

“艾武大模型+”：一种军事大模型系统的开发与实证

崔脩龙, 高志强, 姬纬通, 沈佳楠, 张敏, 邱鑫源

(武警工程大学反恐指挥信息工程教育部重点实验室(立项), 西安 710086)

摘要: 智能化指挥是新型指挥控制理论研究的重要方向, 大模型是智能交互、任务规划和辅助决策等智能化指挥能力实现的重要支撑。本文兼顾理论与实践, 梳理大模型军事能力需求, 设计面向智能化指挥的大模型应用框架, 提出“艾武大模型+”的系统架构、信息流程和协同架构, 梳理工程实现的关键技术, 以智能化指挥实证案例及选型分析验证“艾武大模型+”系统在多模态交互和特定任务军语理解的能力优势, 拓展有/无人平台的末端协同和指令控制, 为重大国防军事专项、智能化指挥研究与落地应用提供参考。

关键词: 军事大模型; 提示工程; 智能化指挥; 专用语料; 多模态交互

中图分类号: TP18 **文献标志码:** A

“Aiwu Large Model+”: Development and Empirical Study of Military Large Model System

CUI Xiaolong, GAO Zhiqiang, JI Weitong, SHEN Jianan, ZHANG Min, QIU Xinyuan

(Key Laboratory of Counter-Terrorism Command & Information Engineering of Ministry of Education (Approval), Engineering University of PAP, Xi'an 710086, China)

Abstract: Intelligent command is an important direction for the new command and control theories, and large language models are important support for the realization of intelligent command capabilities such as intelligent interaction, task planning, and auxiliary decision-making. Combining theory and practice, we outline the military capability requirements of the large model and design a large language model application framework for intelligent command. Then, the system architecture, information process, and collaborative architecture of the “Aiwu large model+” system are proposed and the key technologies for engineering implementation are proposed. Empirical cases of intelligent command are used to verify the multimodal interaction and military language understanding of the system. Collaboration and command control of manned/unmanned platforms can be expanded, which provides reference for research and implementation of the major national defense and military special projects and the intelligent command in the future.

Key words: military large model; prompt engineering; intelligent command; specialized corpus; multimodal interaction

引 言

智能化指挥是新型指控理论研究的重要方向,传统指控方式已不适应智能化战争中海量情报、智能决策和多域协同等智能化指挥需求。智能驱动是智能化指挥的基本特征^[1]。2022年,OpenAI发布的大语言模型 ChatGPT^[2],具有极高的人机交互水平,引起了军地各界的广泛关注,后续涌现的 GPT-4、LLaMA 及 Genie 等系列生成式大模型^[3]可为智能交互、任务规划和辅助决策等智能化指挥能力的实现提供技术支撑,并催生了新型指控模式变革的新方向。鉴于大模型在海量信息参数化全量记忆、多角色长文本生成、即时新知识学习与进化等方面的优势,美国国防信息系统局已将 ChatGPT 等生成式人工智能技术添加到观察名单;美国国防部成立“利马(Lima)”生成式人工智能工作组,负责在整个国防部范围内“评估、协调和使用”生成式人工智能技术;美国 Palantir 公司的人工智能平台 AIP,利用大模型优化决策流程、缩短决策时间,进而获得最优作战方案、保证作战质量;美国 Scale AI 公司推出“多诺万(Donovan)”军事决策辅助系统,利用大语言模型增强分析人员和作战人员对作战环境的理解^[4]。

尽管美国等军事强国已开展军事大模型系统的相关研究,但总体来看该方向尚处于探索试用阶段,军事理论成果和实证应用研究较为匮乏。国内大模型主要应用于智能问答、内容生成和娱乐办公等民用领域,但在性能上与国外大模型仍存在较大差距^[5-6]。国防领域的天机大模型利用部分情报数据进行了特定场景训练,武警部队科研团队提出“艾武大模型+”原型系统,在多模态交互、任务规划和决策建议生成方面进行了探索,相关成果获 2023 年中国国防科学技术信息学会“智研”杯大模型国防应用邀请赛“应用创意陈述”和“应用潜能展示”双赛道第一名,为探索大语言模型的国防应用,共享军事理论成果,本文对“艾武大模型+”的系统架构、信息流程和协同架构进行总结,并对其工程实现的关键技术进行了梳理。

1 相关军事理论

智能化指挥是传统指控模式的重要演进方向,近年来已成为军事智能的研究热点,但将大模型融入新型指控模式研究、系统开发的公开工程应用成果鲜见。而且,简单将大模型叠加到现有指挥架构和范式中,必然无法实现军语理解及智能决策,更无法产生体系倍增、大规模群智涌现。本节将从面向智能化指挥的大模型能力需求梳理和多尺度应用框架设计角度,研究基于大模型技术的智能化指挥理论。

1.1 面向智能化指挥的大模型能力需求

形式上,智能化指挥是人工智能与指挥控制理论全流程、全要素、多维度和多层次的耦合;本质上,是以大模型为代表的智能化技术对指挥控制的模式重塑、能力激发、涌现反馈和迭代提升。通常,大模型以文本、图像和语音等数据为输入,整合各类任务子模型和多源多类任务数据,驱动智能化指挥流程在军事应用场景的应用闭环,涌现指挥智能、混合智能和人机智能等军事智能^[7]。如图 1 所示,面向智能化指挥的大模型能力需求包括军事语义理解能力、多模态交互能力、精准军事数据反哺能力以及决心建议生成能力等。

(1)军事语义理解能力。语言、文字和符号等是信息交流的载体。战场或特定军事任务场景下,指挥员通过口述战斗命令或作战文书来传达战斗命令等信息,即文本和语音承载了大量军事语义,包括对作战构想的理解、作战命令、计划的解析,作战指令的解码等^[8]。为形成智能化指挥的军事语义理解能力,大模型需要基于大量军事文书、军用指令等语料数据,训练形成对军事文本、语音蕴含的语义理解能力,并从指挥员视角理解和认知作战命令、军事术语等,生成符合战场态势或任务场景的作战

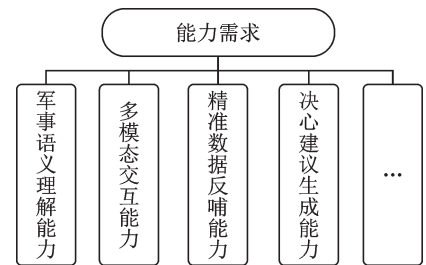


图 1 面向智能化指挥的大模型能力需求
Fig.1 Large language model capability requirements for intelligent command

指令。

(2)多模态交互能力。在现代战争中,战场态势瞬息万变,务必将信息快速地传达到各个战斗单元。在智能化指挥模式下,大模型需支持语音、图像、姿态和手势等多种模态的交互和理解方式^[9]。例如,将口述战斗命令的语音指令精确转换为文字,形成指挥人员对智能系统的精确控制信号。多模态理解和交互能力对突破传统指控交互瓶颈、创新智能化交互模式、提高智能化指挥效率具有重要意义。

(3)军事数据精准反哺能力。尽管大模型在信息检索、要素提取和摘要生成等场景均有良好应用,但军事数据具有典型的精确时空属性^[10],而非多数通用智能所依赖的概率性。因此,为方便指挥员迅速从众多军用文书、案例等军事文档中提取精确管用的作战信息,大模型的知识提取、智能归档分类以及决心建议生成等基础能力需要融合军事数据的精准反哺,支撑智能化指挥所需的精准分类标签、关键要素抽取、摘要生成以及态势情况判断。

(4)模型对齐的决心建议生成能力。智能化指挥的重要特征之一是为指挥员生成辅助决心建议。利用大模型的逻辑推理和理解能力,汇聚情报、研判态势、兵力编成、战场环境和敌情社情等作战信息,对齐系列定制化任务小模型和数据应用服务,整合数据、理解态势、情报分析和任务规划,并基于此生成符合作战文书的任务辅助决心建议。

此外,大模型的军事应用还涉及军语翻译、系统操控、末端执行、军事标图、评估成长和智能体对齐等能力,可分为技术探讨、原型系统、点状应用和成熟推广等发展演进层级^[11]。其中,军事标图能力是指挥员必备的作战指挥能力。从作战数据角度,地图是地理信息数据、(非)结构化数据等信息的综合载体。基于大模型的图像理解,以及大数据整合能力,形成军事决心建议的标图绘图,可以多源高效地加深战场态势融合,提升全局性作战态势理解。

1.2 面向智能化指挥的大模型应用框架

指挥控制是战场胜负的决定性因素之一。战术层面的指挥控制一般以文献[12]提出的观察-判断-决策-行动(Observe-orient-decide-act, OODA)环为指导;战役和战略层面的指挥控制通常具备筹划-准备-执行-评估(Planning-readiness-execution-assessment, PREA)等关键环节^[13]。本节融合战术、战役和战略层面的多尺度指挥控制需求,结合OODA环和PREA环特点,以大模型为核心,融合两个作战尺度,提出面向智能化指挥的大模型应用框架(Intelligent command application framework based on large language model, ICAF-LLM),如图2所示。ICAF-LLM框架中的大模型可选型为国内自主可控的GLM^[14]、文心一言和通义千问等,需要具备多模态交互、文本生成、语言翻译和人类意图对齐等基础能力,以及军事智能所需的生成能力,进而可以从宏观和战术两个尺度支撑智能化指挥过程。

在宏观尺度上,ICAF-LLM框架可概括为大模型+PREA过程,涉及大模型+的筹划、准备、执行和评估4个阶段。(1)筹划阶段。大模型结合海量历史/实时数据演绎得到态势信息,理解上级意图,做出假设和周密决策。根据大模型的认知学习和理解能力,生成任务预案。(2)准备阶段。大模型利用综合推理能力预判直前态势,综合评估预案,调整合适的计划,进一步生成作战指令。(3)执行阶段。大模型以其强大的推理、理解能力和多媒体资源识别能力,监控实施态势与计划设想的偏差,实时评估威胁,实施行动转换,辅助控制战场态势。(4)评估阶段。大模型对作战行动总结、复盘,梳理过程演变态势,度量后续发展趋势,生成并反馈修订方案。

在战术尺度上,ICAF-LLM框架可概括为大模型+OODA流程,涉及观察、判断、决策和行动4个阶段。大模型针对情报侦察单元传回的文字、图片、视频等多模态信息进行识别、抽取、理解,感知实时任务现场态势,完成观察阶段,而后依靠大模型强大的分析推理能力分析情报,抽取关键要素,判断情况,基本形成决策依据,并根据此给出决心建议,快速精准决策。在实时行动中,将大模型+各类任务子模型和多源多类任务数据紧密耦合,可形成快速识别跟踪、应急事件精准响应能力,结合反馈评估,为指挥员临机决策提供支持。

2 系统设计

对标智能化指挥的能力需求,遵循 ICAF-LLM 框架,“艾武大模型+”的系统架构为“数字人+大模型+大系统+大数据”,如图 3 所示,其系统架构自底向上划分 4 个层级,大数据层融合汇聚领域专用、公开和外部等数据,形成涵盖我情、敌情和战场环境等精确数据的领域语料库;大模型层通过预训练、指令微调 and 外挂知识库等,形成大语言模型的军语理解、任务规划和决策建议等能力;大系统层对齐调用系列决策分析(任务)小模型(矩阵),融合嵌入大数据应用系统及服务;数字人层具身为“艾武”虚拟参谋,具备语音、文本等智能交互能力,可支持“云-边-端”有/无人平台的末端协同和指令控制拓展。

按照 ICAF-LLM 框架的两个应用尺度,“艾武大模型+”系统以自主可控大模型为核心,宏观上满足大模型+筹划、准备、执行和评估的 PREA 过程,战术层面满足大模型+观察、判断、决策和行动的 OODA 阶段,其信息流程如图 4 所示。具身数字人“艾武”虚拟参谋,通过语音、文本等多模态交互能力完成外部任务接入,将任务指令正向输入军事领域大模型,分解军事任务要素,通过人在回路的要素确认后,对齐调用大数据应用系统所提供的系列决策小模型,融合精准敌情、我情、战场环境等大数据,形成基于内容生成能力的情况判断和决策建议,并通过军事反馈评价提升大模型自学习自优化能力,以及有/无人平台的末端协同和指令控制能力。

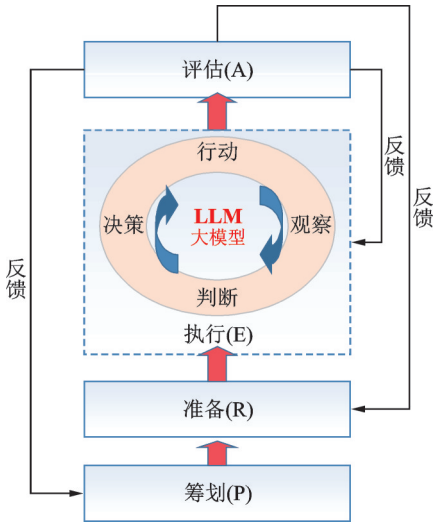


图 2 面向智能化指挥的大模型应用框架
Fig.2 Application framework of large model for intelligent command

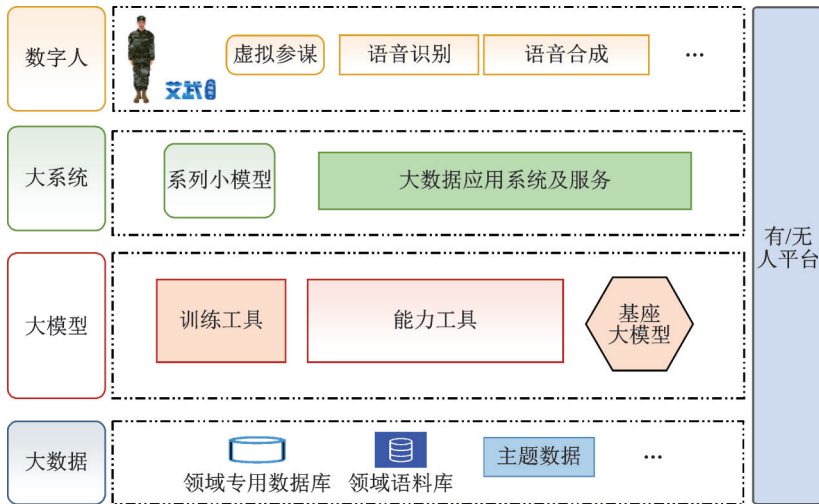


图 3 “艾武大模型+”系统架构
Fig.3 Architecture of “Aiwu large model+” system

“艾武大模型+”系统的实现是涉及作战指挥、大数据、人工智能、软件工程以及最优化等多交叉领域的复杂系统工程,包括数字人、大模型、大系统、大数据和有/无人平台等关键要素,其协同关系如图 5 所示,可归纳为以下 3 点。(1)数字人与大模型协同。作为人机交互的接口,指挥员通过语音、文本等多模态输入方式向“艾武”虚拟参谋发起任务或自主感知任务,“艾武”将多模态输入转化为大模型的文本

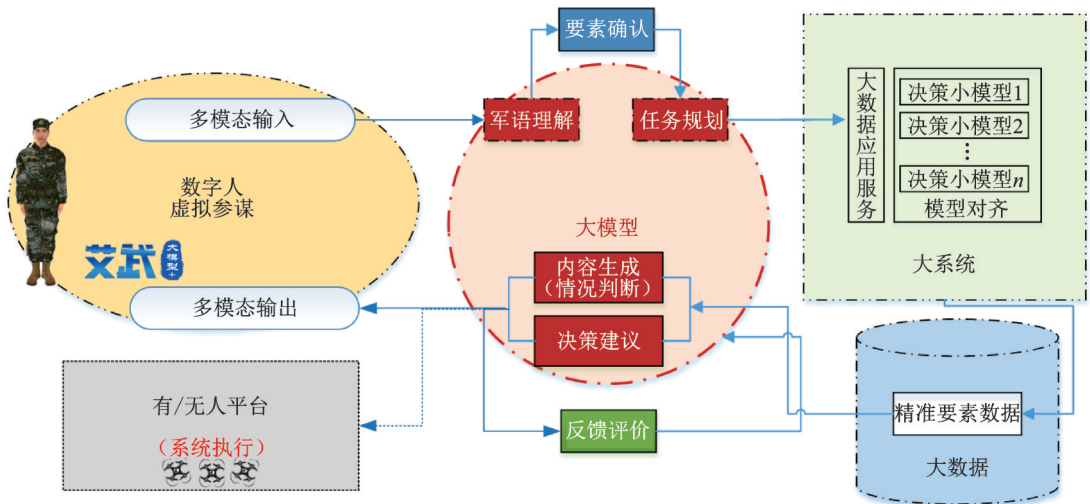


图4 “艾武大模型+”系统信息流程

Fig.4 Information process of “Aiwu large model+” system

输入要素。(2)大模型与大系统协同。大模型通过军语理解、要素确认,形成子任务规划,对齐大数据应用系统的任务小模型矩阵,调用地理信息、特定任务等应用服务。(3)大数据与大模型协同。大模型调用不同业务小模型,结合精准时空属性的军事领域特定知识数据、外挂知识库,形成精准数据分析和决策建议内容生成,增强大模型的精准指挥决策建议生成能力。

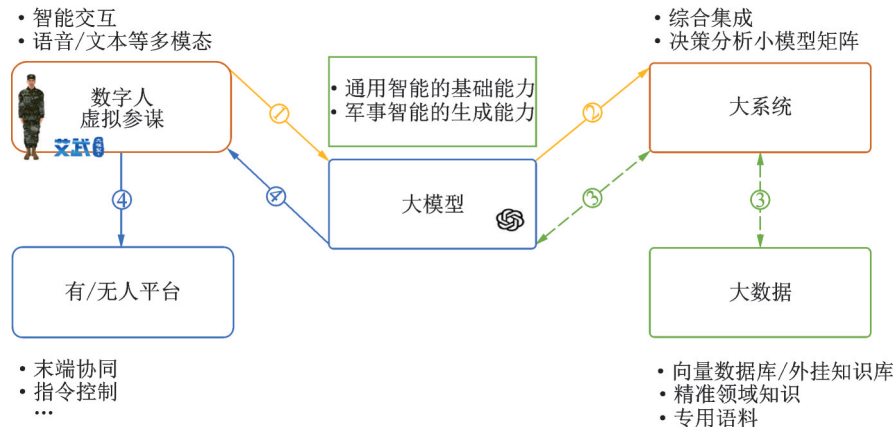


图5 “艾武大模型+”系统协同架构

Fig.5 Collaborative architecture of “Aiwu large model+” system

3 工程实现的关键技术

目前,仅靠大模型的内生能力无法满足智能化指挥所需的军语理解、任务规划以及精准指挥决策建议生成等核心需求。因此,需要从系统工程角度,综合集成基座大模型的多任务模型融合、精准数据与知识增强、反馈评价激励等训练调优和自学习自优化能力,达到“模型+数据+系统”的“1+1+1>3”的智能涌现效果。“艾武大模型+”系统的工程实现涉及如下关键技术。

(1)基于虚拟参谋“语音识别与合成”的综合集成,使大模型具备多模态的人机交互能力。作为智能化指挥的人机交互接口,数字人“艾武”以虚拟参谋为人设,突破鼠标、键盘和触控等交互方式的局

限,拓展大模型输入/输出侧的单一文本模态,增加数字人的语音识别/合成能力。基于语音识别(Automatic speech recognition, ASR)模型和语音合成(Text to speech, TTS)模型^[15],将指挥员的语音命令转化为文本,利用大模型的语义理解能力,转化为大系统可对齐理解和具体操作的指令,当大数据反馈精准数据到大模型后,可完成精准内容生成和文本形式输出和语音合成,实现数字人与指挥员的智能语音交互。

此外,大模型可扩展图像理解、姿态感知等更多模态交互能力,保证指挥员仅专注于指挥业务,集中精力于任务分析、判断、决策和组织协同,提升“艾武大模型+”系统在实际指挥应用中的操作便捷性和指挥效率。

(2)基于“指挥专用语料库”的提示工程训练,激发大模型的指挥领域语义理解能力。通用大模型缺乏垂直业务领域知识,受限于保密、法律法规等约束^[16],在指挥领域基本不具备实质性的内容生成和清晰准确的军语理解能力,因此有必要对大模型做指挥领域的二次训练、监督微调。工程实践中,利用提示工程进行指挥领域知识训练成效极高。按照“指令+上下文+输入文本”模式,完成指挥领域语义理解的简单提示和 Few-shot 提示,体现工程实践的简单可靠及效费比高等特点。此外,构建多层级的智能化指挥专用语料库,并基于向量知识库拓展大模型支持的指挥专用语料长度,完成外挂知识加载读取、分割与向量化、匹配与 Top K 推荐以及 Prompt 与语义回答生成等过程。

(3)基于大系统的“大模型+系列小模型+可视化”模式,使大模型具备任务规划和系统执行能力。鉴于大模型难以支撑复杂的军事任务规划、指控系统操作执行和典型场景可视化,基于团队研发的多任务应用服务大系统及模型算法平台,利用大模型的语义理解能力,对多模态信息进行感知、识别、抽取和理解,分解任务要素和子任务,实现各种小模型与各类指挥业务小模型的语义对齐匹配和按需调用,拉通基于大数据应用服务系统的定制化 OODA 环和 PREA 阶段^[17],拓展有/无人平台的智能操控。

(4)基于大数据的“精准结果数据”反向输入,使大模型具备精准内容生成能力。内容生成是大模型智能涌现的核心能力,但军事数据密级高、实时精准,通用大模型显然无法生成符合作战指挥要求的情报分析、情况判断、决心建议和协同计划等内容。结合团队业已汇聚的领域数据和外部数据,将精准军事数据融合到系列小模型,并反向输入大模型,形成二次分析报告^[18]。本质上,领域大数据可作为大模型的领域外挂知识库。此外,当精准数据实现知识图谱化后,基于大模型的作战指挥内容生成的有效性、准确性会产生质的提升。

(5)基于指挥员评价的“循环反馈提示”,使大模型具备自学习自优化的自我提升能力。指挥员对大模型每次执行任务情况不断“暗示”、反馈和评价^[19],能有效提高其垂直领域的业务能力。指挥员每次指令的内容生成都包含评价机制,并记录完整的语义理解、任务规划、系统执行过程和-content生成质量,双向优化对齐情况,循环反馈、评价奖励,不断提高大模型的自学习自优化能力和指挥决策建议的水平^[20],为后续能交互、能决策、能学习和自成长的演化奠定基础。

4 实证案例与选型分析

4.1 实证案例

军事智能场景具有极高的复杂性、多样性、不确定性,“艾武大模型+”系统是智能化指挥模式的全新探索,无先例可循,采用边研发、边优化调整的云原生“原型+增量”敏捷开发模式,围绕大模型的智能交互、语义理解、任务规划以及精准内容生成等功能,探索“数字人+大模型+大数据+大系统”的工程实践。

下面以公开网络信息和脱敏军事信息支撑“艾武大模型+”系统的任务案例数据组织,以开源大模型 GLM-6B 为基座,以抢险救援任务为背景,对“艾武大模型+”系统进行零样本测试,任务案例为某地发生地震,要求“艾武大模型+”系统查询周边救援力量,生成机动路线,形成救援行军计划。具体案例

实证过程中,按照大模型+OODA阶段信息流程,指挥员通过语音唤醒“艾武大模型+”系统,数字人将语音任务案例和智能交互信息转化为文本,基座大模型通过军语理解,抽取任务、人员和地点等关键要素信息,结合人在回路的要素确认,对齐调用抢险救援大数据平台,以及救援力量查询、机动路线规划等系列小模型,可视化呈现任务地域周边救援力量部署情况、推荐机动开进路线等救援要素;结合精准时空属性的救援任务数据和内部军事信息知识库,形成具有精准数据特征的救援行军计划报告,为后续抢险救援任务的高效实施奠定数据分析和模型决策基础。

4.2 选型分析

考虑到国防军事领域的独立网络部署环境和探索试验阶段的实际,综合大模型的中文语义理解、轻量级私有化部署、安全可控等测评维度,“艾武大模型+”系统的基座模型为GLM-6B,目前已适配到GLM3-6B版本。在4.1节实证案例中涉及大模型的提示工程、外挂知识库和任务子模型对齐等技术,具体能力指标包括军事实体抽取、计划报告生成和结果可解释性等。其中,实体抽取是理解军事提示输入的第一步,是对齐任务子模型的关键基础,而后利用大小模型对齐形成的关联性与因果性,有助于提升计划报告方生成的可解释性。

尽管中文语言理解测评基准(CLUE)、中文通用大模型综合性评测基准(SuperCLUE)以及综合性中文评估基准数据集(CMMLU)等在语言理解、生成、推理和跨模态交互等方面对大模型进行通用维度的测评,但军事领域大模型应用环境的网络封闭性、安全可控性,以及对模型结果精准性与可解性的要求极高。因此,在前期的选型分析中,利用有限算力环境(2块V100S GPU算力卡,单卡双精度浮点运算能力7T FLOPS,显存32 GB)模拟封闭的网络环境,本地部署了LLAMA-7B、GLM-6B、GLM3-6B大模型;同时,为比对他大模型的能力边界,利用互联网在线访问方式,分析文心大模型3.5、GLM-3、GLM-4等大模型服务。相关选型分析结果如表1所示。

表1 基座大模型选型分析
Table 1 Analysis of base large model selection

| 类别 | 大模型 | 参数规模/B | 联网检索能力 | 自主可控 | 结果可解释性 | 实体抽取能力 | 计划报告生成能力 |
|------|----------|--------|--------|----------|--------|--------|-------------|
| 本地部署 | LLAMA-7B | 7 | 无 | 美国Meta公司 | 差 | 可用 | 可参考(中文能力较弱) |
| | GLM-6B | 6 | 无 | 国产 | 差 | 可用 | 可参考 |
| | GLM3-6B | 6 | 无 | 国产 | 差 | 可用 | 可参考 |
| 在线联网 | 文心大模型3.5 | 20 | 无 | 国产 | 差 | 可用 | 可参考 |
| | GLM-3 | 130 | 无 | 国产 | 差 | 可用 | 可参考 |
| | GLM-4 | 130 | 有 | 国产 | 较强 | 可用 | 可参考 |

大语言模型的出色表现多源于Transformer架构、预训练和微调技术等,在基座大模型的选型对比过程中,Meta于2023年2月发布LLAMA-7B^[21],接受了20种语言、超1万亿tokens数据的训练,尽管2023年7月发布了免费可商用版本LLAMA-2^[22],将预训练的语料扩充到2T token,但中文训练数据只占很少一部分。因此,LLAMA系列模型是典型的非自主可控大模型,LLAMA-1的训练数据情况如表2所示。

GLM-6B^[5]针对中文问答和对话进行了优化,经过约1 TB tokens的中英双语训练,具备通用的关键词抽取和文本生成能力,INT4量化级别只需7 GB显存即可启动微调,但6 B的模型容量较小,记忆和语言能力较弱,在偏向性指示下易被误导;GLM3-6B于2023年发布,具有对话流畅、部署门槛低等特性,采用更多样的训练数据、更充分的训练步数和更合理的训练策略,在语义、数学、推理、代码和知识

表2 LLAMA-1的训练数据情况
Table 2 Training data of LLAMA-1

| 数据集 | 语种 | 采样比例/% | 数据量/GB | 描述 |
|----------------|-----|--------|--------|---|
| CommonCrawl | 英语 | 67 | 3 379 | 2017年至2020年的5个CommonCrawl数据集,用fastText分类器去除非英语页面,用n-gram模型过滤低质量内容 |
| C4 | 英语 | 15 | 783 | 预处理的CommonCrawl数据集 |
| Github | 英语 | 4.5 | 328 | 保留在Apache、BSD和MIT许可下分发的项目,对结果数据集进行去重 |
| 维基百科 | 多语种 | 4.5 | 83 | 2022年6月至8月间的维基百科数据,涵盖拉丁或西里尔文等20种语言 |
| 书籍 | 英语 | 4.5 | 85 | Gutenberg项目(包含公共领域的书籍)和ThePile的Books3的两个书籍语料库 |
| ArXiv | 英语 | 2.5 | 92 | ArXiv的Latex文件 |
| Stack Exchange | 英语 | 2 | 78 | 涵盖从计算机科学到化学等多种领域高质量问题和答案的网站数据 |

等不同角度数据集上的测评结果性能较强,是目前GLM-6B用户友好的升级版;2024年发布的GLM-4为130B底座,目前采用云服务访问,在中文对齐能力上整体超过GPT-4,适用于复杂的对话交互和深度内容创作设计场景,支持自行规划检索任务、与信息源交互;文心大模型3.5^[23]发布于2023年,安全体系方面能力突出,在伦理道德、隐私保护、违法犯罪以及负面诱导等方面均展示出强大的防护能力。

在实际联网使用过程中,GLM-3(参见官网<https://chatglm.cn/main/detail>)在通用内容生成后,会提供3个相关问题的联想,但军事场景需要精准问答,简单的联想推荐作用较弱,无法满足精准的时空数据保障需求;GLM-4(参见官网<https://chatglm.cn/main/alltoolsdetail>)会检索最新的网页资源,总结与问题相关的内容生成输出,但仅通过最新的网页公开资源,无法满足精准高效的决策保障需求;文心大模型3.5(参见官网<https://yiyi.baidu.com/>)的安全防护能力可作为本地化大模型安全的参考基准,但模型参数规模大,对算力要求高,在封闭网络环境下较难实现。此外,由于互联网版大模型通过采样温度、核取样等参数控制输出的随机性,通常多次问答结果会存在不一致现象。尤其,直接的军事问题会触发API策略阻止错误,提示输入或生成内容可能包含不安全或敏感内容。

综合上述选型分析,验证阶段的军事大模型应用需要平衡本地化部署成本和联网应用的不确定性、大规模参数的计算成本与中文自主可控能力,并兼顾系统集成的轻量级可调用需求。因此,“艾武大模型+”系统经过多轮迭代选型后,以GLM-6B作为基座模型,目前已适配到GLM3-6B版本,作为中文军事实体抽取、计划报告生成所需的军语理解与生成能力的核心;利用任务大数据系统和系列任务小模型增强军事任务执行的可解释性,弥补GLM-6B大模型参数规模的劣势;对标GLM-4的联网检索能力,“艾武大模型+”系统调用外挂的领域专用精准时空数据知识,可以生成体例规范、要素齐全和数据精准的计划报告。最终,综合集成语音识别、虚拟人、大模型、任务数据应用服务和外挂数据库,实现“数字人+大模型+大系统+大数据”架构,在一定程度上验证了利用大模型技术构建智能辅助决策系统的可行性。

目前,“艾武大模型+”系统通过信息系统系统集成,可以实现本地化安全部署、军事语义理解以及精准内容生成能力,性能基本符合预期。但由于当前场景训练数据规模较小,数据质量不高,后续仍需从体系架构角度持续重点攻关、系统集成,真正发挥大模型智能涌现的核心优势。

5 结束语

本文按照“理论研究-系统设计-工程实现-案例实证”的思路,梳理了面向智能化指挥的大模型能力需求,提出了面向智能化指挥的大模型“艾武大模型+”应用框架,给出了该模型筹划、准备、执行和评估的PREA过程,以及观察、判断、决策和行动的OODA阶段的两个尺度分析,设计了“数字人+大模型+大系统+大数据”的系统架构、信息流程以及3种协同关系,总结了多模态交互、提示工程训练、大小模型矩阵对齐、精准知识外挂以及循环反馈提示等关键工程技术。并以抢险救援任务的智能化指挥案例验证了“艾武大模型+”系统的实体抽取、计划报告生成能力,并进行了选型分析。下一步工作将聚焦于基座大模型二次预训练、偏好反馈优化在军语理解方面的性能提升。

参考文献:

- [1] 张维明,黄松平,朱承,等. 指挥控制的新范式:边缘指挥控制[J]. 指挥信息系统与技术, 2021, 12(1): 1-7.
ZHANG Weiming, HUANG Songping, ZHU Cheng, et al. A new paradigm of command and control: Edge command and control[J]. Command Information System and Technology, 2021, 12(1): 1-7.
- [2] OpenAI. Introducing ChatGPT[EB/OL]. (2022-12-30) [2024-03-30]. <https://openai.com/blog/chatgpt>.
- [3] 夏润泽,李丕绩. ChatGPT大模型技术发展与应用[J]. 数据采集与处理, 2023, 38(5): 1017-1034.
XIA Runze, LI Piji. Development and application of ChatGPT large model technology[J]. Journal of Data Acquisition and Processing, 2023, 38(5): 1017-1034.
- [4] 王亚坤,陈浩,葛悦涛,等. 2023年人工智能领域科技发展综述[J]. 战术导弹技术, 2024(1): 20-32, 67.
WANG Yakun, CHEN Hao, GE Yuetao, et al. Overview of technological development in the field of artificial intelligence in 2023 [J]. Tactical Missile Technology, 2024(1): 20-32, 67.
- [5] DU Z, QIAN Y, LIU X, et al. GLM: General language model pretraining with autoregressive blank infilling[EB/OL]. (2022-03-17) [2024-03-30]. <https://arxiv.org/abs/2103.10360>.
- [6] 罗文,王厚峰. 大语言模型评测综述[J]. 中文信息学报, 2024, 38(1): 1-23.
LUO Wen, WANG Houfeng. Review of large language model evaluation[J]. Chinese Journal of Information Science, 2024, 38(1): 1-23.
- [7] 孟燕兵,房超,薛颖. 智能化分布式指挥与控制建设及关键技术应用[J]. 国防科技, 2024, 45(1): 55-61.
MENG Yanbing, FANG Chao, XUE Ying. Intelligent distributed command and control construction and key technology applications [J]. National Defense Technology, 2024, 45(1): 55-61.
- [8] 徐皮克. 面向军事知识问答的问题语义解析关键技术研究[D]. 长沙: 国防科技大学, 2021.
XU Pike. Research on key technologies for semantic analysis of questions in military knowledge Q&A[D]. Changsha: National University of Defense Technology, 2021.
- [9] 崔中良. 通感、具身与意识:人机交互实现的感知基础[J]. 湖北大学学报(哲学社会科学版), 2024, 51(2): 143-152.
CUI Zhongliang. Synesthesia, embodiment, and consciousness: Perceived foundations of human computer interaction implementation[J]. Journal of Hubei University (Philosophy and Social Sciences Edition), 2024, 51(2): 143-152.
- [10] 程霄. 基于知识图谱的军事信息系统设计与实现[D]. 北京: 北京邮电大学, 2023.
CHENG Xiao. Design and implementation of military information systems based on knowledge graph[D]. Beijing: Beijing University of Posts and Telecommunications, 2023.
- [11] 邓启文,刘书雷,吴集. 国防颠覆性技术推动军事智能化发展[J]. 国防科技, 2023, 44(5): 46-50.
DENG Qiwen, LIU Shulei, WU Ji. Defense disruptive technology promotes military intelligence development[J]. Defense Technology, 2023, 44(5): 46-50.
- [12] 胡晓峰,郭圣明,贺筱媛. 指挥信息系统的智能化挑战——“深绿”计划及AlphaGo带来的启示与思考[J]. 指挥信息系统与技术, 2016, 7(3): 1-7.
HU Xiaofeng, GUO Shengming, HE Xiaoyuan. The intelligent challenge of command information systems: Inspiration and reflection from the “Deep Green” plan and AlphaGo[J]. Command Information Systems and Technology, 2016, 7(3): 1-7.

- [13] 徐国卿. 军事智能化建设发展战略问题研究[D]. 长沙: 国防科技大学, 2020.
XU Guoqing. Research on the development strategy of military intelligence construction[D]. Changsha: National University of Defense Technology, 2020.
- [14] ZENG A, LIU X, DU Z, et al. GLM-130B: An open bilingual pre-trained model[EB/OL].(2022-08-04)[2024-03-27]. <https://arxiv.org/abs/2210.02414>.
- [15] 王皓宇, JEON Eunah, 张卫强, 等. 基于声学模型共享的零资源韩语语音识别[J]. 数据采集与处理, 2023, 38(1): 93-100.
WANG Haoyu, JEON Eunah, ZHANG Weiqiang, et al. Zero resource Korean speech recognition based on acoustic model sharing [J]. Journal of Data Acquisition and Processing, 2023, 38(1): 93-100.
- [16] 余跃, 刘欣, 蒋芳清, 等. 支持鹏程系列开源大模型应用生态演化的可持续学习能力探索[J]. 智能科学与技术学报, 2022, 4(1): 97-108.
YU Yue, LIU Xin, JIANG Fangqing, et al. Exploration of sustainable learning capability supporting the ecological evolution of Pengcheng series open source large model applications[J]. Journal of Intelligent Science and Technology, 2022, 4(1): 97-108.
- [17] 马武彬, 朱承, 吴亚辉, 等. 面向边缘指控的数据赋能模式研究[C]//中国指挥与控制学会第十届中国指挥控制大会论文集(上册). 北京: 兵器工业出版社, 2022: 418-423.
MA Wubin, ZHU Cheng, WU Yahui, et al. Research on data empowerment models for edge accusations[C]//Proceedings of the 10th China Command and Control Conference. Beijing: Ordnance Industry Press, 2022: 418-423.
- [18] 岳增营, 叶霞, 刘睿珩. 基于语言模型的预训练技术研究综述[J]. 中文信息学报, 2021, 35(9): 15-29.
YUE Zengying, YE Xia, LIU Ruiheng. A review of pretraining techniques based on language models[J]. Chinese Journal of Information Science, 2021, 35(9): 15-29.
- [19] 熊明辉, 池骁. 论生成式大语言模型应用的安全性——以ChatGPT为例[J]. 山东社会科学, 2023(5): 79-90.
XIONG Minghui, CHI Xiao. On the security of generative large language model applications: Taking ChatGPT as an example [J]. Shandong Social Sciences, 2023(5): 79-90.
- [20] 韩毅, 乔林波, 李东升, 等. 知识增强型预训练语言模型综述[J]. 计算机科学与探索, 2022, 16(7): 1439-1461.
HAN Yi, QIAO Linbo, LI Dongsheng, et al. Overview of knowledge enhanced pre trained language models [J]. Computer Science and Exploration, 2022, 16(7): 1439-1461.
- [21] TOUVRON H, LAVRIL T, IZACARD G, et al. LLAMA: Open and efficient foundation language models[EB/OL]. [2023-04-02]. <https://arxiv.org/abs/2302.13971>.
- [22] TOUVRON H, MARTIN L, STONE K, et al. LLAMA 2: Open foundation and fine-tuned chat models[EB/OL]. (2023-07-19) [2024-03-14]. <https://arXiv.org/abs/2307.09288>.
- [23] SUN Y, WANG S, FENG S, et al. ERNIE 3.0: Large-scale knowledge enhanced pre-training for language understanding and generation[EB/OL]. (2021-07-05)[2024-03-21]. <https://arxiv.org/abs/2107.02137>.

作者简介:



崔翛龙(1973-),男,教授,
研究方向:智能化指挥与
指挥信息系统。



高志强(1989-),通信作者,
男,副教授,研究方向:军
事智能、联邦学习,E-mail:
1090398464@qq.com。



姬纬通(1990-),男,工程
师,研究方向:指挥智能
化,E-mail:jwt37223354@
126.com。



沈佳楠(2001-),男,本科,
研究方向:人工智能,
E-mail: 1120331465@qq.
com。



张敏(1988-),女,编辑,研究
方向:智能化指挥,E-mail:
270648115@qq.com。



邱鑫源(1999-),女,博士
生,研究方向:人工智能、
智能化指挥,E-mail:
787942392@qq.com。

(编辑:刘彦东)