

面向具身智能与机器人物联网的模型上下文协议研究综述： 从语义控制到实时安全保障

郭帅帅¹, 朱家兵², 叶佳³, 张安邦¹, 倪艺洋⁴

(1. 山东大学控制科学与工程学院, 山东 济南 250062; 2. 浪潮软件集团有限公司, 山东 济南 250101;
3. 重庆大学电气工程学院, 重庆 400044; 4. 江苏第二师范学院计算机工程学院, 江苏 南京 211200)

摘要: 具身智能与机器人物联网 (IoT, Internet of things) 的发展呈现自主化、协同化和多模态融合的趋势, 大量异构具身智能体接入物联网系统, 对网络架构的互操作性和安全性提出了更高的要求。面对传统协议在语义表达与动态上下文适应性上的局限, 引入并系统地阐述了模型上下文协议 (MCP, model context protocol) 作为统一“语义控制平面”的核心作用。提出并详细地构建了“Context→QoS/Sec 编译器”这一创新框架, 旨在将高层任务意图与环境上下文 (包括优先级、时延抖动容忍度、安全敏感等级等) 精确、可验证地映射至底层数据平面 (如 ROS2/DDS、OPC UA/TSN、5G/6G) 的实时服务质量 (QoS, quality of service) 参数与安全侧写。进一步聚焦当前研究面临的深层挑战, 包括代价模型的可识别性、异构约束的同构化、系统可验证性与可解释性以及工程演化中的稳定性与最优性权衡, 并针对性地提出了未来可行的研究路线。

关键词: 具身智能; 机器人物联网; 模型上下文协议; 语义控制平面; Context→QoS/Sec 编译器

中图分类号: TN915.08

文献标志码: A

doi: 10.11959/j.issn.2096-3750.WLW25125

A survey on model context protocol for embodied intelligence and robotic Internet of things: from semantic control to real-time security assurance

Guo Shuaishuai¹, Zhu Jiabing², Ye Jia³, Zhang Anbang¹, Ni Yiyang⁴

1. School of Control Science and Engineering, Shandong University, Jinan 250062, China

2. Inspur Software Group Co., Ltd., Jinan 250101, China

3. School of Electrical Engineering, Chongqing University, Chongqing 400044, China

4. School of Computing Engineering, Jiangsu Second Normal University, Nanjing 211200, China

Abstract: The development of embodied intelligence and the robotic Internet of things (IoT) is characterized by increasing autonomy, collaboration, and multimodal fusion. The large-scale integration of heterogeneous embodied intelligent agents into IoT systems poses higher requirements for interoperability and security in network architectures. In this context, the introduction of a unified model context protocol (MCP) to standardize device capability descriptions and task in-

收稿日期: 2025-08-12; 修回日期: 2026-05-18

通信作者: 郭帅帅, shuaishuai_guo@sdu.edu.cn

基金项目: 国家自然科学基金资助项目 (No. 62171262, No. 62471204); 中央高校基本科研业务费资助项目 (No. 2024CDJXY006); 江苏省自然科学基金优秀青年基金资助项目 (No. BK20250160); 江苏省高校自然科学研究重大项目 (No. 24KJA510003)

Foundation Items: The National Natural Science Foundation of China (No. 62171262, No. 62471204), The Fundamental Research Funds for the Central Universities (No. 2024CDJXY006), the Natural Science Foundation of Jiangsu Provincial Excellent Young Scientist (No. BK20250160), the Major Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (No. 24KJA510003)

teractions becomes essential. A comprehensive survey of MCP in the integration of embodied intelligence and robotic IoT was presented in this paper. The core concepts of MCP and its roles in multimodal perception, communication, computation, and control collaboration were introduced. The challenges in standardization, scalability, and cross-domain adaptation were summarized and future research directions were discussed. Studies indicate that MCP has the potential to significantly enhance resource utilization efficiency, task execution coordination, and system security in embodied intelligence and robotic IoT, providing a unified protocol foundation for the next-generation intelligent IoT systems.

Key words: embodied intelligence, robotic IoT, MCP, semantic control plane, Context-to-QoS/Sec compiler

0 引言

随着人工智能、机器人技术和物联网的深度融合^[1]，具身智能体在人形机器人、四足机器人、自主无人系统等领域快速发展^[2]，并在工业5.0、智慧城市、无人仓储、应急救援等场景中展现出广阔的应用前景^[3]。具身智能体通过与环境持续交互，实现感知、认知、决策和执行的闭环自治^[4]，这不仅要求设备具备多模态信息处理和实时决策能力^[5]，还对网络的智能化、自主性、低时延和高可靠性提出了前所未有的挑战。在这种趋势下，机器人物联网逐渐成为支撑具身智能体群体协作、跨域任务执行的重要基础设施^[6-7]。

机器人物联网以多源异构的传感器、执行器、控制器为基础，将机器人、环境感知节点与云-边计算平台互联互通，实现跨设备、跨网络的任务协作与信息共享^[8]。具身智能与机器人物联网的深度融合，虽然带来了巨大的潜力，但也引发了一系列系统性挑战^[9]。首先是互操作性瓶颈：大量异构具身智能体接入物联网系统，对网络架构的互操作性提出了更高的要求。传统的机器人通信协议多针对单一任务和固定拓扑设计，难以适应具身智能体在动态环境中快速上下文切换和跨模态信息融合的需求^[10]。其次是实时性与确定性要求：具身智能体需要具备多模态信息处理和实时决策能力，这对网络的智能化、自主性、低时延和高可靠性提出了前所未有的挑战。最后是安全与隐私风险：开放、自治、多模态的融合特性极大地扩大了攻击面，对设备身份、零信任边界、机密计算和生成式安全提出了更高的要求。此外，异构性、高动态性以及多模态数据的大规模流动，使系统在可扩展性、任务适配性以及资源利用效率方面面临显著瓶颈，尤其是在算力、带宽和能耗优化等方面。

面对上述挑战，引入统一的模型上下文协议

(MCP, model context protocol) 显得尤为关键^[11]。MCP作为一种新型协议框架，旨在统一描述设备的功能能力、输入输出模式及运行上下文，为跨模态、多任务、多域的模式调用与交互提供标准化接口。其被业界视为连接人工智能(AI, artificial intelligence)助手与数据和工具世界的开放标准。MCP的核心作用在于提供强大的标准化与互操作性。2024年底，Anthropic公司率先公开了MCP标准，旨在弥补AI代理与外部数据源及工具之间连接的空白，并定义了将现有数据源(如文件系统、数据库、应用程序接口(API, application programming interface)等)连接到AI工作流的方式。MCP提供了一种标准化的方式^[12]来构建自定义集成，并允许用户在不同应用之间灵活切换并携带上下文信息。这种标准化对于解决具身智能体间的语义鸿沟至关重要。

在具身智能与机器人物联网的融合场景中，MCP不仅能够实现异构智能体之间的语义互操作，还能够基于上下文信息动态地选择最优的任务执行路径，从而显著提升系统的可扩展性与任务执行效率^[13]。这种动态决策能力是传统静态协议所不具备的。MCP的出现也发出了强烈的生态信号：2025年，多家大型技术公司(如微软、谷歌、OpenAI、阿里巴巴)已将MCP纳入“面向代理与开放网络”的协议化路线图中，预示着MCP将从“概念验证”阶段迅速迈向“系统基础设施”阶段。微软在Build大会上提出的“Agentic Web/NLWeb 每个端点即MCP Server”设想，更是将网站与应用转化为可被代理直接调用的能力端点，为机器人获取领域知识、业务工具与数据资产搭建了标准化入口。

目前，国内外学者在MCP相关研究上主要集中在3类方向：一是协议架构设计与标准化^[14]，包括接口规范、上下文建模、语义描述等；二是MCP在多模态感知融合与边缘-云协同计算中的应用，包括

跨模态对齐、协同推理和资源优化调度；三是 MCP 在安全与隐私保护中的探索^[15]，包括访问控制、可信执行和加密通信等。然而，现有研究多针对单一技术环节，缺乏对 MCP 在具身智能与机器人物联网融合中的系统性综述与对比分析，尚未形成涵盖架构、算法、应用及安全机制的完整技术体系。

本综述的核心贡献在于围绕“MCP 作为语义控制平面→Context→QoS/Sec 编译器”这一核心主线，系统性地重构 MCP 在具身智能与机器人物联网融合中的技术脉络。具体而言，本文深入剖析了 MCP 作为“语义控制平面”的定位，阐明了其如何解耦底层协议差异，实现“上下文-能力-意图”的统一描述。MCP 的价值并非在于取代现有的实时传输协议，而在于其作为语义抽象层，能够显式化“任务-模型-资源”的语义边界。这种语义绑定替代了传统的接口绑定，使上层应用能够以统一的高层次语义来编排任务，而无须关心底层异构网络的具体实现细节。这种解耦极大地提升了系统的可扩展性和灵活性，降低了集成成本，是实现大规模、动态、异构具身智能体协作的关键。本文形式化定义“Context→QoS/Sec 编译器”，详细阐述其将抽象语义需求精确映射至实时服务质量（QoS, quality of service）与安全侧写（DDS QoS、时间敏感网络（TSN, time-sensitive networking）时隙、5G/6G 切片、可信执行环境（TEE, trusted execution environment）约束等）的机制与优化问题。该编译器将 MCP 描述的任务意图与环境上下文（包括优先级、时延抖动容忍度、安全敏感等级等）作为输入，生成跨机器人操作系统 2/数据分发服务（ROS2/DDS, robot operating system 2/data distribution service）、开放平台通信统一架构/时间敏感网络（OPC UA/TSN, open platform communications unified architecture/time-sensitive networking）与 5G/6G 的实时 QoS 参数与安全侧写。通过形式化建模，它将原本模糊的“需求”转化为可量化、可优化的“参数”，从而实现系统性能（时延、可靠性、吞吐量、可扩展性等指标）与安全性的联合优化，而非传统上将两者独立考量或牺牲另一方性能。这种自动化和形式化转换机制是实现具身智能系统在动态环境中自适应性能和安全保障的关键。最后，聚焦当前研究面临的核心挑战，包括代价模型的可识别性、异构约束的同构化、系统可验证性与可解释

性、工程演化中的稳定性与最优性权衡，针对性地提出了未来可行的研究路线与落地实施途径。

1 具身智能与机器人物联网：演进、挑战与融合需求

具身智能与机器人物联网的融合是当前人工智能和机器人领域极具发展前景的方向之一。这一融合不仅拓宽了智能系统能力的边界，也带来了前所未有的复杂性和挑战，促使学术界重新思考传统网络和协议的设计范式。

1.1 具身智能体的概念与前沿发展

具身智能（embodied AI）的核心在于强调智能体通过其物理身体与真实世界环境进行持续的感知、交互、学习和行动，从而实现从感知到认知、决策再到执行的完整闭环自治^[16]。这种范式与传统的、基于抽象数据处理和符号推理的人工智能形成鲜明对比，它更侧重于物理交互、多模态感知和运动在实现复杂多目标任务中的关键作用。具身智能的这种“物理性”需求，是其对底层通信协议和计算架构提出根本性变革需求的原因。

具身智能体应用领域示意如图 1 所示，具身智能体在多个领域展现出了巨大的应用潜力。在工业 5.0 与智能制造领域，具身智能体，特别是协作机器人和自主移动机器人（AMR, autonomous mobile robot），能够显著地提升生产效率和灵活性^[17]。具身智能体协作辅助如图 2 所示，在智慧城市与公共服务领域，具身智能体可用于环境监测、基础设施巡检和应急救援等任务。在医疗与康复领域，具身智能机器人能够辅助手术、提供康复训练，甚至实现远程医疗服务。

大模型驱动的具身智能^[18]正成为研究热点。大型语言模型（LLM, large language model）和视觉语言动作（VLA, visual language action）模型被应用于机器人控制、路径规划和智能决策，极大地提升了人机交互的自然性和自主性。LLM 能够显著增强机器人的自然语言理解和生成能力，使其能够处理复杂的自然语言输入，并提供多样化和个性化的交互体验。然而，大模型在具身智能中的应用也带来了新的挑战^[19]。尽管 LLM 在机器人领域潜力巨大，但其训练和部署需要庞大的计算资源和数据，这对于资源受限的机器人平台而言是一个严峻的挑战。基础模型在真实物理环境中的性能仍有不足，需要数十亿次的物理交互数据进行训练，而非仅仅依赖标

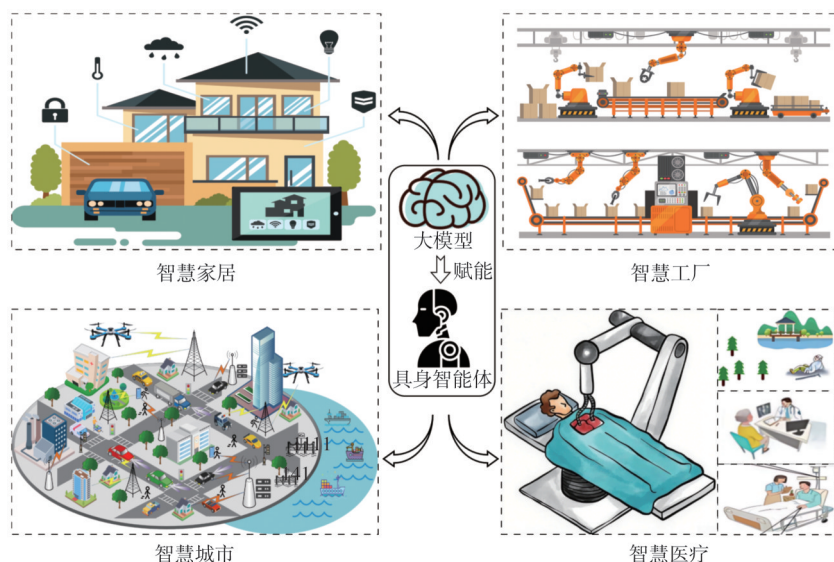


图1 具身智能体应用领域示意

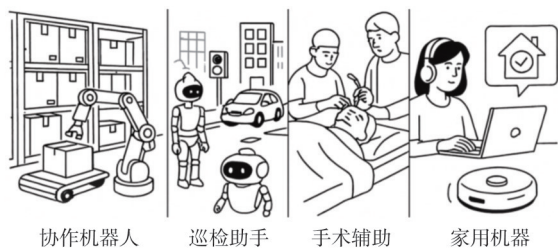


图2 具身智能体协作辅助

注图像。这种计算密集性与机器人实时、低功耗的需求之间存在内在矛盾，这进一步突显了MCP作为语义控制平面的重要性。MCP可以在高层编排大模型的能力调用，同时将低层实时性约束传递给边缘计算和通信层，从而实现资源的高效利用。

1.2 机器人物联网的架构与特性

机器人物联网（IoRT, Internet of robotic things）是连接具身智能体与数字世界的桥梁^[20]，其核心在

于构建一个高度互联互通的生态系统，机器人物联网的端边云架构如图3所示。IoRT的架构设计旨在克服传统机器人系统在异构性、动态性和可扩展性方面的局限。IoRT的基石在于多源异构设备的互联互通。它将各种类型的机器人（包括其集成的传感器和执行器）、分布式的环境感知节点（如智能摄像头、温度传感器、激光雷达）以及强大的云-边计算平台紧密互联。这种互联形成了一个复杂且高动态的网络拓扑，其中涉及多种通信技术和协议的协同工作。在IoRT架构中，云-边计算平台协同扮演着至关重要的角色^[21]。系统通常采用分层架构，将计算任务智能地分布在云端和边缘端，以平衡实时性、资源效率和复杂性。在网络边缘，即数据源附近进行数据处理和智能分析，以满足具身智能体对低时延、高带宽和隐私保护的严苛需求。其中，

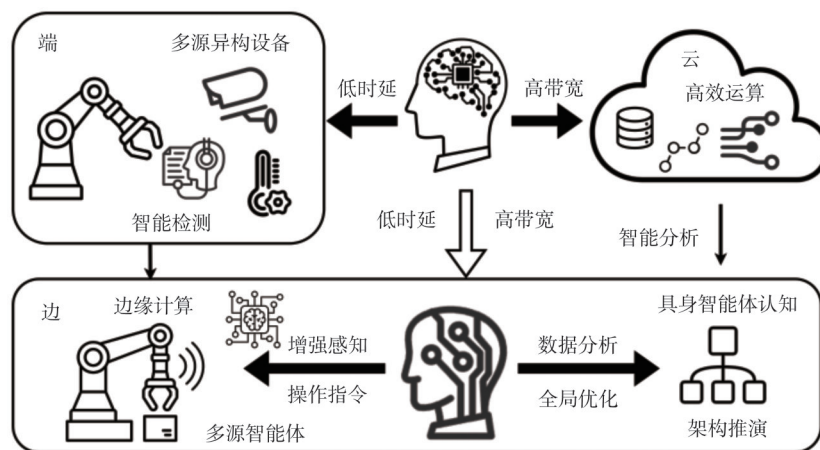


图3 机器人物联网的端边云架构

边缘计算将云的计算、存储和网络能力下沉到网络边缘,使数据能够被实时处理和响应,从而避免了数据回传云端带来的高时延和带宽消耗。这对于时间敏感的机器人应用^[22](如图像处理 and 路径规划)至关重要。云端提供大规模存储、弹性计算资源和复杂模型训练能力,主要用于非实时或计算密集型任务。云端可以存储海量数据,进行复杂的数据分析、模型训练和全局优化,为边缘智能提供强大的后台支持。这种边云协同的分工模式,使边缘设备能够专注于时间敏感的本地控制和初步数据处理,而云端则负责更复杂的全局规划、模型重训练和知识共享。因为任务需要在边缘和云之间无缝迁移和协作,这种模式对协议的互操作性和上下文感知能力提出了更高要求。

IoRT 的另一个关键特性是任务协作与信息共享。其目标是实现跨设备、跨网络的任务协同执行与信息高效共享^[23]。例如,在复杂的物流场景中,多台自主移动机器人可以协同搬运大型货物;在智能工厂中,不同传感器网络可以共享环境数据,以构建全面的态势感知。然而,这种协作并非没有挑战。系统的异构性、高动态性以及多模态数据的大规模流动,使 IoRT 在协议互操作性、任务适配性、实时响应以及安全可信性方面面临显著的瓶颈。特别是,多智能体系统在可伸缩性、异构环境中的适应性以及学习机制方面仍存在诸多开放挑战,需要更先进的协调机制来解决。

1.3 融合带来的系统性挑战

具身智能与机器人物联网的融合,虽然孕育着巨大的变革潜力,但也暴露了现有技术栈的深层局限性,并带来了多方面的系统性挑战。

1) 互操作性瓶颈:异构智能体与设备间的语义鸿沟

具身智能体和 IoRT 系统通常由来自不同制造商、采用不同通信协议(如 ROS2、DDS、消息队列遥测传输协议(MQTT, message queuing telemetry transport)、OPC UA)且具有不同能力描述方式的异构设备组成。这些设备之间缺乏统一的语义理解和交互标准,导致数据孤岛和系统集成复杂性极高。传统的协议互操作性通常停留在语法和格式层面,而无法解决深层次的语义理解问题。例如,一个传感器报告“温度高”,另一个机器人如何理解这个“高”意味着需要采取何种行动,这需要对

“温度高”的“上下文”和“意图”有统一的理解。这种语义鸿沟是实现真正“智能”协作而非仅仅“连接”的根本障碍。语义鸿沟导致系统难以实现真正的即插即用,新设备的集成需要大量定制开发,极大地限制了 IoRT 的可扩展性和灵活性。

2) 实时性与确定性要求:动态环境中的低时延与高可靠性

具身智能体在与物理环境交互时,需要极低的感知-决策-执行环路时延和高确定性^[24]。例如,机器人抓取操作、自动驾驶中的避障,任何微小的时延或数据丢失都可能导致任务失败,甚至引发安全事故。传统的机器人通信协议多针对单一任务和固定拓扑设计,难以适应具身智能体在动态环境中的快速上下文切换和跨模态信息融合需求。缺乏确定性的实时通信和计算能力,将严重制约具身智能体在复杂、动态、安全关键场景中的应用,使其无法满足工业自动化、医疗手术等领域对高可靠性的要求。

3) 安全与隐私风险:开放、自治、多模态带来的攻击面扩大

具身智能与 IoRT 系统的开放性、自治性以及多模态数据流的特性,显著地扩大了潜在的攻击面^[25]。从边缘设备到云端平台,链路性风险无处不在,包括设备身份伪造、越权访问、数据泄露、模型篡改、拒绝服务攻击等。仅仅依靠传统的接入鉴权机制,不足以覆盖“会话内越权、模型侧数据渗漏与链路侧窃听”等更深层次的攻击。安全漏洞可能导致物理世界的破坏、敏感数据的泄露以及对智能体行为的恶意控制,严重威胁系统可靠性、用户隐私和公共安全。安全性不再是事后补丁,而是系统设计的内生属性。

4) 可扩展性与任务适配性:快速上下文切换与跨模态信息融合

具身智能体需要在动态变化的环境中执行多样化任务,这要求系统能够快速进行上下文切换,并有效地融合来自不同模态的信息。例如,一个服务机器人可能需要根据环境变化,从“引导访客”的任务快速切换到“处理紧急情况”,并在此过程中无缝融合视觉、语音和传感器数据。传统系统由于其僵化的设计和缺乏上下文感知能力,难以适应这种高动态性和复杂性,系统难以扩展新功能,且在面对未知或突发情况时表现脆弱。

5) 资源利用效率：算力、带宽、能耗的优化

具身智能体通常部署在资源受限的边缘设备上，但其执行的任务（如复杂视觉推理、运动规划）又通常需要大量的计算资源^[16, 23]。同时，无线通信带宽有限，且能耗是移动机器人续航能力的关键约束。如何在有限的资源下，最大化任务达成度，同时满足实时性和能耗约束，是一个复杂的优化问题。低效的资源利用会导致系统性能瓶颈、高运营成本以及移动设备续航能力不足，限制了具身智能体在实际场景中的广泛部署。

1.4 现有协议与中间件的局限性分析

为了理解MCP的引入价值，有必要审视当前具身智能和IoRT领域广泛使用的协议和中间件，并分析其在应对上述挑战时的局限性。

ROS2/DDS：实时 QoS、安全与跨域部署的进展，但语义表达偏底层。

优势：ROS2作为机器人领域的实时标准^[26]，与DDS的结合，在实时QoS方面取得了显著进展，包括截止时间、时延预算、可靠性等参数的精细控制，以及数据分发服务安全提供的认证和加密功能。近期研究表明，DDS在复杂拓扑和高负载下的可调优空间不断增大，为具身智能体在数据平面进行实时通信提供了坚实的基础。例如，ROS2的性能基准测试关注消息传递、内存管理和传输层效率，并致力于优化这些方面的瓶颈。

局限性：尽管ROS2/DDS在时延与确定性方面能提供工程级保证，但它们对任务语义、模型能力与运行上下文的表达仍然偏底层。这些协议更多关注“如何高效可靠地传输数据”，而非“这些数据代表什么语义意图”或“在何种上下文下使用这些数据”。DDS的QoS参数通常需要开发者手动配置，这在动态、复杂的具身智能任务中难以适应，因为任务的实时性需求、环境变化，甚至安全等级都可能动态调整，而DDS本身无法根据这些高层语义变化自动调整底层参数。ROS2/DDS是强大的“数据管道”，但缺乏高层语义抽象能力，无法直接理解任务意图和环境上下文。

OPC UA/TSN：车间级确定性通信的稳定性，但缺乏高层语义抽象。

优势：OPC UA与TSN的结合，正将工业以太网推向时间同步与时隙调度的硬实时领域。这种组合在车间级确定性通信中展现出卓越的稳定性和工

程可移植性，为工业自动化提供了可靠的通信基础。面向嵌入式系统的增强型时间敏感网络（eTSN, enhanced time-sensitive networking）的实现进一步降低了内存占用，并稳定了确定性时延。

局限性：与ROS2/DDS类似，OPC UA/TSN专注于底层确定性传输，例如，通过预先分配时隙来保证确定性。这对于固定、重复的工业任务非常有效。然而，它们对高层任务语义、模型能力与运行上下文的表达能力不足。当具身智能体在动态环境中执行新任务时，固定的时隙分配无法满足其动态变化的实时性需求。OPC UA/TSN缺乏对任务语义的理解，无法根据任务的动态变化进行资源调度。

现有协议和中间件虽然在各自领域表现出色，但它们普遍缺乏一个统一的、高层次的“语义控制平面”来协调异构设备、管理动态上下文、并实现跨层级的性能与安全优化。它们更像是“数据管道”，负责高效传输比特流，但无法理解这些比特流背后的“意图”和“上下文”^[27]。这种局限性导致了系统集成复杂、可扩展性差、难以适应开放动态环境等问题。

具身智能体对“意图理解”和“上下文感知”的需求，是传统“点对点”或“消息总线”模式无法满足的。MCP的引入正是为了填补这一空白，将“语义”提升为协议的核心要素^[27]，从而实现从“接口互联”向“语义协作”的本质跃迁。

2 模型上下文协议：语义控制平面的核心构建

模型上下文协议的出现，标志着具身智能与机器人物联网领域从传统的“设备互联”向“语义协作”迈出了关键一步。它不仅仅是一个新的通信协议，更是一个旨在统一语义表达、实现智能体间高效协作的“语义控制平面”。

2.1 MCP的起源、定义与生态演进

MCP的诞生源于人工智能领域对更高效、更智能的人工智能体与外部世界交互的需求^[28]。2024年底，Anthropic公司率先正式发布了MCP，将其定义为一个开放标准，旨在标准化模型/工具能力暴露与上下文化调用。MCP的核心价值在于弥合AI助手与外部数据和工具世界之间的鸿沟，使LLM能够像“即插即用”的设备一样，高效、安全地连接到各种数据源和工具。MCP定义了连接现有数据源（如文件系统、数据库、API等）到

AI工作流的方式,提供了一种标准化的方法来构建自定义集成,并允许用户在不同应用之间灵活切换并携带上下文信息。

2025年,MCP的生态演进迎来了重要里程碑,微软、谷歌、OpenAI、阿里巴巴等科技公司纷纷加入。微软公司在当年的Build大会上,战略性地将MCP纳入Windows的“Agentic OS”路径,并在Copilot Studio中宣布了通用可用性(GA, general availability)与生态工具链。这一举措清晰地表明,MCP正从“概念验证”阶段迅速迈向“系统基础设施”的早期阶段。微软的“Agentic Web/NLWeb 每个端点即MCP服务器”设想,更是将传统的网站和企业应用转化为可被AI代理直接调用的“能力端点”,为机器人和AI代理获取领域知识、业务工具和数据资产搭建了标准化的入口。这种由大型科技公司推动的标准化和大规模可用性,预示着MCP将在未来的AI代理生态系统中扮演核心角色,其影响力将远超单一研究领域。

MCP的经典架构示意如图4所示,MCP的架构遵循经典的客户端-宿主-服务器(client-host-server)模型^[29]。这种分层设计清晰地定义了MCP在AI应用生态中的职责,即作为AI模型与外部世界交互的标准化桥梁。

MCP宿主:作为AI应用程序的容器和协调者,接收用户请求,并通过MCP寻求上下文访问。它包含编排逻辑,负责创建和管理多个客户端实例,控制客户端连接权限和生命周期,执行安全策略和用户授权,并协调AI/LLM集成和上下文聚合。例如,集成开发环境(IDE, integrated development environment)如Cursor或Claude Desktop可以作为MCP宿主。

MCP客户端:存在于宿主内,是宿主与服务端之间通信的桥梁。它将用户请求转换为MCP可处理的结构化格式,并解析服务器响应、处理错误、验证响应的相关性。一个宿主可以有多个客户端,但每个客户端与一个MCP服务器之间是一对一关系。IBM BeeAI、Microsoft Copilot Studio、Claude.ai等都是MCP客户端的典型示例。

MCP服务器:作为外部服务的封装或中介,通过将用户请求转换为服务器操作,为LLM提供上下文信息和能力。MCP服务器可以访问各种外部系统、工具和数据源,如数据库、客户关系管理(CRM, customer relationship management)、本地文件系统、版本控制系统(如Git),甚至可以包装LLM推理能力本身。服务器通过3种核心原语暴露其能力:资源(用于信息检索,不执行副作用)、工具(执行带副作用的操作,如API调用或计算)和提示(可重用的模板和工作流)。

MCP的传输机制基于JSON远程过程调用协议(JSON-RPC, JavaScript object notation remote procedure call)构建,提供了一个专注于上下文交换和采样协调的状态会话协议。JSON-RPC的选择确保了其轻量级和广泛的兼容性,便于快速部署和集成。这种清晰的架构设计,将AI应用、协议适配和外部能力封装的职责分离,使各部分可以独立开发和演进,从而加速了MCP机制的落地和扩展。

2.2 MCP的核心概念与体系结构

MCP的强大之处在于其对“上下文”的深刻理解与标准化,这使其能够超越传统协议,实现真正的语义协作。

MCP的核心概念在于其能够统一描述设备的功能能力、输入输出模式及运行上下文。这里的

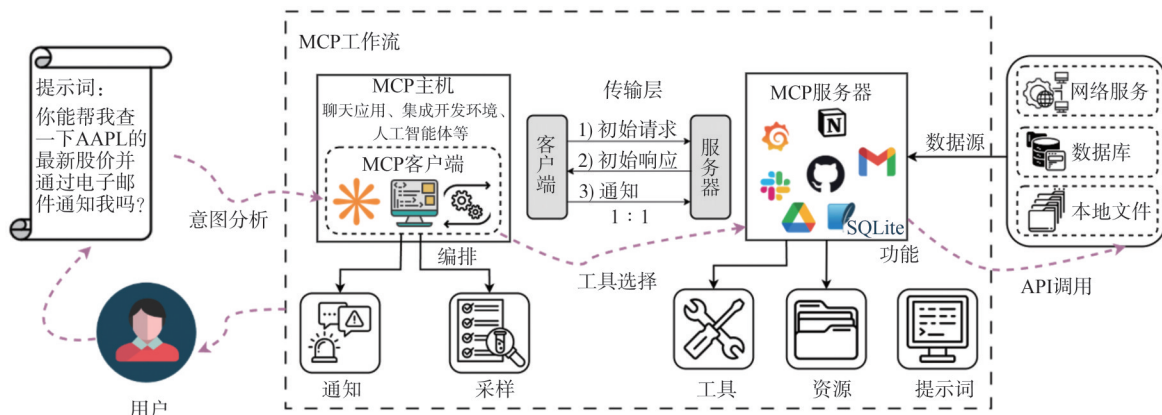


图4 MCP的经典架构示意

“上下文”不仅是简单的数据，更是数据背后的环境状态、任务意图、操作约束和系统配置^[30]。通过对这些信息的标准化描述，MCP使具身智能体能够在多模态、多域与跨平台的复杂场景中，获得可组合、可审计、可迁移的调用通道。例如，当机器人需要执行一个抓取任务时，MCP可以描述抓取工具的能力、所需物体的输入模式（如3D点云、颜色信息），以及当前的运行条件（如光照强度、机器人姿态、任务优先级）。这种详细的上下文信息使智能体能够更智能地理解和响应环境变化^[31]，从而实现真正的自适应行为。

MCP的体系结构实现了语义绑定替代接口绑定，这是其与传统中间件的根本区别。传统的系统集成通常依赖于接口绑定，即每个组件都需要了解并适配其他组件的特定API。当系统规模扩大、组件异构性增加时，这种方式会导致集成复杂性呈指数级增长。MCP则更像是一个抽象层，显式化了“任务-模型-资源”的语义边界。这意味着上层应用通过MCP完成“该做什么、在什么约束下做”的编排，而具体的实时性与确定性保证则由下层传输与调度机制来承载。例如，机器人不需要知道某个传感器具体是ROS2节点还是OPC UA设备，只需要通过MCP知道该传感器能提供何种“感知能力”及其“上下文”（如精度、刷新率、数据敏感等级）。这种语义层面的抽象，使系统更灵活和可重构，极大地简化了异构系统间的互操作性。

MCP通过“能力端点”的概念，实现了模型、工具和数据源的标准化暴露。在MCP框架下，无论是部署在云端的大型模型，还是运行在边缘设备的轻量级模型，都可以被视为MCP服务器所暴露的“能力端点”。MCP在运行时能够根据上下文信息动态路由对这些能力的调用^[32]，例如，根据网络时延和计算负载选择在边缘或云端执行推理。此外，数字孪生体也可以被暴露为MCP的“能力端点”，从而允许AI代理在离线/在线之间无缝切换，进行风险评估和策略预演，而无须直接操作物理实体。微软的“Agentic Web/NLWeb每个端点即MCP Server”设想，进一步将传统网站和企业应用转化为可被AI代理直接调用的能力端点。这种统一的“能力端点”抽象，极大地简化了复杂系统的构建和管理，使各种异构的计算资源、模型、工具和数据源能够被统一地发现、调用和组合，从而实现具

身智能体在开放环境中的可组合性和可扩展性。

2.3 MCP作为语义控制平面的定位与作用

MCP作为语义控制平面，核心价值在于提供一个高层次的抽象层，使具身智能与机器人物联网系统能够从“接口互联”升级到“语义协作”。这种定位并非取代底层实时数据平面，而是对其进行智能化的赋能和编排。

MCP通过“上下文-能力-意图”的统一描述，实现了底层协议差异的解耦，从而提供了一个可组合、可审计、可迁移的调用通道。它在现有的实时栈（如ROS2/DDS、OPC UA/TSN、5G/6G）之上叠加了一层新的抽象，使高层应用能够以统一的语义来表达任务需求，而无须关注底层通信协议的具体实现细节。这种分层设计，既利用了底层协议的实时性和确定性优势，又通过语义抽象解决了上层应用的灵活性和互操作性需求。它实现了“语义绑定替代接口绑定”的范式转变，使系统能够更灵活地适应动态环境和多变的任务需求。

在具身智能体的“感知-决策-执行”闭环中，MCP扮演着至关重要的角色，它负责将任务意图与上下文约束编译为可落地的执行计划。例如，在群体智能机器人协作任务中，MCP能够对多源感知数据、通信链路和计算资源进行统一建模与动态调度，从而实现多机器人之间的高效协同。当机器人接收到高层任务指令时，MCP会根据当前环境的实时上下文（如网络状况、设备负载、任务优先级等）将抽象的任务意图转化为一系列具体的、可执行的子任务，并为每个子任务分配相应的资源和性能约束。这个编译过程是动态且自适应的，能够根据实时反馈进行调整。

MCP与实时数据平面（ROS2/DDS、OPC UA/TSN、5G/6G）之间形成清晰的分工与协同关系。MCP负责高层“该做什么、在什么约束下做”的编排，即将抽象的语义需求转化为可执行的链路与调度计划。而底层的数据平面，如ROS2/DDS和OPC UA/TSN，则负责“如何在确定时限内把它做完”，提供实时、确定性的数据传输和调度服务。这种协同使MCP能够将高层语义需求精确地映射到DDS的QoS参数（如截止时间、时延预算、可靠性）和TSN的时间窗/调度策略，甚至5G/6G网络的切片和超可靠低时延通信（URLLC, ultra-reliable low-latency communication）配置。通过这

种方式, MCP将“抽象的语义需求”落实为“可执行的链路调度计划”, 从而实现了跨层协同优化。

3 “Context→QoS/Sec 编译器”: 形式化与优化

将MCP引入具身智能与机器人物联网的关键, 在于构建一个能够将高层任务语义与运行上下文精确转译为可执行的实时与安全侧写的机制。本节将详细阐述这一方案“Context→QoS/Sec 编译器”的核心思想、形式化建模、优化目标与求解策略, 以及其可复现研究价值的评测指标与基线。

3.1 编译器核心思想与形式化建模

“Context→QoS/Sec 编译器”是连接MCP语义控制平面与底层实时数据平面的核心组件。它接受由MCP描述的任务意图与环境状态, 并生成跨ROS2/DDS、OPC UA/TSN与5G/6G的实时QoS和安全参数, 同时决定边云算力分配与数据路径选择, 从而在强约束下实现语义控制平面与数据平面的闭环协同。

编译器的输入并非抽象的文字描述, 而是可度量的上下文张量 $\mathcal{C}$ 。这个张量包含了任务的详细语义与约束。

1) 任务语义与约束: 包括任务的优先级、可容忍的时延抖动 δ 、误检风险上限 ϵ 以及数据的敏感等级 s 。这些参数直接量化了任务对性能和安全的

要求。

2) 环境态: 涵盖实时网络链路拥塞情况、时变带宽、边缘节点和云端的算力负载以及能耗预算。这些反映了系统当前可用物理资源状况。

3) 数字孪生侧的风险评估结果: 数字孪生作为物理系统的虚拟镜像, 能够对潜在的风险进行预演和评估, 并将这些结果反馈给编译器作为输入, 指导决策。

编译器的输出是跨层的决策变量 z 。这些变量是底层协议和资源管理机制的配置参数。

1) DDS 参数: 如截止时间、时延预算、可靠性和历史深度。这些参数直接影响消息的实时传输和可靠性。

2) TSN 参数: 包括周期 T_c 、时隙向量 $\mathbf{g}$ 和时间偏置 ϕ 。这些参数定义了时间敏感网络的调度策略, 确保关键数据流的确定性传输。

3) 6G 参数: 如网络切片类型和数据包时延预

算/可用性 (PDB/AR, packet delay budget/availability)。这些参数决定了蜂窝网络为不同服务提供的QoS保障。

4) 边云分配与安全侧写: 包括是否启用可信执行环境约束、所需加密等级 κ 以及密钥轮换周期。这些参数直接影响计算任务的部署位置和数据的安全保护级别。

MCP的“意图-能力-资源”三元组在此被映射为可编排的参数组, 实现了高层语义到低层配置的精确转换。端到端时延与可靠性模型是编译器形式化建模的核心。给定一个操作有向无环图(DAG, directed acyclic graph) $\mathcal{G} = (V, E)$, 其中, V 是节点(计算任务), E 是边(数据流), 每条任务流 π 的最坏情况下界可以通过网络演算表达为

$$L_{\text{c2c}}(\pi) \leq \sum_{e \in \pi} \underbrace{q_e(z) + t_e(z)}_{\text{排队+传输}} + \sum_{v \in \pi} \underbrace{c_v(z)}_{\text{计算}} \quad (1)$$

其中, $q_e(z)$ 是边 e 上的排队时延, $t_e(z)$ 是传输时延, $c_v(z)$ 是节点 v 的计算时延。TSN的门控机制为每条优先流提供速率-时延型服务曲线 $\beta_e(t)$, 而DDS的时延预算和截止时间则限制了 q_e 的上界。

可靠性在多跳网络中可以从乘性转换为可加约束。如果每跳的丢包率为 $p_e(z)$, 则端到端可靠服务等级目标(SLO, service level objective) ε 可以线性化为可求解的凸/混合整数子问题:

$$1 - \prod_{e \in \pi} (1 - p_e) \leq \varepsilon \Leftrightarrow \sum_{e \in \pi} \ln(1 - p_e(z)) \geq \ln(1 - \varepsilon) \quad (2)$$

这种转换使端到端可靠性约束能够被纳入优化问题中。

安全侧写在编译器中并非事后补丁, 而是作为优化约束的内生组成部分。当数据流的敏感等级 s 达到或超过某个阈值 s 时, 决策变量 z 需要强制满足特定安全要求, 例如, 计算节点必须启用TEE, 即 $\text{TEE}(v) = 1$, 且链路加密强度 κ 必须达到或超过 κ^* 。随之引入的可信计算基(TCB, trusted computing base)提升和加密开销会被并入计算时延 $c_v(\cdot)$ 和传输时延 $t_e(\cdot)$ 的代价模型中。通过这种方式, “安全”与“实时”在同一目标函数中被共同优化, 而不是互相牺牲。

3.2 优化目标与求解策略

编译器的优化目标不应简单地最小化均值时延, 因为具身智能体的闭环控制对尾部风险更敏

感。因此，本框架采用条件风险价值（CVaR, conditional value at risk）作为统一的风险度量，以在分位数层面保障服务等级目标。令 L 为端到端时延的随机变量，主目标函数将 $\alpha(L)$ 作为主项，并辅以语义效用与成本正则项：

$$\min_z \sum_{\pi} \mathcal{S}_{\pi} + \mu \mathcal{E}(z) + \nu \mathcal{K}(z) \quad (3)$$

其中， $\mathcal{S}_{\pi}$ 表示任务流 π 的语义达成度，由语义通信返回的信息价值评分驱动（例如，任务相关熵降或正确率边际）。 $\mathcal{E}(z)$ 表示系统能耗或费用， $\mathcal{K}(z)$ 表示密钥管理与TEE带来的额外成本。语义通信在此并非辅助功能，当 $\mathcal{S}_{\pi}$ 的边际收益可以通过压缩或丢弃部分模态数据来维持不变时，编译器会倾向于在空口“少传但不降质”，将带宽和计算资源分配给更关键的数据流，从而在物理资源固定的情况下提高整体任务达成度。

直接求解上述问题将陷入NP难的混合整数非凸区域，但该问题具备清晰的分解结构。因此，有效求解策略是拆分为3部分进行交替迭代优化。

1) 固定边云部署与TSN时隙：针对DDS/5G的连续QoS参数求解一个凸近似子问题。CVaR可以通过线性化处理，从而实现高效求解。

2) 固定链路与QoS参数：对TSN的门控和时隙进行混合整数线性规划（MILP, mixed-integer linear programming）求解，同时引入最小化时延抖动的二范数正则项，以抑制频繁的重编排。

3) 固定网络配置：解决边云放置与容器安全侧写的二分配问题。利用拉格朗日乘子对能耗和费用进行影子定价，促使低收益的云外包任务回迁至边缘执行。

每轮交替迭代结束后，利用数字孪生返回的扰动集进行一次分布鲁棒校准。这使编译器能够在最坏情况下评估“看起来最优”的解决方案在一组可能的网络状态下的表现，从而避免现实部署中出现“脆弱的最优解”。

具身系统是强时变的，因此可用的方案需要事件触发的增量重编译机制。当测得的链路时延、抖动或算力有效频率偏离预设阈值 τ 时，编译器仅在受影响的子图上进行局部重求，以减少计算开销和系统扰动。当安全态从“低”切换到“高”时，则执行“安全优先”覆盖，将TEE/加密切换作为强约束即时落地，并在下一次完整编译时再做细化。这种两阶段策略将工程可用性与优化品质有效结合。

3.3 实证分析与性能验证

为验证所提出“Context→QoS/Sec编译器”在动态环境下的有效性，文献[33]基于MCP构建了事件驱动仿真平台，模拟具身智能体在多源动态网络与安全场景中的运行行为。实验场景包括带宽波动、深度衰落及安全态变化3类扰动窗口，用以考察系统在不同扰动模式下的自适应性。比较对象包括静态策略（static MCP）与仅任务驱动策略（task-only MCP）。仿真以端到端时延均值、尾时延以及服务等级目标违约率为评价指标。

文献[33]结果表明，动态模型上下文协议（dMCP, dynamic model context protocol）在多源动态环境中表现出显著优势。与静态控制相比，平均时延分别降低约16%、95%与99%，尾时延降低约6%，SLO违约率下降56%。其性能提升主要得益于事件触发机制与大模型驱动的策略再编译能力，可在链路拥塞、抖动及安全威胁变化时动态调整TSN周期、时延预算及加密强度，实现时延、可靠性与安全的联合调优。

4 挑战与未来研究方向

尽管模型上下文协议及其驱动的“Context→QoS/Sec编译器”为具身智能与机器人物联网的跨层协同提供了创新的解决方案，但该领域仍面临诸多挑战。本节将对这些挑战进行深入剖析，并提出未来值得优先推进的研究方向。

4.1 代价模型的可识别性与在线校准

编译器的所有决策都隐含在“排队-传输-计算-安全侧写”的联合代价模型中。然而，这些代价项并非静态常数，它们是高度动态且相互耦合的。例如，压缩比的改变会影响空口误码率和重传次数，从而影响传输代价；启用TEE和强加密会显著增加端到端尾部时延，提升计算和传输代价；困难样本的上送策略会改变边云分工的边际收益。如果这些代价项只能依靠经验阈值给定，编译器的最优性将无法保证，且难以适应具身系统强时变的特性。

未来的研究方向应着眼于将代价建模视为一个可学习问题。在保证物理合理性的前提下，利用在线观测数据对排队时延、加密开销、计算负载等参数进行持续校准。这需要开发鲁棒的在线学习算法，能够从噪声数据中准确识别代价模型的动态变

化,并将“模型可识别”作为编译器收敛性的前置条件,而非事后调参的经验性工作。例如,可以借鉴上下文感知 QoS 预测框架(CQEC, context-aware QoS prediction framework)中的卷积神经网络和注意力机制来提取上下文特征,以及鲁棒 QoS 预测模型(RA-QoS, robust QoS prediction model)处理噪声数据和偏差的方法,以提高代价模型预测的准确性和鲁棒性。

4.2 约束的同构化与统一风险泛函

具身智能系统的要求看似异构且相互矛盾:既要低时延、高可靠性,又要强安全、低能耗。在传统设计中,这些目标通常需要进行权衡和牺牲。然而,若能将它们写入同一个风险泛函,问题就变得同构,从而实现全局最优。

研究方向应聚焦于将所有异构约束同构化,并纳入一个统一的“风险-效用-成本”泛函中进行优化。例如,将端到端时延的尾部风险用 CVaR 表示,将可靠性写成对数可加的违约边界,将安全与能耗转化为对计算和链路的不确定性惩罚。这样做不仅使目标函数在数学上更“优美”,更重要的是,它使任何方案的改动都可以在同一泛函上讨论收益与代价,避免了“指标准则漂移”的问题。这种统一的优化框架将促进跨领域技术的协同,使编译器能够理性地在性能、安全和资源之间进行权衡。

4.3 可验证性与可解释性:轻量证明物与审计链

当 MCP 触发一次上下文切换,系统必须能够回答 3 个关键问题:为什么要切换(充分性)、切换到哪里(可行性)、切换后是否仍满足 SLO(可验证性)。仅依靠经验规则无法提供这种严格的保证^[34-35]。在安全关键的具身智能系统中,决策的透明度和可追溯性至关重要。

可行的研究路径是引入形式化的轻量证明物。在每次编译输出时,除了生成决策变量,还应同时给出网络演算意义下的时延上界与可靠性下界。这些形式化的界限应与数字孪生的“预演结果”联合存证。日后回溯时,上下文切换的因果链可以被完整复现,这种“审计友好”的可解释性,才是 MCP 在工业场景落地的决定性差异。未来的工作应探索如何自动化生成这些证明物,并将其集成到 MCP 的协议栈中,形成一个可信的审计链。

4.4 工程演化与稳定性:渐进最优策略

具身系统对频繁重编排高度敏感,即使平均性能有所提升,也可能造成控制回路抖动,影响系统稳定性。因此,在工程实践中,稳定性通常优先于纯粹的最优性。

编译器设计应引入变更约束,以抑制频繁的系统重构。例如,对单位时间内可改写的 QoS/时隙数量设置上限,对参数变化加入“二范数惩罚”,将“少改一次、改得有把握”作为解空间的内在偏好。与之配套的是双时间尺度策略:在小时间尺度上,只在受影响的局部子图上进行微调;在大时间尺度上,在作业边界进行全局再优化。这种策略将“渐进最优”变成系统性的设计原则,而非仅依赖于调参经验,从而在保证系统稳健性的前提下实现性能的持续优化。

4.5 与大模型的深度融合

LLM 在机器人控制和决策方面展现出巨大潜力,但其计算资源需求、实时性约束以及在真实物理世界中的泛化能力仍是挑战。LLM 的训练和部署需要大量的计算资源和数据,这对于资源受限的机器人平台而言是一个严峻的挑战。

未来的研究应探索 MCP 如何作为 LLM 与物理世界交互的标准化桥梁,实现 LLM 驱动的具身智能体^[36]更高效、安全地执行复杂任务,其中包括以下 3 个方面。

1) LLM 作为 MCP 服务器的能力暴露:将 LLM 的推理能力封装为 MCP Server,使机器人可以按需调用云端或边缘的 LLM 服务,实现“近端控制+远端重算”的 LLM 应用。

2) 上下文感知 LLM 调用:MCP 可以为 LLM 提供丰富的上下文信息,包括任务意图、环境状态、安全敏感等级等,从而引导 LLM 生成更符合实际约束的决策和行动计划。

3) LLM 驱动的 QoS/Sec 策略生成:探索 LLM 如何辅助 Context→QoS/Sec 编译器生成更智能、更自适应的 QoS 和安全策略,例如,根据自然语言的任务描述自动推断 QoS 需求。

4.6 标准化与跨域互操作性

尽管 MCP 已获得 Anthropic 和微软等主要参与者的支持,但其在不同厂商、不同领域间的实现兼容性仍是挑战。要实现具身智能与机器人物联网的普适性,需要更广泛的行业共识和标准化努力。

研究方向应推动在行业可复用的最小上下文集上达成共识。将“任务语义-时延容忍-安全级别-能耗预算”等核心上下文变量固化为MCP的标准字段，使不同系统能够共享同一“源语言”。这不仅有助于促进学术研究的横向比较，也将极大地降低工程落地的集成成本，加速MCP在工业界和学术界的广泛采纳，促进跨协议栈的无缝集成。

5 结束语

本综述以“MCP作为语义控制平面→Context→QoS/Sec编译器”为主线，综述了面向具身智能与机器人物联网的MCP，并强调通过统一的源语言将其编译为可执行的跨层参数，最终由DDS/TSN/5G等数据平面提供确定性承载。这种创新的最小闭环范式，使具身智能系统能够在精度、时延与安全3个维度上实现协同优化。

面向未来，具身智能与机器人物联网的进一步发展，并非依赖于“大而全”的平台，而是需要聚焦于以下3项“小而坚实”的关键工作。

1) 在行业可复用的最小上下文集上达成共识，并固化为MCP的标准字段。将任务语义、时延抖动容忍、误检风险上限、能耗与安全等级等核心约束，作为跨系统可共享的公共接口，为不同研究团队的工作提供统一的“源语言”，极大地促进互操作性和研究成果的复用。

2) 给出可验证的“上下文→QoS/Sec”映射原理，明确代价模型的可识别边界与收敛条件，并随观测进行在线校准。这要求将“最优性”建立在可测可证的基础上，通过数据驱动和在线学习，确保编译器在动态环境中的决策是基于准确的代价预测。

3) 构建轻量的证明物与审计链，将每次编译输出的时延上界、可靠性下界与安全策略一并固化。这将使系统的每一次上下文切换都“说得清、验得过、追得到”，从而形成可审计的知识库，提升系统在安全关键场景中的可信度和透明度。

当上述3点落地，具身智能与机器人物联网才算真正完成从“接口互联”向“语义协作”的本质跃迁。沿着本综述所提的主线继续推进——以最小完备的上下文标准为锚、以形式化的编译与证明为桥、以确定性数据面为底座——将有望把具身智能体的“自适应、可迁移与可审计”从宏伟愿景变为可复现

的常态。这将为下一代智能物联网系统提供统一、可编程、可审计的协议支撑，最终实现具身智能体在开放、动态、复杂环境中的自主、高效、安全运行。

参考文献：

- [1] Wang B C, Tao F, Fang X D, et al. Smart manufacturing and intelligent manufacturing: a comparative review[J]. Engineering, 2021, 7(6): 738-757.
- [2] 王文晟, 谭宁, 黄凯, 等. 基于大模型的具身智能系统综述[J]. 自动化学报, 2025, 51(1): 1-19.
Wang W S, Tan N, Huang K, et al. Embodied intelligence systems based on large models: a survey[J]. Acta Automatica Sinica, 2025, 51(1): 1-19.
- [3] Ghobakhloo M, Iranmanesh M, Tseng M L, et al. Behind the definition of Industry 5.0: a systematic review of technologies, principles, components, and values[J]. Journal of Industrial and Production Engineering, 2023, 40(6): 432-447.
- [4] 郝肇铁, 郭斌, 赵凯星, 等. 从规则驱动到群智涌现: 多机器人空地协同研究综述[J]. 自动化学报, 2024, 50(10): 1877-1905.
Hao Z T, Guo B, Zhao K X, et al. From rule-driven to collective intelligence emergence: a review of research on multi-robot air-ground collaboration[J]. Acta Automatica Sinica, 2024, 50(10): 1877-1905.
- [5] Duncan J A, Alambeigi F, Pryor M W. A survey of multimodal perception methods for human-robot interaction in social environments[J]. ACM Transactions on Human-Robot Interaction, 2024, 13(4): 1-50.
- [6] 牛志升. 面向6G网络的高可靠低延时通信计算与控制[J]. 中国科学: 信息科学, 2024, 54(5): 1267-1282.
Niu Z S. uRLLC³: ultra-reliable and low-latency communication, computing, and control for 6G networks[J]. Scientia Sinica (Informationis), 2024, 54(5): 1267-1282.
- [7] Krejčí J, Babiuch M, Suder J, et al. Internet of robotic things: current technologies, challenges, applications, and future research topics[J]. Sensors, 2025, 25(3): 765.
- [8] Ullah S, Khan S, Vanecek D, et al. Machine learning and digital-twins-based Internet of robotic things for remote patient monitoring[J]. IEEE Access, 2025, 13: 57141-57165.
- [9] Li Q H, Ye Q, Zhang N, et al. Digital-twin-enabled industrial IoT: vision, framework, and future directions[J]. IEEE Wireless Communications, 2025, 32(6): 173-181.
- [10] 孙长银, 袁心, 王远大, 等. 具身智能自主无人系统技术[J]. 自动化学报, 2025, 51(4): 762-777.
Sun C Y, Yuan X, Wang Y D, et al. Embodied intelligence autonomous unmanned systems technology[J]. Acta Automatica Sinica, 2025, 51(4): 762-777.
- [11] Liu Z X, Du H Y. Model context protocol-based Internet of experts for wireless environment-aware LLM agents[PP]. V1. arXiv

- (2025-05-03)[2025-05-30]. arXiv: 2505.01834.
- [12] Hou X Y, Zhao Y J, Wang S N, et al. Model context protocol (MCP): landscape, security threats, and future research directions[PP]. V1. arXiv (2025-03-30)[2025-05-30]. arXiv: 2503.23278.
- [13] Ahmadi A, Sharif S, Banad Y M. MCP bridge: a lightweight, LLM-agnostic RESTful proxy for model context protocol servers[J]. V1. arXiv (2025-04-11)[2025-05-30]. arXiv: 2504.08999.
- [14] Ehtesham A, Singh A, Gupta G K, et al. A survey of agent interoperability protocols: model context protocol (MCP), agent communication protocol (ACP), agent-to-agent Protocol (A2A), and agent network protocol (ANP)[PP]. V1. arXiv (2025-05-04)[2025-05-23]. arXiv: 2505.02279.
- [15] Jiao J F, Park J, Xu Y M, et al. SafeMate: a modular RAG-based agent for context-aware emergency guidance[PP]. V1. arXiv (2025-05-05)[2025-05-30]. arXiv: 2505.02306.
- [16] 刘阳, 柏永杰, 林惊. 面向人机物高效融合与协作的具身智能技术体系[J]. 机器人, 2025, 47(4): 559-580.
- Liu Y, Bai Y J, Lin L. Technological system for embodied intelligence towards efficient integration and coordination of human, robot, and PhysicalWorld[J]. Robot, 2025, 47(4): 559-580.
- [17] Li Z W, Zhang T Z, Zhou M H, et al. MIPD: a multi-sensory interactive perception dataset for embodied intelligent driving[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(11): 21320-21334.
- [18] Mon-Williams R, Li G, Long R, et al. Embodied large language models enable robots to complete complex tasks in unpredictable environments[J]. Nature Machine Intelligence, 2025, 7(4): 592-601.
- [19] 曾凯, 王耀南, 谭浩然, 等. AI大模型驱动的具身智能人形机器人技术与展望[J]. 中国科学: 信息科学, 2025, 55(5): 967-992.
- Zeng K, Wang Y N, Tan H R, et al. Prospects and technology of embodied intelligent humanoid robots driven by AI large models[J]. Scientia Sinica (Informationis), 2025, 55(5): 967-992.
- [20] Chinedu Okafor K, Oname A, Chinebu T I, et al. HAC-19: a coinfection model for infectious diseases using IoT-networked robots[J]. IEEE Internet of Things Journal, 2025, 12(16): 33545-33560.
- [21] Chen C, Cheng Y J, Dou Z, et al. A UAV power line patrolling system with edge intelligence and Beidou SMS in signal loss area[J]. IEEE Internet of Things Journal, 2025, 12(15): 30020-30035.
- [22] Ren K J, Wang G T, Morgan A S, et al. Collision-inclusive manipulation planning for occluded object grasping via compliant robot motions[J]. IEEE Robotics and Automation Letters, 2025, 10(9): 9120-9127.
- [23] Kayataş Y, Kabadayı S. An AI-assisted autonomous IoRT agent for smart spaces[C]//Proceedings of the 2024 9th International Conference on Computer Science and Engineering (UBMK). Piscataway: IEEE Press, 2024: 613-618.
- [24] 沈甜雨, 陶子锐, 王亚东, 等. 具身智能研究的关键问题: 自主感知、行动与进化[J]. 自动化学报, 2025, 51(1): 43-71.
- Shen T Y, Tao Z R, Wang Y D, et al. Key problems of embodied intelligence research: autonomous perception, action, and evolution[J]. Acta Automatica Sinica, 2025, 51(1): 43-71.
- [25] 徐文渊, 冀晓宇, 闫琛, 等. 具身智能安全治理[J]. 中国科学院院刊, 2025, 40(3): 429-439.
- Xu W Y, Ji X Y, Yan C, et al. Embodied artificial intelligence security and governance[J]. Bulletin of Chinese Academy of Sciences, 2025, 40(3): 429-439.
- [26] Toyomoto Y, Oshima T, Hatanaka T. Robot zoo aqua: aquatic control testbed of multiple USVs on robot operating system 2[C]//Proceedings of the 2024 SICE Festival with Annual Conference (SICE FES). Piscataway: IEEE Press, 2024: 196-201.
- [27] Singh A, Shetty A, Ehtesham A, et al. A survey of large language model-based generative AI for text-to-SQL: benchmarks, applications, use cases, and challenges[C]//Proceedings of the 2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC). Piscataway: IEEE Press, 2025: 15-21.
- [28] Model context protocol[R]. 2025.
- [29] Sunavala F. Introducing model context protocol (MCP) in azure AI foundry: create an MCP server with azure AI agent service[R]. 2025.
- [30] Szeider S. MCP-solver: integrating language models with constraint programming systems[PP]. V1. arXiv (2025-04-06)[2025-05-30]. arXiv: 2501.00539.
- [31] Cao X G, Shi L Y, Wang X L, et al. MCP: a named entity recognition method for shearer maintenance based on multi-level clue-guided prompt learning[J]. Applied Sciences, 2025, 15(4): 2106.
- [32] Mueller M. How to use model context protocol the right way[R]. 2025.
- [33] Guo S S, Zhu J B, Ye J, et al. Task-oriented communications for agentic IoT: an LLM-driven QoS/security policy generation via dynamic model context protocol[C]//Proceedings of the 2025 Seventeenth International Conference on Wireless Communications and Signal Processing (WCSP). Piscataway: IEEE Press, 2025: 1-5.
- [34] Tanveer M. Model context protocol (MCP): the next frontier in information security[R]. 2025.
- [35] Santos O. AI model context protocol (MCP) and security[R]. 2025.
- [36] Nitesh Palamakula S. Agent-to-agent protocol vs. model context protocol (MCP)[R]. 2025.

[作者简介]



郭帅帅(1990—), 男, 博士, 山东大学控制科学与工程学院教授、博士生导师, 主要研究方向为大模型 AI 智能体、智能通信、通信系统性能分析。



朱家兵(1975–)，男，浪潮软件集团有限公司首席技术官、正高级工程师，主要研究方向为人工智能。



张安邦(2000–)，男，山东大学控制科学与工程学院硕士生，主要研究方向为语义通信、机器学习。



叶佳(1997–)，女，博士，重庆大学电气工程学院教授、博士生导师，主要研究方向为无线通信系统设计分析与优化、通信与电力系统交叉。



倪芝洋(1986–)，女，博士，江苏第二师范学院计算机工程学院副教授、硕士生导师，主要研究方向为人工智能在无线通信中的应用、未来网络与资源分配。