize throughput and flow completion time, whereas Internet of Things (IoT) applications additionally emphasize energy consumption and network lifetime constraints (Table 2). Furthermore, inter-metric correlation matrices reveal highly differentiated coupling mechanisms across environments (Fig. 5, Fig. 6). In data center scenarios, throughput correlates negatively with flow completion time and queuing delay. Conversely, edge computing environments exhibit a negative correlation between end-to-end delay and resource utilization, reflecting an inherent tradeoff between real-time computing offloading latency and system resource efficiency.

Conclusions The strong coupling between evaluation metrics and application scenarios necessitates a multidimensional perspective for future network architecture. Based on the extracted scenario-metric dependencies, the SMV analysis framework is conceptualized (Fig. 7). By integrating differentiated service quality requirements (Service), physical infrastructure overhead (Multiplexing), and adaptive reconfiguration agility for emerging services (Versatility), this paradigm enables adaptive resource orchestration in AI-native networks. Future research will focus on evolving the static literature mining pipeline into a dynamically updatable knowledge base and validating the SMV framework's engineering applicability in programmable and multimodal networks.

Key words: Network metric; Large Language Model (LLM); Metric system; Polymorphic network