电磁频谱大模型赋能资源认知决策理论与方法

汇报人：周福辉 教授/博导

工作单位：南京航空航天大学

感谢电磁频谱认知智能实验室成员：张浩、刘纯玉、吴雨航、袁璐、王梓、丁锐、徐铭、杨媛、刁羿赫、曾智杰等博士

微信二维码

1

研究背景

2

频谱大模型的构建

3

频谱大模型的初步成果

4

频谱大模型验证平台

5

研究展望

频谱资源是国家战略资源

《“十四五”国家无线电管理和发展规划(2021年)》明确指出电磁频谱资源是**国有并具战略意义的稀缺资源**

工业和信息化部组织召开《“十四五”国家无线电管理和发展规划》
宣贯电视电话会议

发布时间: 2021-11-30 15:10 来源: 无线电管理局

2021年11月29日,《“十四五”国家无线电管理和发展规划》宣贯电视电话会议在京召开。工业和信息化部党组成员、总工程师田玉龙出席会议并讲话。

- 《中长期科技发展规划纲要》将“电磁频谱”列为**前沿技术和重大科技主题**。—2010年1月国务院
- 全面贯彻新发展理念, **统筹规划和高效利用频谱资源**, 要**统筹发展和安全**, **维护好空中电波秩序**
- 随着“**低空经济**”、“**商业航天**”、“**中国制造2025**”“**等新质生产力和国家战略的实施**, **迫切需要更多的频谱资源**

电磁空间是第六维作战空间

2019年美国发布《制胜无形之战：在电磁频谱内获取持续性优势》报告明确指出失去制电磁权，必失去制海与制空权

- 2020年10月，美军发布了《电磁频谱优势战略》
- 美国战略司令部，要实现陆、海、空、天和赛博空间优势，必须首先夺取电磁频谱优势
- 俄乌战争中，俄罗斯未获取显著的电磁频谱制胜，导致作战进展缓慢
- 阿塞拜疆利用先进的电磁频谱战设备致盲了亚美尼亚的防空电磁体系

- 《国家中长期频率战略规划》为首次制定的国家层面**中长期电频谱战略**
- 2023年11月16日，**北斗标准**正式被国际民航组织采纳，列入《国际民航公约》附录10并正式生效，与GPS和GLONASS一同成为了**国际公认的GNSS系统之一**

频谱战略和政策的研究为保障我国**频谱利益**，推进国家重大工程走出国门奠定基础

频谱监测设备

频谱相关**国际标准的制定**有力推动相关产业发展，将产生**显著经济社会效益**

移动互联网

通信速率更快

物联网

终端接入更多

6G天地一体化信息网络

网络覆盖更广

万物互联、6G天地一体的网络空间，迫切需要更多的电磁资源

电磁空间智能认知和决策是实现电磁空间频谱智能管控的核心，具有重要意义

2023.10

国家频谱战略

明确指出美国将**无线电频谱**作为最重要的国家资源之一

2024.03

国家频谱实施计划

为频谱战略做出**更详细工作分解**

2024.10

国家频谱研发计划

强调频谱相关**创新与优先发展**领域

美国一直以来高度重视**人工智能**应用于频谱管控的重要战略意义，持续布局相关研究

➤ 电磁环境和频谱任务具有**高度动态性**：

随着越来越多的无线终端、通信系统、雷达、干扰机和其他设备在频谱环境中运行，电磁环境变得越来越复杂，任务不仅呈现高度动态性，而且还具突发性，**使得对电磁频谱智能管控算法和模型的泛化能力要求极高。**

360伪基站追踪系统展示的北京全局态势

➤ 频谱领域知识难以利用和迁移:

电磁频谱领域知识的利用是解决电磁频谱智能管控面临高质量样本数据少、复杂电磁空间认知难、和资源决策速度要求高的关键。然而，低、中、高频段频谱演变规律、衰落模型等特性不同，**频谱的差异性导致跨频段知识难以有效利用与迁移。**

➤ 可解释性和推理能力有限

传统深度学习方法通常面临**黑盒决策**问题，难以清晰解释其识别或决策过程，在复杂电磁环境下，新场景，**推理能力严重不足**；而大模型通过**多层次知识关联和注意力机制**，能够对结果提供**直观解释**，并能基于历史经验**推理未知模式**，这是实现高可靠电磁认知与决策的关键突破。

通用智能：采用**自监督预训练方法**从海量通用域**无标注数据**中自动学习知识，以**大模型参数**存储知识

优势：无标注数据廉价；大模型有力支持知识的学习和存储

2020年1750亿参数大模型GPT-3发布，具备**解决数十种文本任务的能力**，人工智能进入**通用智能时代**

2023年发布的ChatGPT和GPT-4 **具有通用认知能力**，可通过文字、视觉、语音信号**与人类交互**，可使用人类工具处理复杂任务，已被用于政务、商务等众多领域进行**智慧赋能**

通用域无标注数据

自监督预训练学习

通用大模型

问：计算机英文是什么？
答：Computer.
问：如何治疗疟疾？
答：青蒿素。
问：1+2=?
答：3

通用知识

通用智能的出现为应对电磁空间诸多挑战提供了新范式

大模型技术提供了从无标注大数据萃取世界知识的通用范式

能力涌现

随着计算量和数据量的增大，大模型会涌现出解决复杂任务的能力

知识筑基

人类对齐

大模型可以与人类的需求、期望、价值观、伦理道德等进行深度对齐

能力激发

序列化世界

大模型技术能够对任何可以被序列化的事物进行压缩和学习

范式拓展

大模型可以解决频谱领域的诸多难题

各国频谱规划不一致，难以利用

大模型可以用于构建
频谱知识库，助力全球
频谱管理

频谱任务众多，现有方法难以通用

大模型在频谱认知、决策等
多个任务中具有较强的泛化性

用频场景多变，现有规划不够全面

大模型作为通用工具的存在，能全面应对多种
用频场景

面向国家战略需求和国际学术前沿，突破关键技术，提升关键指标

1

研究背景

2

频谱大模型的构建

3

频谱大模型的初步成果

4

频谱大模型验证平台

5

研究展望

- 原始IQ数据：监测设备收集的原始信号数据形式
- 功率谱密度：记录信号的功率如何在不同频率上分布
- 时频图：信号随时间变化的频率特性
- 空间信息：信号的空间分布或定位

公开数据集

自建数据集

- ◆ 实时监测数据：实时监测频谱环境，监测信号、频谱态势等
- ◆ 历史监测数据：监测信号、频谱态势等历史记录
- ◆ 设备信息数据：设备技术参数等信息
- ◆ 专项任务数据：重大活动保障、投诉处理数据、违法信号监测、干扰源定位、专项检查数据等

实时监测数据

实时监测数据

设备信息数据

专项任务数据

- ◆ 软件仿真数据：MATLAB、Python、GNU Radio、LabVIEW、Simulink仿真数据
- ◆ 场景仿真数据：室内场景、室外场景、复杂电磁环境、动态移动场景、多用户环境仿真数据
- ◆ 算法仿真数据：信号检测算法、信号调制识别、参数估计算法、定位算法、干扰分析算法仿真数据
- ◆ 信道仿真数据：传播效应、信道环境、干扰噪声等仿真数据

软件仿真数据

场景仿真数据

算法仿真数据

信道仿真数据

- ◆ 时间信息：监测环境对应的**时间**信息及记录
- ◆ 位置信息：监测环境对应的**位置**信息及记录
- ◆ 环境信息：监测环境对应的**环境**信息及记录
- ◆ 用频场景信息：监测环境对应的**用频**信息、**地形**信息等

时间信息

位置信息

环境信息

用频信息

- ◆ **用户行为记录**：用户的频谱接入等相关记录
- ◆ **频谱管理记录**：频谱管理等相关记录
- ◆ **异常信号记录**：异常信号登记信息
- ◆ **干扰信息记录**：干扰信号登记信息

用户行为记录

频谱管理记录

异常信号记录

干扰信息记录

- ◆ 标准文件：ITU、3GPP、IEEE等国际国内标准文件
- ◆ 规则及政策：国际频谱分配规则、国内划分规定等文件
- ◆ 文献及专利：专业文献、百科、专利等文件
- ◆ 其他文本：设备说明、用户手册等内容

标准文件

规则及政策文件

文献及专利

其他文本

- 数学理论基础：傅里叶变换、小波变换、希尔伯特变换等**数学工具**
- 物理原理：电磁波传播理论、频率与波长关系、频谱能量**分布规律**
- 信号处理原理：采样定理、频域分析方法、调制解调技术等**信号处理原理**
- 测量原理：频谱分析仪工作原理、测量方法学、误差分析等**测量原理**

数学理论基础

物理原理

信号处理原理

测量原理

- **仿真结果**：通过大量仿真实验积累的数据和现象认知，形成对系统行为的**经验性理解和预判**。
- **频谱演化规律**：在长期频谱监测和分析中，总结出的频谱资源使用**变化趋势和特征性规律**。
- **逻辑规则**：在频谱管理和应用实践中，归纳出的各种判断准则和处理流程的**经验性方法**。

仿真结果

频谱演化规律

逻辑规则

- 属性知识：指对频谱特征的**描述性知识**，包括频率、带宽、发射功率、调制方式等基本属性参数及其定义和含义
- 频谱语义：基于**语义特征提取模型**生成知识图谱下的**语义嵌入向量**

属性知识

频谱语义

- 采用**自注意力机制**，可以直接建模序列中任意位置间的依赖关系
- 基于编码器-解码器架构，通过多头注意力实现并行计算和多层次特征提取
- 可并行计算，**训练效率高**，时空复杂度为 $O(N*N)$ ，**全局建模能力强**

- **关系推理**: 基于**知识表达**、**概率图模型**、**深度学习**的关系推理方法
- **规则匹配**: 利用**预定义规则**对知识数据进行模式识别和处理
- **知识补全**: 基于**已知信息推测缺失**的知识。

基于概率图模型的关系推理技术路线示意图

基于深度学习的关系推理技术路线示意图

关系推理

规则匹配

知识补全

- **特征层对齐**: 负责处理不同模态数据在表征层面的映射关系, 确保基础特征的匹配准确性。
- **语义层对齐**: 专注于更高层面的语义理解, 协调不同模态在含义层面的一致性表达。
- **时序层对齐**: 主要解决多模态数据的时序匹配问题, 处理不同模态之间的时间依赖关系。
- **对比学习**: 通过构建合适的学习策略, 增强不同模态数据之间的相关性, 提升整体对齐效果。

- ◆ 监督训练：利用具有**标签**的数据进行训练
- ◆ 半监督/自监督训练：通过**少量标签**进行半监督训练或**无标签**数据进行自监督训练
- ◆ 掩码训练：通过隐去部分数据，进行掩码学习
- ◆ 自回归训练：通过自身数据的下个时刻，进行训练
- ◆ 对比训练：利用同类、异类样本进行训练

监督训练

自监督训练

掩码训练

自回归训练

对比训练

- 专用任务层：负责将输入信息转化为特定领域的分析和决策结果。
- 任务输出适配：多任务损失加权、任务依赖建模、输出校准优化

频谱大模型

专用任务层

- ◆ **模型压缩与优化**: 负责降低模型复杂度, 通过量化、剪枝等技术实现模型的轻量化。
- ◆ **云边端协同部署**: 聚焦于不同计算资源的调度和分配, 实现云端和终端的高效协作。
- ◆ **无线通信资源优化**: 专注于优化数据传输效率, 降低网络延迟和带宽消耗。

利用知识图谱增强大模型领域知识推理能力

- ◆ 把知识图谱当作一个准确的知识库，作为一个**外部检索的知识源**，提升领域知识方面的推理能力；
- ◆ 利用**知识图谱的多跳路径**作为大模型的输入，提升模型专业性、可信性、真实性与可解释性；
- ◆ 使用**知识模型**根据问题生成知识陈述。

利用知识图谱增强大模型推理的可解释性

- ◆ 通过动态知识图谱融合增强大模型的推理能力，从知识图谱中较容易地获得一系列规则，既从数据中总结得出的特征子图，作为COT、TOT的指令。
- ◆ 通过**检索增强知识融合增强大模型的推理能力**，知识图谱是典型的图结构，有大量的路径。借助这种链式关系可提升大模型推理的可解释性。
- ◆ 利用大语言模型对基于知识图谱生成的问题进行预测，验证大模型的可解释性。

1

研究背景

2

频谱大模型的构建

3

频谱大模型的初步成果

4

频谱大模型验证平台

5

研究展望

SpectrumFM: A Foundation Model for Intelligent Spectrum Management

Fuhui Zhou, *Senior Member, IEEE*, Chunyu Liu, Hao Zhang, *Graduate Student Member, IEEE*, Wei Wu, *Member, IEEE*, Qihui Wu, *Fellow, IEEE*, Derrick Wing Kwan Ng, *Fellow, IEEE*, Tony Q. S. Quek, *Fellow, IEEE*, and Chan-Byoung Chae, *Fellow, IEEE*

论文链接: <https://arxiv.org/abs/2505.06256>

代码链接: <https://github.com/ChunyuLiu188/SpectrumFM.git>

- **数据收集与处理**：从多个来源收集不同的频谱数据并进行预处理，以确保数据集之间的一致性和兼容性。
- **预训练阶段**：模型通过自监督学习任务，即**掩码重建**和**下一时隙信号预测**，学习可迁移的频谱嵌入。
- **微调阶段**：调整预训练模型以适应特定的下游任务，包括**调制方式识别**，**协议识别**，**频谱感知**和**异常检测**。

数据采集与大模型部署平台

下游任务二：无线接入方式识别

无线接入方式识别是识别接收信号所属的接入协议类型（如 Wi-Fi、蓝牙、ZigBee、LTE 等），大模型通过端到端学习协议特征与通信行为，在**多协议共存、信号混叠和低信噪比**等复杂场景中具备更强的**区分能力、鲁棒性与小样本能力**。

Models	Precision	Recall	F1 Score
ResNet	0.6723	0.6950	0.6764
VGG	0.7542	0.7404	0.7356
CNN2	0.7851	0.7509	0.7407
AMC_Net	0.6492	0.6472	0.6405
DAE	0.6826	0.6735	0.6487
CGDNN	0.7815	0.7488	0.7397
MCNet	0.7210	0.7320	0.7116
GRU2	0.7286	0.7429	0.7225
MSNet	0.7696	0.7546	0.7481
Transformer	0.7617	0.7463	0.7415
Our	0.8270	0.8253	0.8226

下游任务三：频谱感知

频谱感知是识别无线信道在特定时刻是否被占用的任务，大模型能够在低信噪比和非理想条件下更准确地判断频谱空闲状态，并显著提升了频谱感知的鲁棒性与检测性能。

下游任务四：异常检测

- 正常信号由 Ceyear 1435B-V RF 信号源生成，采用 QPSK 调制，载频为 2.9 GHz，带宽为 20 MHz
- 单音干扰：使用一个正弦信号在 2.89–2.91 GHz 频段内扫频，模拟简单的窄带干扰
- 多音干扰：由多个正弦信号组成，在同一频段内扫频，模拟频域上更复杂的宽带干扰情形
- 混叠干扰：生成带宽为 10 MHz 和 30 MHz 的随机信号，频率范围覆盖 2.885–2.915 GHz，与正常信号在频域产生混叠；进一步对这些随机信号施加 QPSK 调制，使其在调制特性上更接近正常通信信号，从而显著增加异常检测的难度

(a) ROC curve for single-tone interference.

(b) ROC curve for multi-tone interference.

(c) ROC curve for 10M aliased-signal interference.

(d) ROC curve for 30M aliased-signal interference.

预测未来频谱占用情况

Models	metrics			
	Accuracy(%)	Precision(%)	Recall(%)	F1
VLM4SP(proposed)	97.0207	91.1894	99.3892	0.9522
VLM4SP_no_text	90.5251	81.0256	91.6577	0.8601
VLM4SP_no_global	95.6189	92.1501	96.4286	0.9424
VLM4SP_no_local	93.8318	86.5277	93.6306	0.8993
VLM4SP_simple_fusion	88.2568	78.1642	97.3569	0.8671

- 提出了一种基于视觉Transformer的频谱图构建大型模型——GenSpectraLM。该模型受BERT掩码语义推理和MAE局部-全局构建机制的启发，采用自监督掩码预训练方式，从多样化数据集中隐式学习电磁传播模式，并通过微调实现跨场景泛化。
- 仿真结果表明，GenSpectraLM在25%的采样率下实现了均方根误差为1.3286的精确频谱图构建，性能比基准方法提高约40%，有效解决了复杂环境中的数据效率挑战。

GenSpectraLM架构

GenSpectraLM重建效果 (掩码比75%)

RMSE性能对比

名称	RMSE	Mask Ratio
GenSpectraLM(fine-tuned)	1.3286	75%
GenSpectraLM(pretrained)	1.6849	75%
IDW	2.2607	75%
KNN	2.2111	75%
Kriging	2.3465	75%

突破基于电磁频谱大模型的资源快速决策方法，通过融合全局感知、强泛化能力和端到端优化机制，为复杂动态的频谱资源分配任务提供了更智能、高效与可拓展的解决方案，提升了频谱效率和电磁对抗能力

- 现状分析：开展频段使用调查、占用度评估、效率分析和干扰评估
- 需求预测：研究业务趋势、量化预测需求、评估新技术需求和区域特点
- 方案制定：规划频段划分、频率复用、干扰协调和频谱共享
- 效益评估：评估经济价值、定价机制、投资收益和社会效益

频谱大模型

低空频率规划

- 融合了来自 3GPP、IEEE、ITU 的标准文档、论文与专利信息，构建了一个大规模频谱知识图谱，用于组织频谱相关的结构化知识，支撑大模型的理解与推理能力。
- 通过引入结构化频谱知识，实现基于知识图谱的频谱问答，提高大模型对专业领域问题的响应准确性与可解释性。

频谱大模型智能管理平台

首页 对话 频谱 知识图谱 知识库 警告

频谱大模型

IEEE

访问受限：仅有授权机构（如清华大学）可获取完整文档，其他用户需通过 IEEE 官方渠道购买。
技术覆盖不完整：当前信息未提供具体的术语修订细节，需结合原文进一步分析。

若您需要更详细的技术修订内容或具体案例分析，请补充相关文档或具体问题。

Pro/deepseek-ai/DeepSeek-R1 人知识图谱

2.4GHZ

推理过程

好吧，我需要处理用户关于2.4 GHz频段的查询，并结合给的知识图谱信息生成对话。先看看知识图谱里的数据点。有不同日期和时间的数据，比如2006年7月28日的几个时间点的频率和信号强度数据。用户提供了一些Marker的具体数值，比如Marker1在2.456 GHz处-29.90 dBm，还有256 QAM调制方式，操作频段是16。

首先，我需要理解这些数据的含义。不同的Marker可能代表不同的频率点上的信号强度测量结果。比如，在04:26:52这个时间点，各Marker的频率分别是2.456、2.4375、2.4905、2.4692 GHz，对应的信号强度从-29.9 dBm到-60.07 dBm不等。随后，在05:10:33，又有一个测量结果。还有提到操作频段16和256QAM调制，这可能涉及Wi-Fi或者其他通信标准。

需要确定的是，用户可能是在分析某个时间段内2.4GHz频段的信号分布或干扰情况。这些数据可能来自于频谱分析仪的输出。256QAM是一种高阶调制方式，通常用于提高数据传输速率，但需要较高的信噪比。操作频段16可能涉及具体的信道划分，比如Wi-Fi的某些信道。

接下来，我需要将这些信息结合起来，解释不同时间和频率的信号强度变化、可能的原因，以及调制方式对系统性能的影响。

输入问题。

1

研究背景

2

频谱大模型的构建

3

频谱大模型的初步成果

4

频谱大模型验证平台

5

研究展望

室内验证平台

计算平台

1

研究背景

2

频谱大模型的构建理论

3

频谱大模型的利用方法

4

频谱大模型的应用场景

5

研究展望

民用电磁频谱应用

电磁干扰精准识别与消除

所需数据:

- 各类干扰源特征库
- 多种环境下干扰传播模型数据

所需知识:

- 干扰分类与特征提取技术
- 干扰消除与抑制算法
- 干扰源定位方法

电磁波传播预测与规划

所需数据:

- 高精度地形地貌数据
- 建筑物电磁特性数据库

所需知识:

- 射线追踪与衍射计算方法
- 多尺度传播模型
- 大气及天气影响建模

低空经济频谱规划

所需数据:

- 无人机飞行走廊电磁环境数据
- 城市低空三维电磁场分布

所需知识:

- 低空通信专用频段规划
- 无人机集群通信协议
- 空地一体化网络架构

卫星通信频谱优化

所需数据:

- 卫星轨道与覆盖区域数据
- 空间电磁环境与干扰数据

所需知识:

- 星地频谱共享技术
- 星地传播模型

军事电磁频谱应用

复杂电磁环境下信号识别

所需数据:

- 多源异构信号样本库
- 低信噪比环境下信号数据

所需知识:

- 深度学习信号特征提取
- 新型调制解调技术
- 信号盲分离算法

主动式频谱感知与管控

所需数据:

- 战场电磁环境实测数据
- 敌方频谱使用模式数据库

所需知识:

- 认知无线电技术
- 分布式频谱感知方法
- 频谱态势评估模型

电子对抗效能评估

所需数据:

- 干扰效能实测数据集
- 对抗环境下通信质量数据

所需知识:

- 电子对抗效能评估方法
- 抗干扰通信技术
- 战场电磁环境建模

电磁信号智能伪装与欺骗

所需数据:

- 目标特征电磁信号库
- 敌方信号识别系统特性数据

所需知识:

- 信号形态学变换技术
- 对抗样本生成方法
- 智能欺骗策略

感谢电磁频谱认知智能实验室成员：吴雨航、张浩、袁璐、王梓、丁锐、徐铭、刘纯玉、杨媛、刁翊赫、曾智杰等博士

服务电磁频谱空间国家利益
谢谢!

Cog.nuaa.edu.cn

(天地一体频谱认知智能实验室: S³CI Lab)