大模型赋能移动通信系统

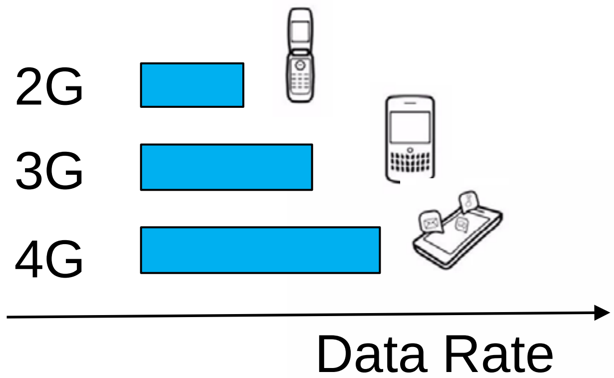
向路平、彭于波
南京大学智慧网络与通信研究院

2025年9月19日

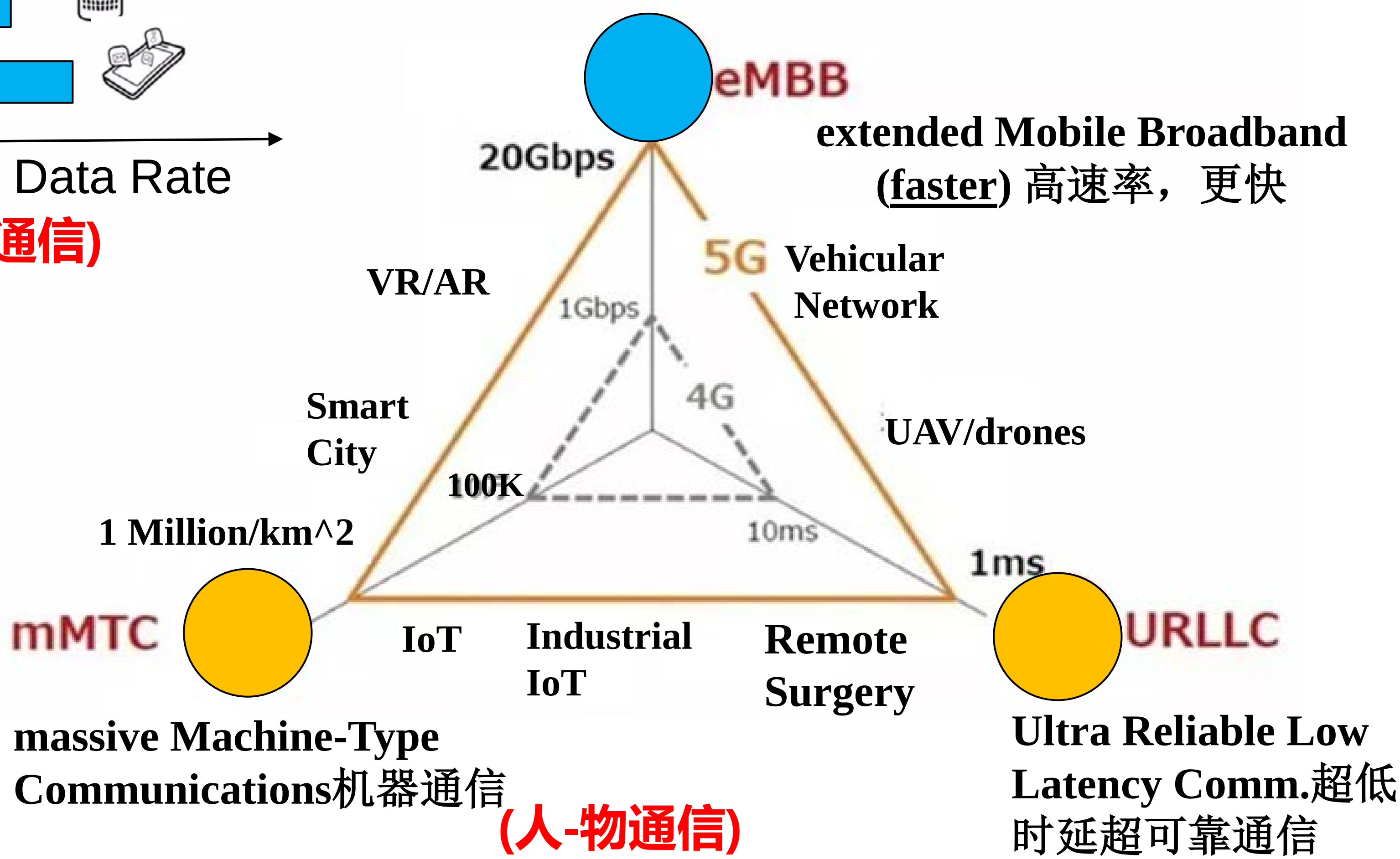
Agenda

- 背景和驱动
- 大模型赋能语义通信
- 大模型驱动数字孪生网络
- 大模型增强O-RAN
- 总结

移动通信系统的演进: 2G-5G



(人与人通信)



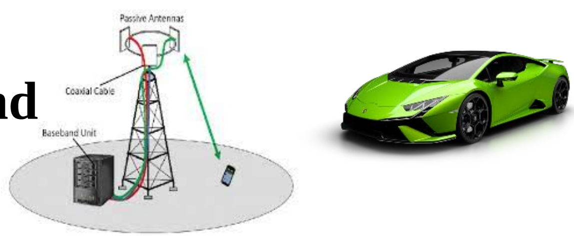
(人-物通信)

移动通信系统的演进：6G

Space
空

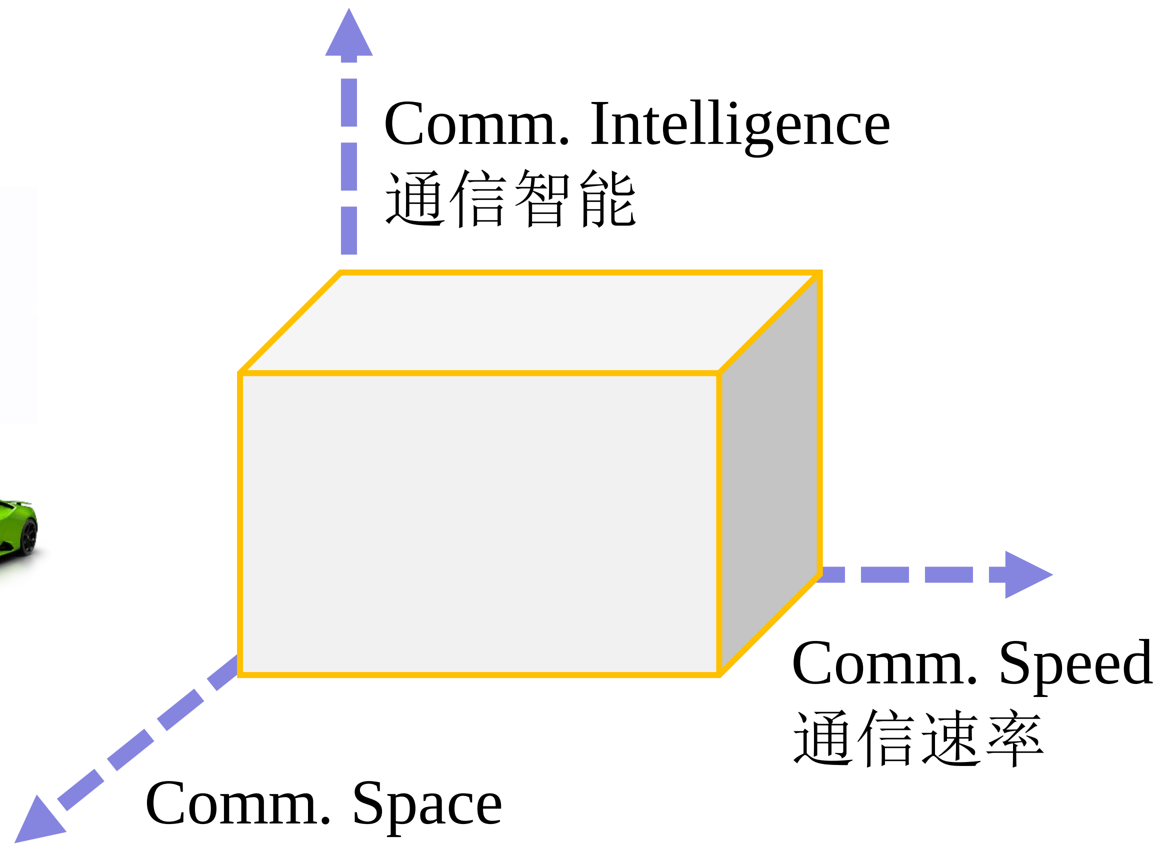
Sky
天

Ground
地

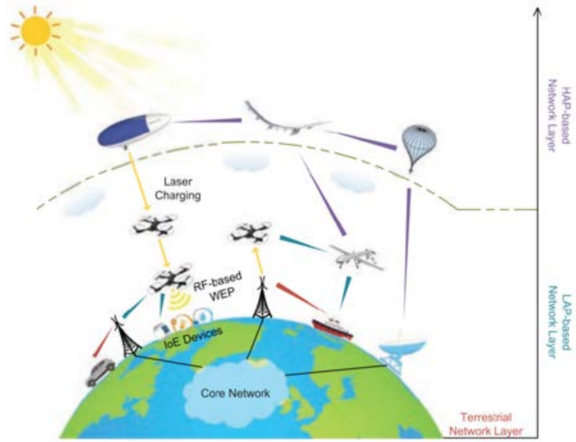


Sea
海

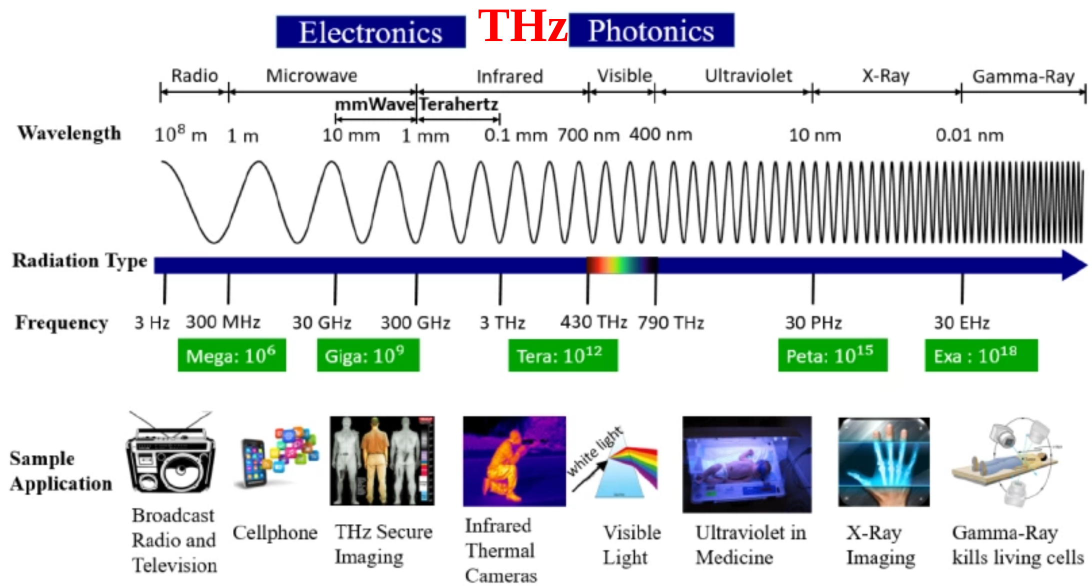
Under water



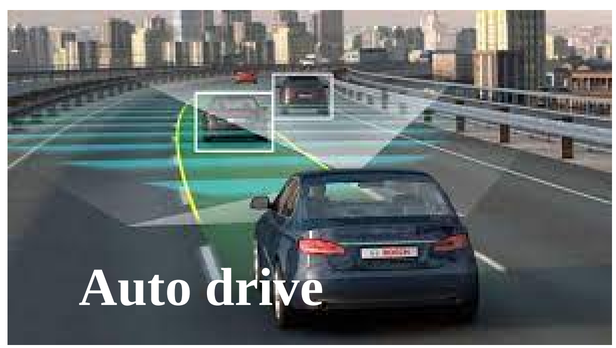
全场景-全频段-全服务



All scenarios: space-sky-ground-sea 全场景-空天地海



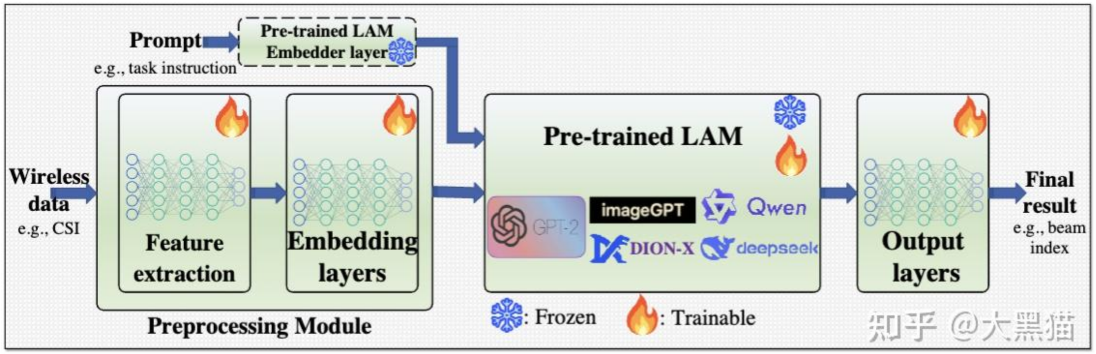
Full Spectrum全频谱



All Services 全服务

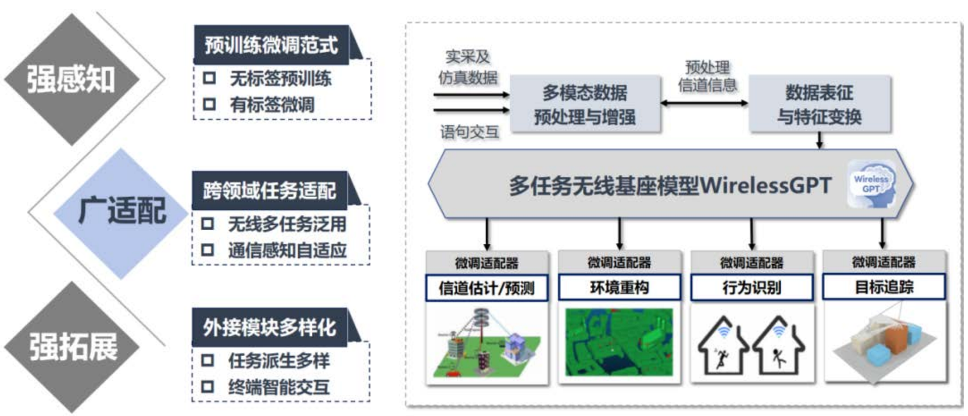
如何构建未来的移动通信网络?

大模型赋能通信行业



东南大学提出用于无线物理层的大模型

华为提出面向多模态任务的盘古大模型



鹏城实验室研发了全球首个面向通信与感知一体化（ISAC）的多任务预训练无线大模型WirelessGPT

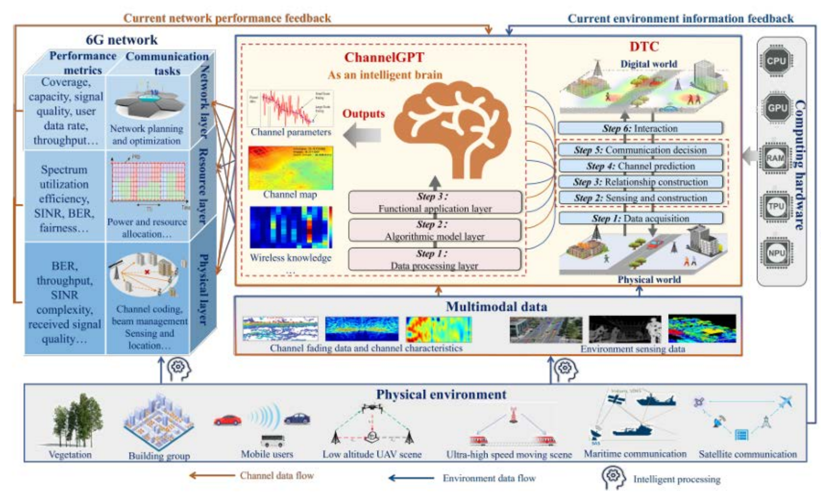
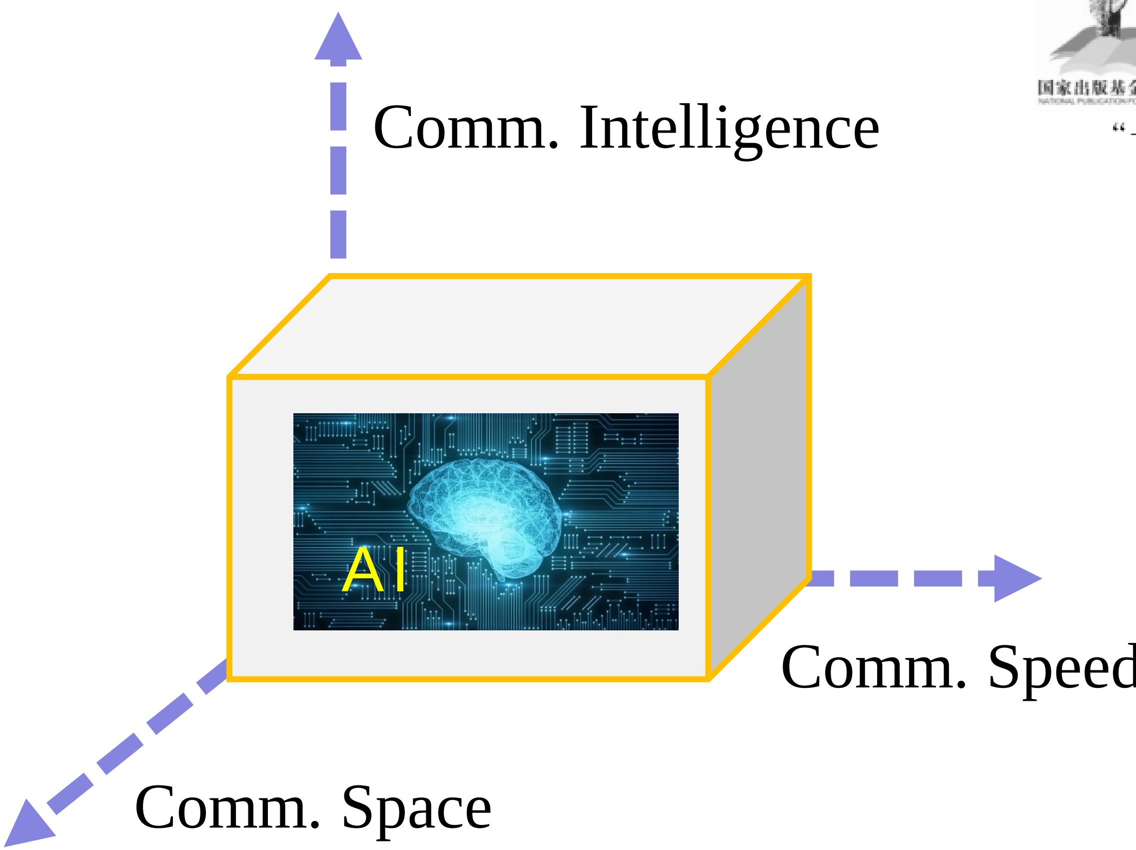


Fig. 1. ChannelGPT enabled 6G architecture with EI.

北邮团队提出ChannelGPT

通信行业积极拥抱大模型！

Native Intelligence in 6G
6G内生智能



All unified via AI

国家出版基金项目
NATIONAL PUBLICATION FOUNDATION

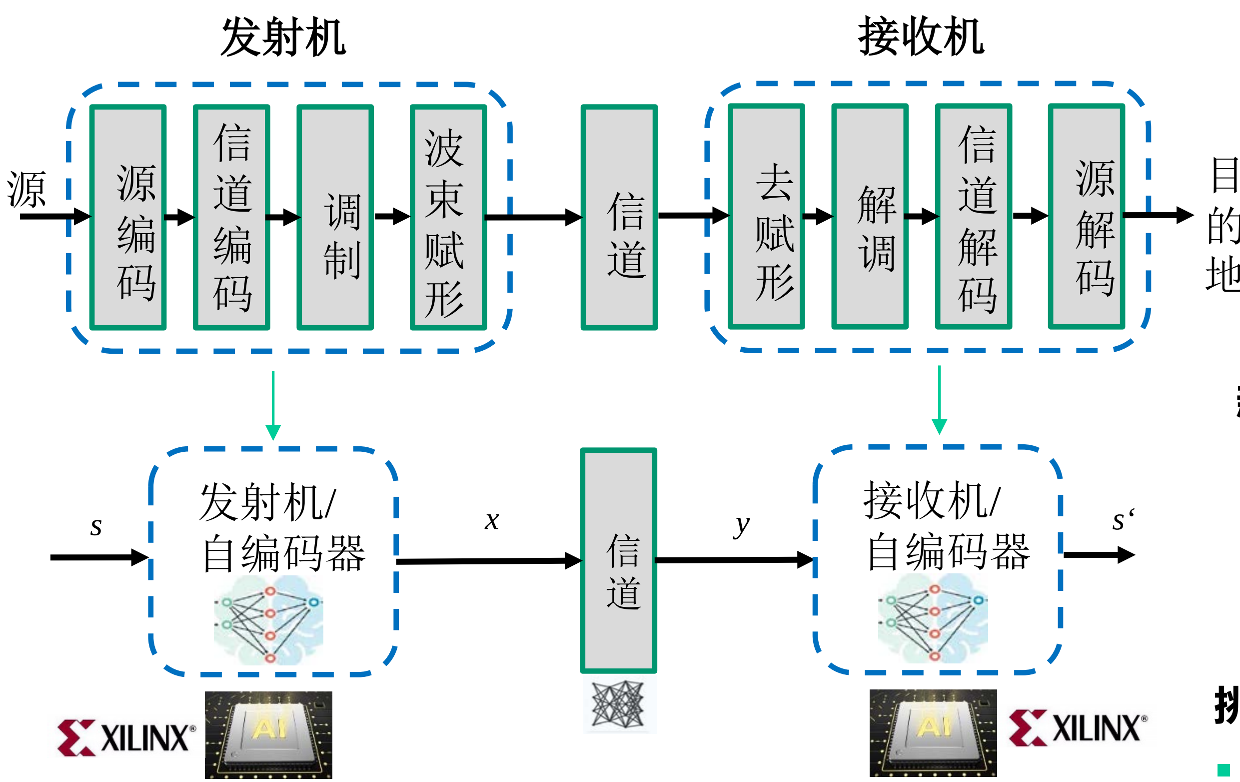
“十四五”时期国家重点出版物出版专项规划项目
新一代人工智能理论、技术及应用丛书

通信网络内生智能

杨 鲲 向路平 高 阳 著

- 端到端设计
  - 物理层
  - 网络架构：O-RAN/AI RAN
- 语义通信
- 数字孪生网络

端到端设计6G内生智能



传统通信系统

- 功能模块独立设计，难以全局优化
- E.g., 源编码**压缩**vs信道编码**加冗余**

新型自驾通信系统

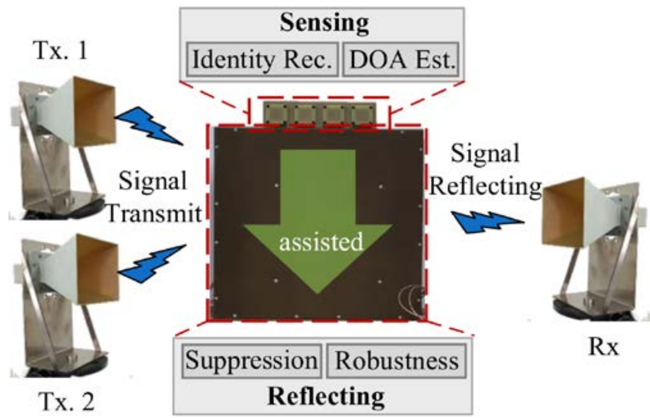
- 端到端设计，便于全局优化
- 使用机器学习
- 革命性变革

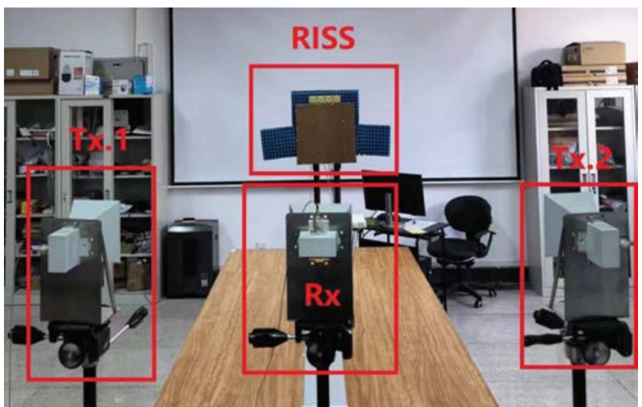
挑战

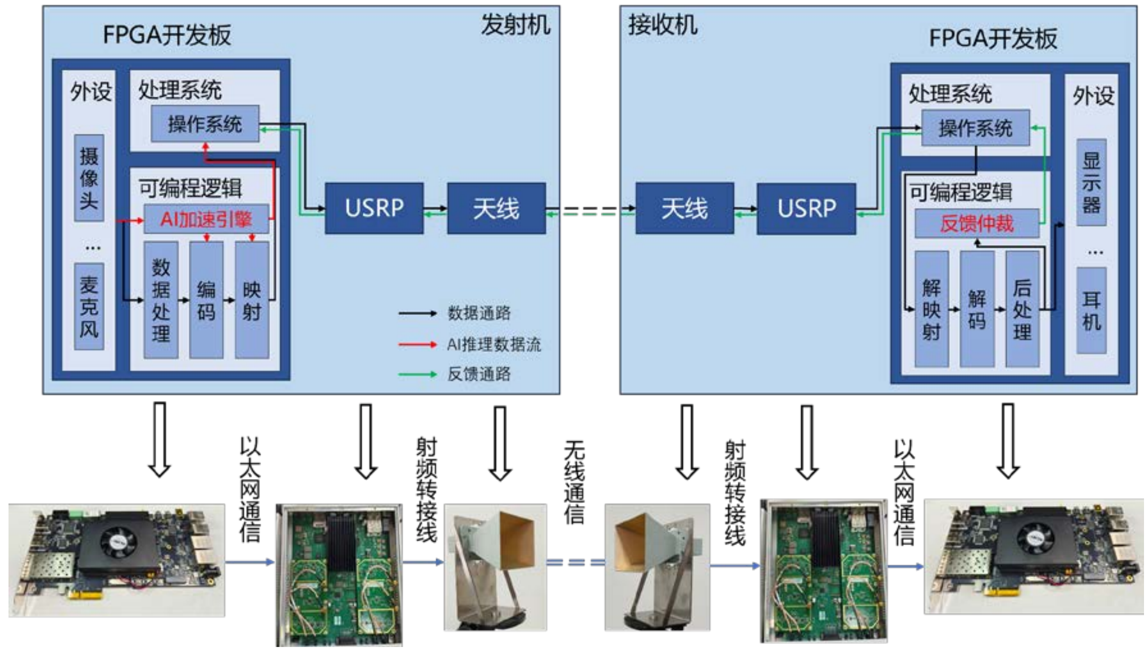
- **通用/泛化**机器学习模型
- **数据**来源和质量

自驾通信网络架构： O-RAN/AI-RAN + RIC (Radio Intelligent Controller) + AI，给出支持内生智能的新网络架构

物理层智能通信原型系统







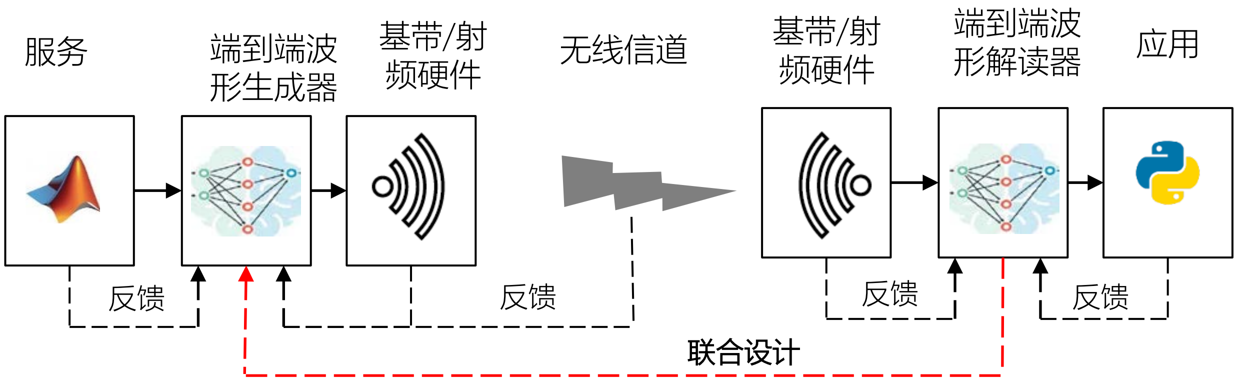
▶ 设计核心：**AI加速引擎，反馈仲裁模块**

▶ 设计核心：**感知辅助通信，干扰抑制**

系统演示视频链接 https://www.bilibili.com/video/BV1Xm2gYCEDt/?spm_id_from=333.337.searchcard.all.click&vd_source=1b193fe53cdb1b55c9ce57b2fa798a9d

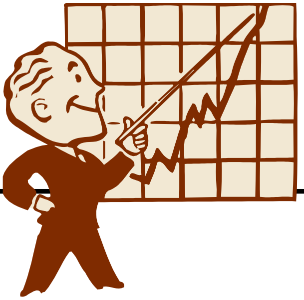
大模型打破标准桎梏

物理层：5G/6G要赋能千行百业，但应用场景各异，对通信波形的要求也各自不同，长周期的标准化进程难以适应应用要求的快速变化，设计**非标准化波形**可能是未来，这就需要端到端（收发两端）的联合设计



MAC/网络层：用神经网络（包括**大模型**）学习并理解一个未知协议，甚至**生成**一个全新的、符合当前网络环境和业务需求的**协议?**

Agenda



- 背景和驱动

- 大模型赋能语义通信
- 大模型驱动数字孪生网络
- 大模型增强O-RAN

- 总结

Semantic Communications: background
语义通信：背景

- Weaver defined three levels of communication problems: [1]
  - **Technical** problem: How accurately can the symbols of communication be transmitted ? （语法）
  - **Semantic** problem: How precisely do the transmitted symbols convey the desired meaning? （语义）
  - **Effectiveness** problem: How effectively does the received meaning affect conduct in the desired way? （意图）

Bird

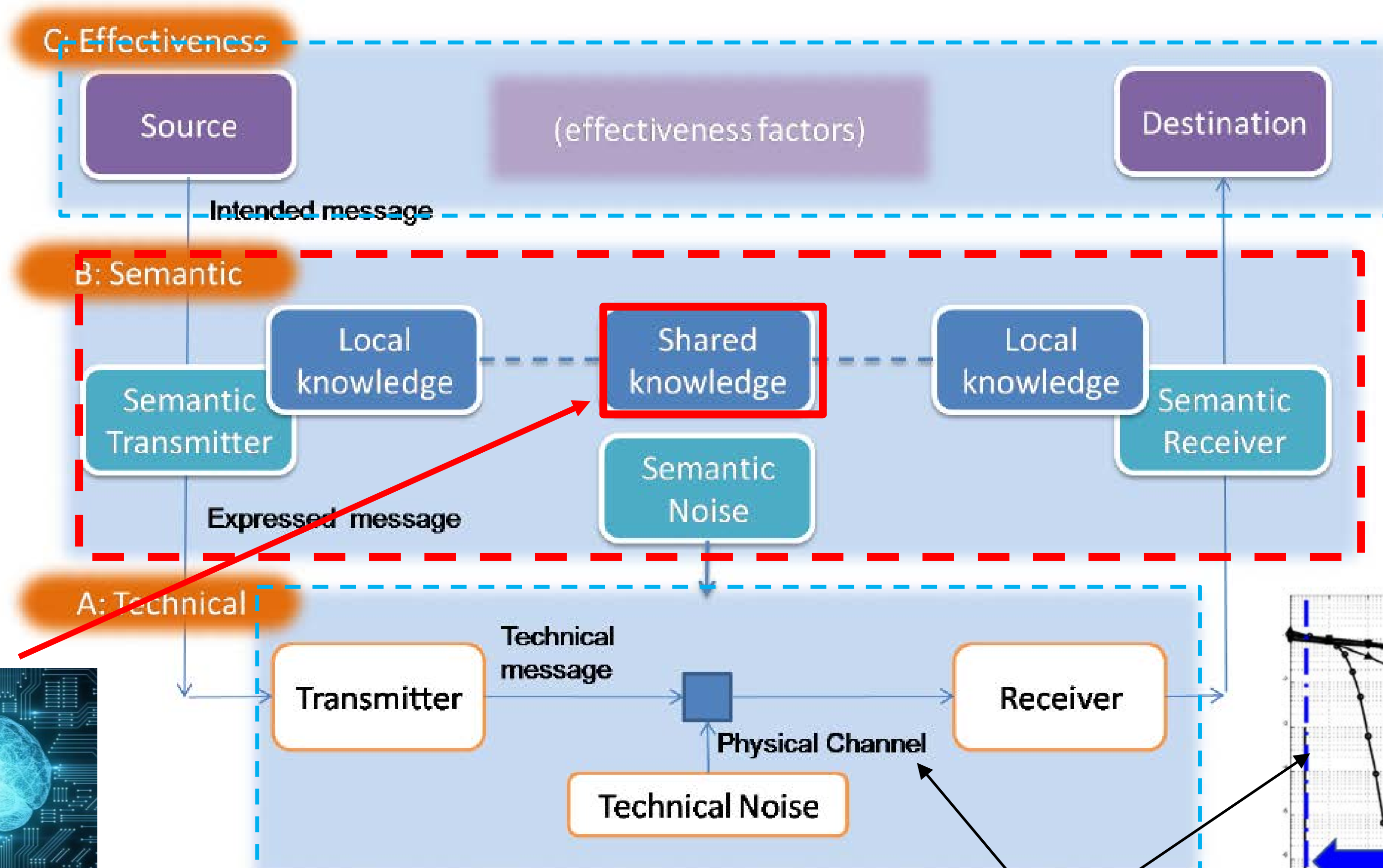
To reduce bandwidth usage and to improve comm. efficiency and to increase reliability esp. at low SNR 弱通信

NLP (natural language processing): Mom says to daughter: “do your piano when you feel like it.” -> “do your piano” or even “piano”; “do yuor piaxo whn yu ~~feel~~ like it”

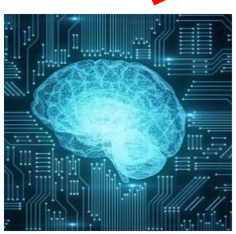
Why now? – driven and empowered by native intelligence in 6G and “all services” (e.g., metaverse)

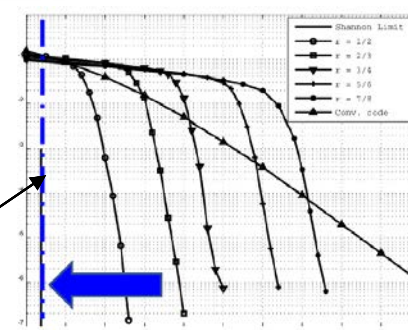
[1] Weaver. “Recent Contributions to The Mathematical Theory of Communication”, 1949

Semantic Comm vs Traditional Comm
语义通信 vs 传统通信



Semantic channel Capacity (*Turing Capacity?*)



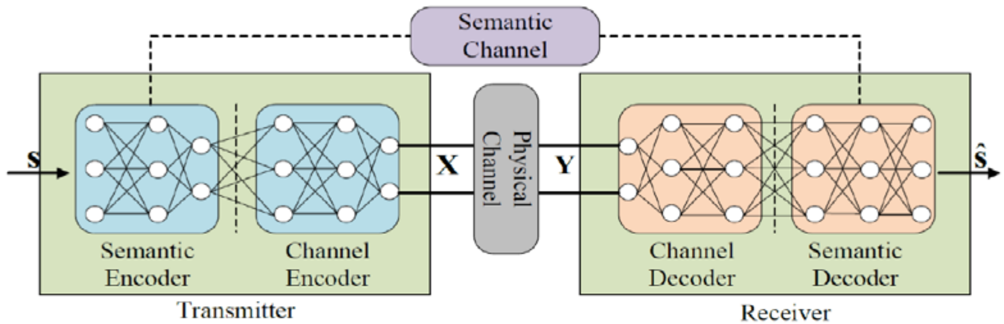


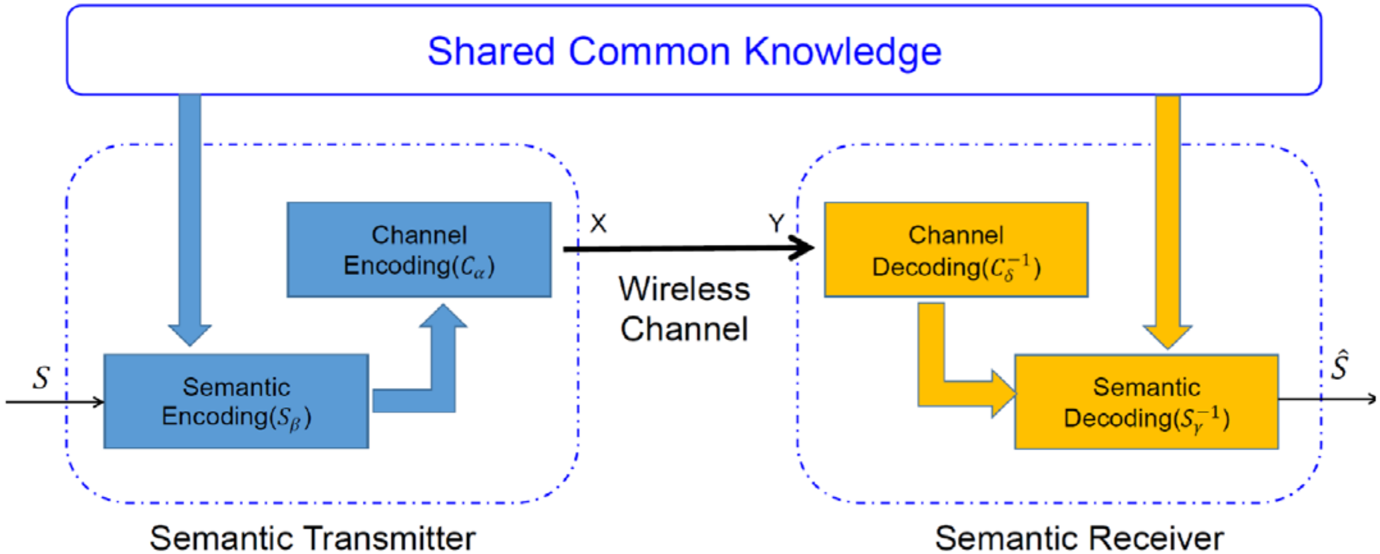
Shannon Capacity

J. Bao et al., "Towards a theory of semantic communication," 2011 IEEE Network Science Workshop, 2011, pp. 110-117

Semantic Comm: technical approach
语义通信：技术路线

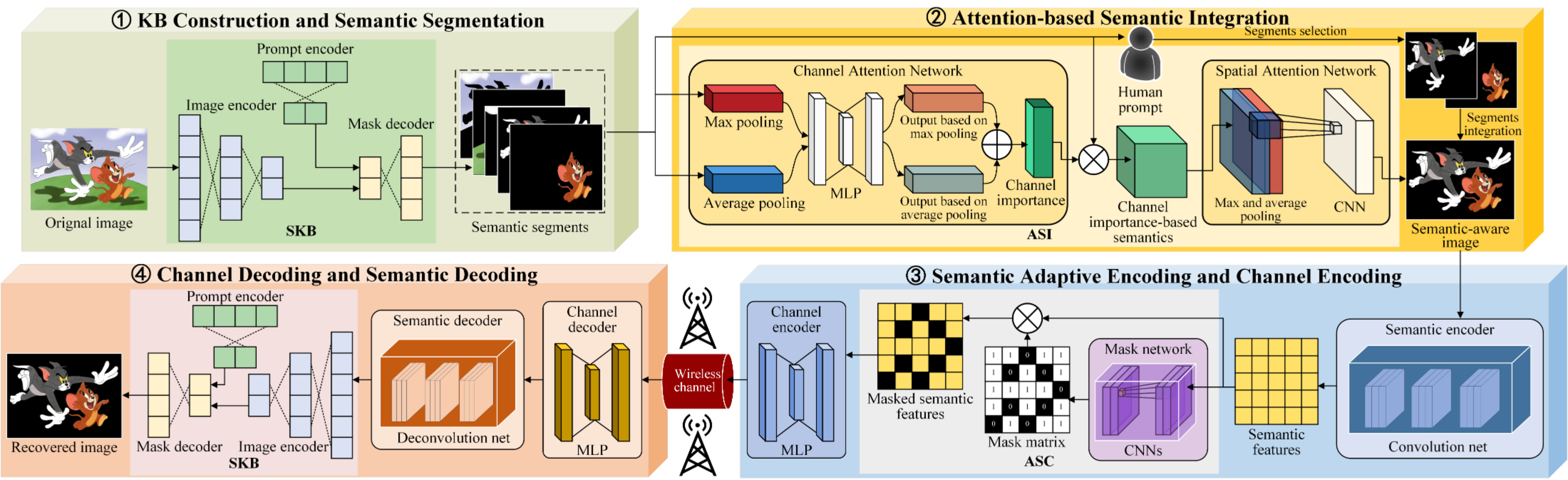
- 机器学习– ref. data-driven native AI in 6G





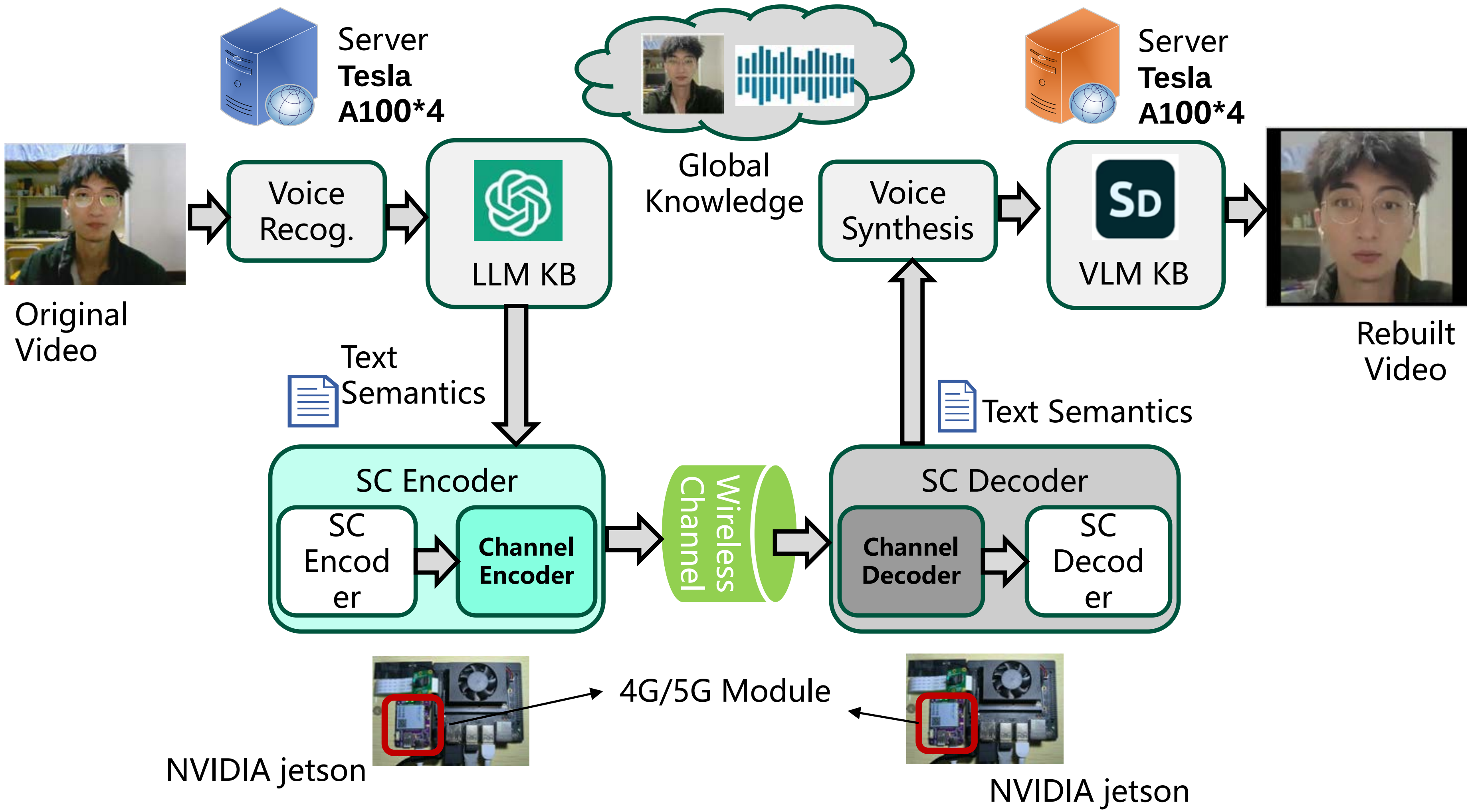
基于大模型的图像知识库赋能的语义通信

采用SAM作为语义知识库，用于根据用户的意图实现对发送的原图像进行特定目标的选取与保留，从而在只传输用户感兴趣的图像信息。



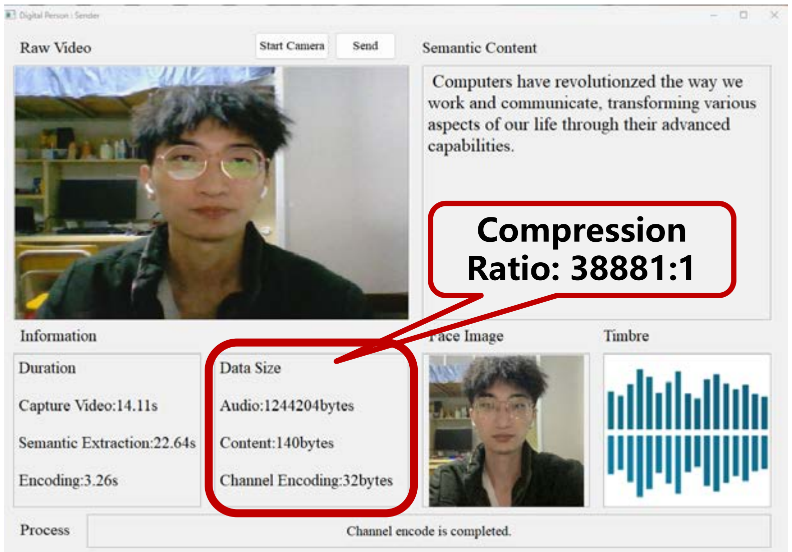
Jiang F, Peng Y, Dong L, et al. Large AI model-based semantic communications[J]. IEEE Wireless Communications, 2024, 31(3): 68-75.

基于语义的多模态端侧内容生成

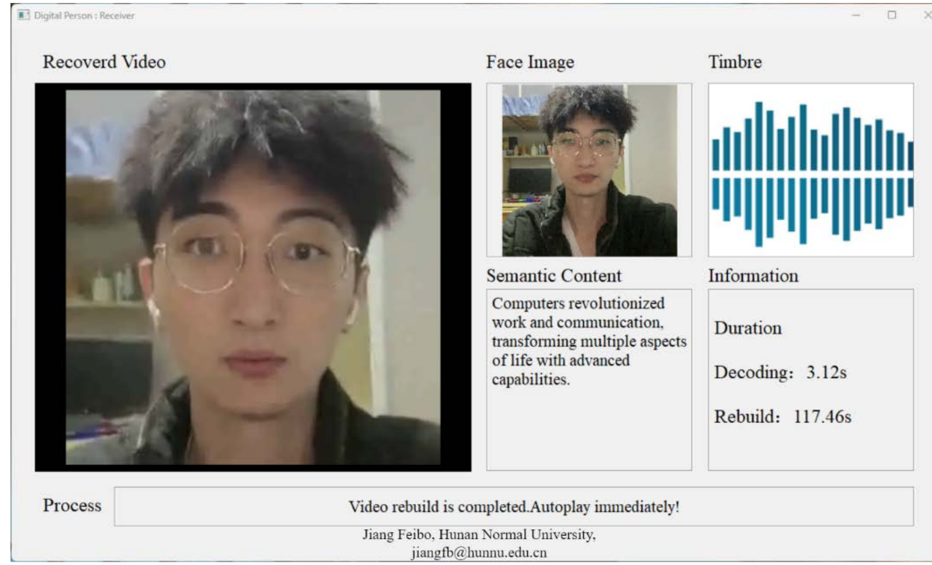


Jiang F, Dong L, Peng Y, et al. Large AI model empowered multimodal semantic communications[J]. IEEE Communications Magazine, 2024.

基于语义的多模态端侧内容生成



SC Sender



SC Receiver

Original Video

Computers have revolutionzed the way we work and communicate, transforming various aspects of our life through their advanced capabilities.

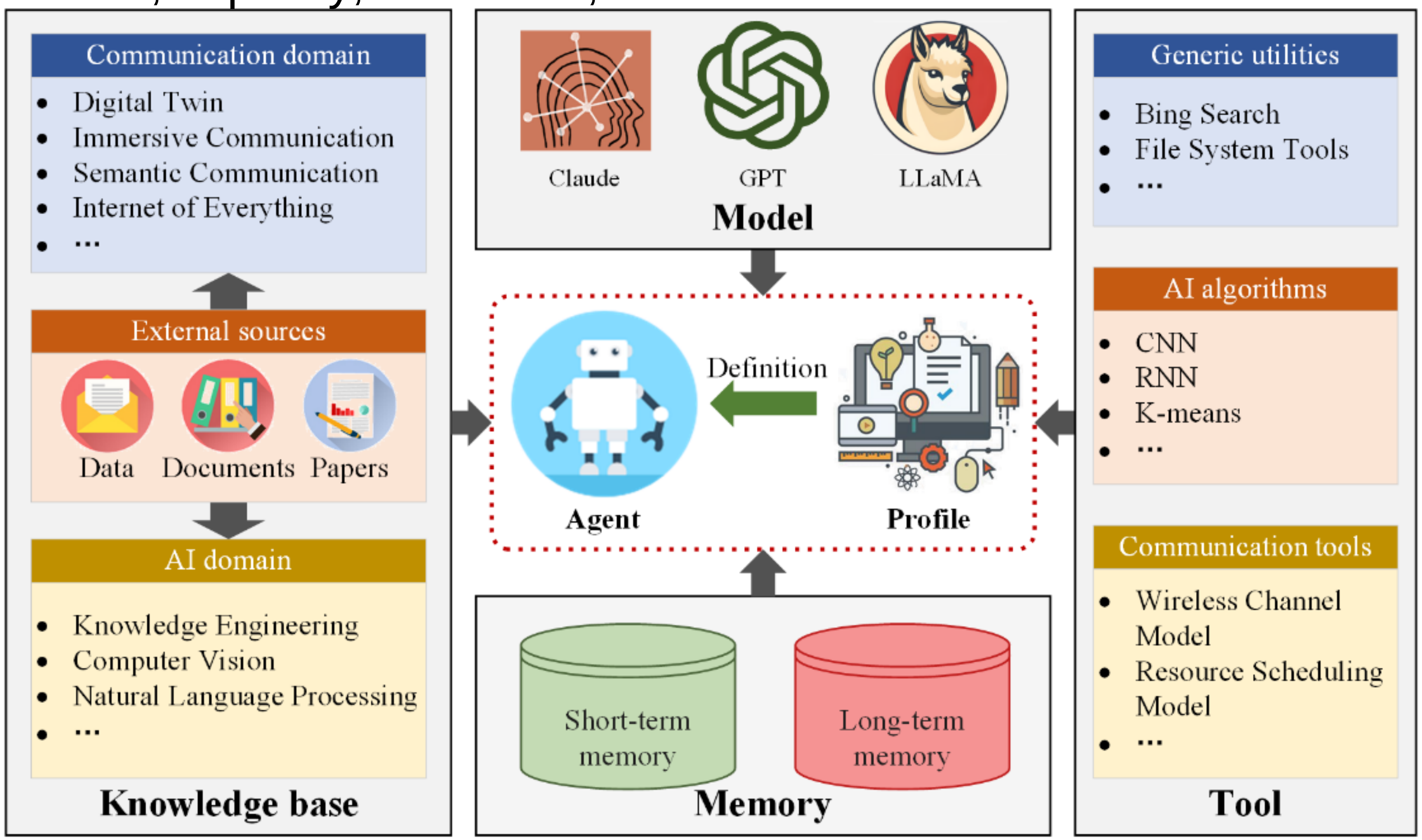
Text Transmitted

Rebuilt Video

语义通信系统的软件代码自动生成
(系统结构)

Agent: each Agent is a Chain in Langchain (思维链). It sends its input and prompts to ChatGPT to get the results

Profile includes a set of parameters and rules, describing Agent's properties such as roles, objectives, capacity, behaviors, etc.

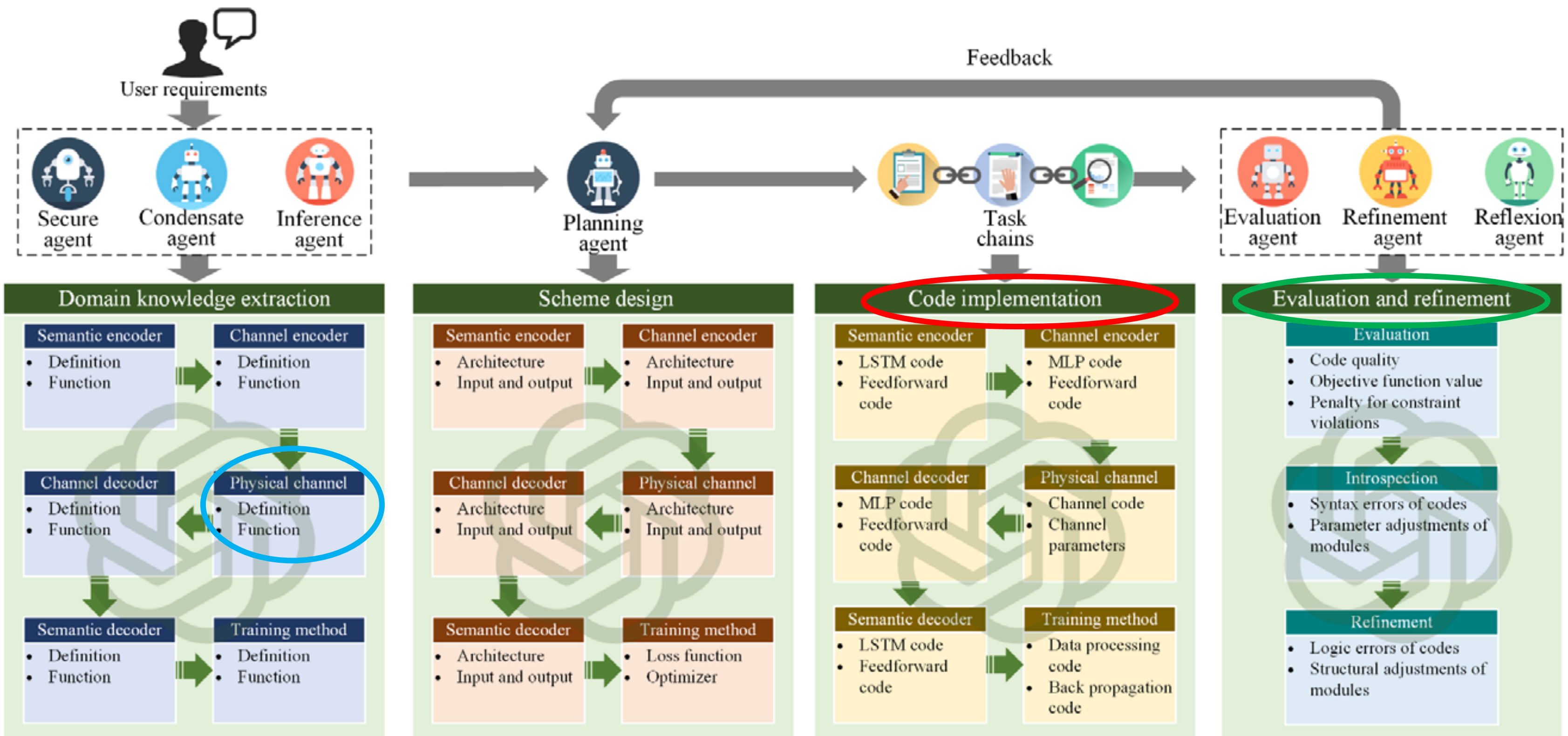


Jiang F, Peng Y, Dong L, et al. Large language model enhanced multi-agent systems for 6G communications[J]. IEEE Wireless Communications, 2024.

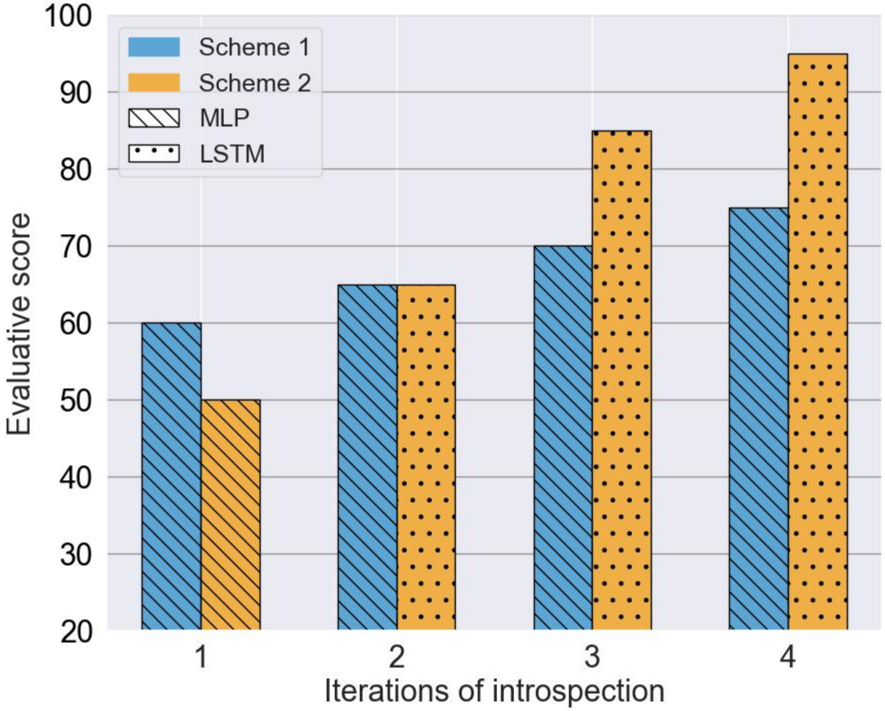
语义通信系统的软件代码自动生成
(具体实现)

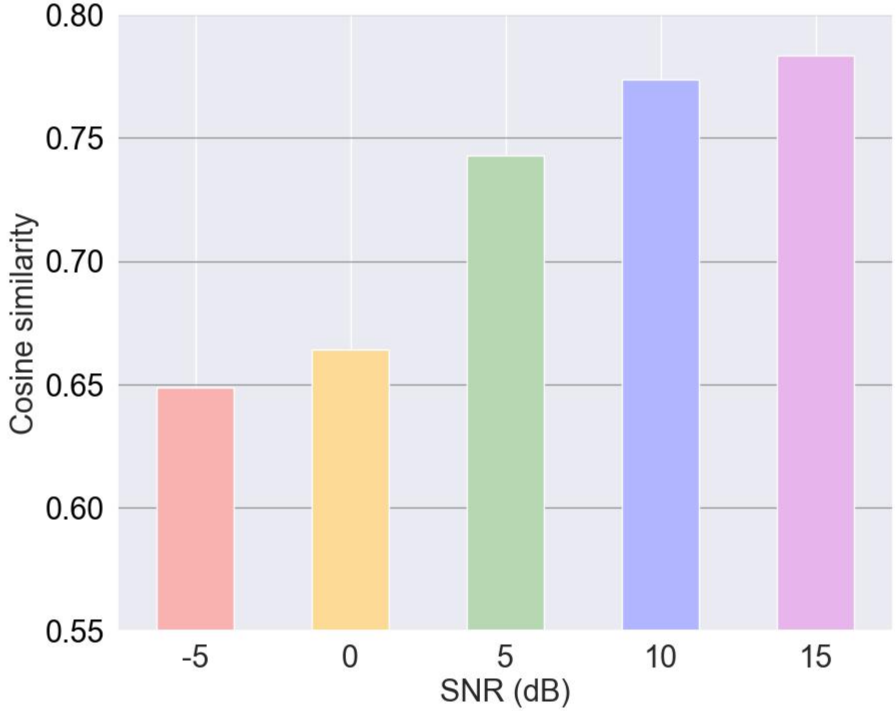
To describe requests using natural language:

- "Please produce **Python code** that implements an SC model for text transmission.
- **Additive Gaussian White Noise (AWGN)** can serve as the physical channel.
- The **bilingual evaluation understudy (BLEU)** score is adopted as the metric that evaluates the SC model.
- The generated SC achieves no less than a 0.6 BLEU score when the Signal-to-Noise Ratio (SNR) is 10dB. The total number of model parameters does not exceed 2,000,000 for resource constraints of devices."



语义通信系统的软件代码自动生成
(性能)





- Measuring the source code generated by GPT3.5
- Two sets of generated source code against iterations of introspection

- Scheme 2 is selected here as it has the best performance
- To use the generated SC to transmit text and then to compare the received text against the original text using Cosine similarity

思考

- 什么是语义通信？它和传统通信的区别是什么?
- 目前大模型与语义通信结合的主要方式是什么?
- 语义通信系统引入大模型会产生什么新的问题?

Agenda

- 背景和驱动

- 大模型赋能语义通信
- 大模型驱动数字孪生网络
- 大模型增强O-RAN

- 总结

我们认为的数字孪生是什么? What is DTN?

- 数字孪生在各行各业都是热词 Digital Twin is a buzzword everywhere!

漂亮的可视化
Visualization

虚拟现实/增强现实
VR/XR

$$DT_k(t) = \{e_k(t), \hat{e}_k(t)\}, \forall k \in \mathcal{K}$$

Difference from real entity

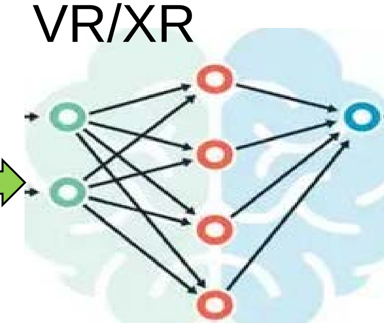
公式中的新变量

New variables

知识图谱

Knowledge Graph

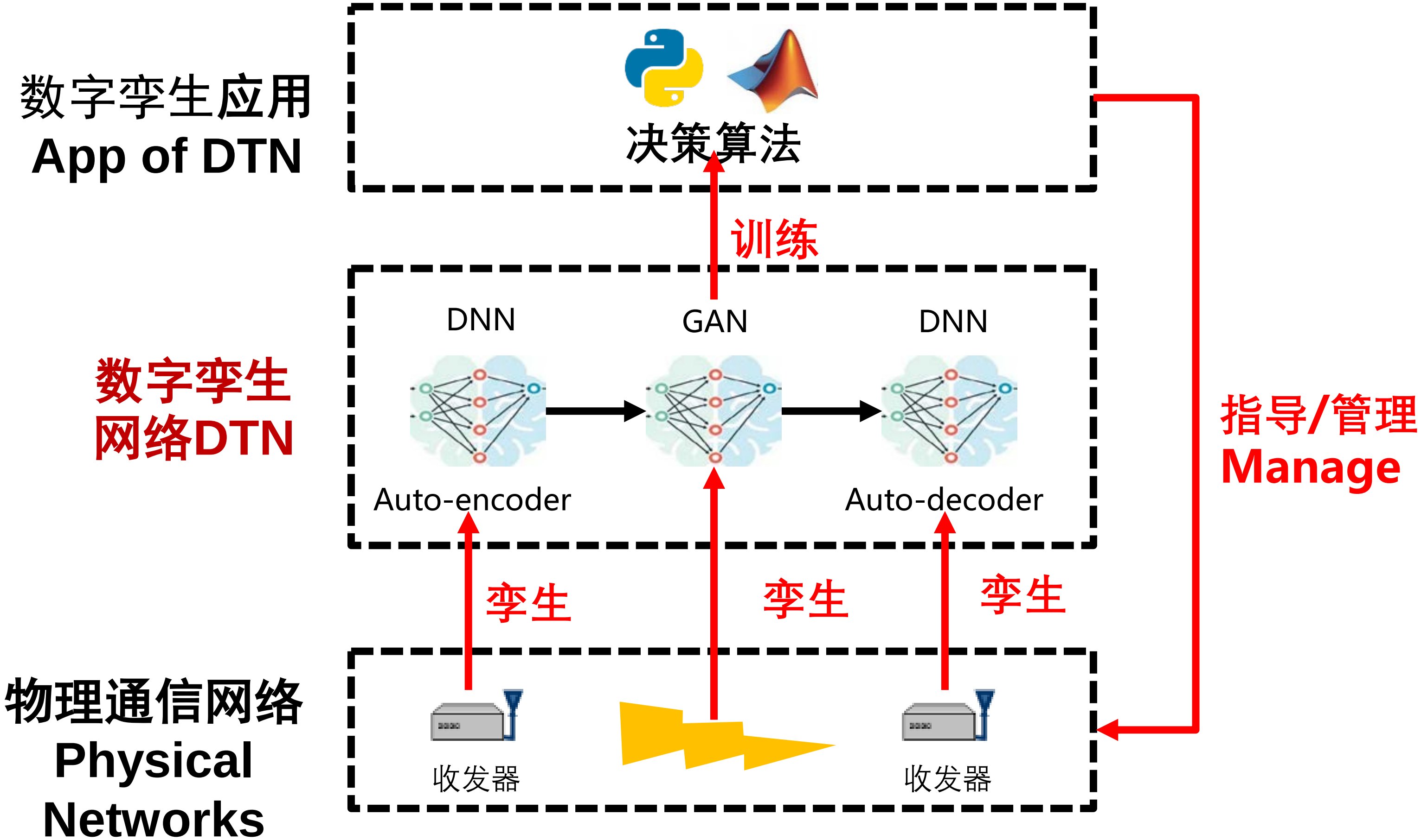
数据
data



深度神经网络

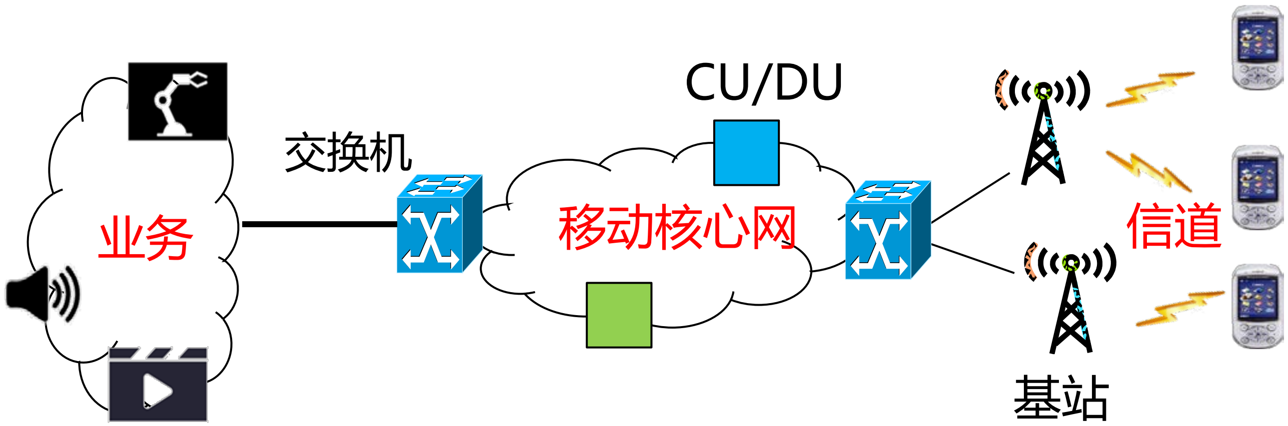
Deep Neural Network

数字孪生网络 (Digital Twin Network: DTN)



- 该DTN不是对网元的一对一映射，而是特征提取
- 通过该DTN来辅助实体网络的管理和运维，（注意：这里不是基于日志的**事后被动**分析，而是**在线主动**决策），进而降低能耗并提升通信网络效率
- **预测**未来而非仅仅基于**当下** **future** (predictive) vs **now**

总体技术思路

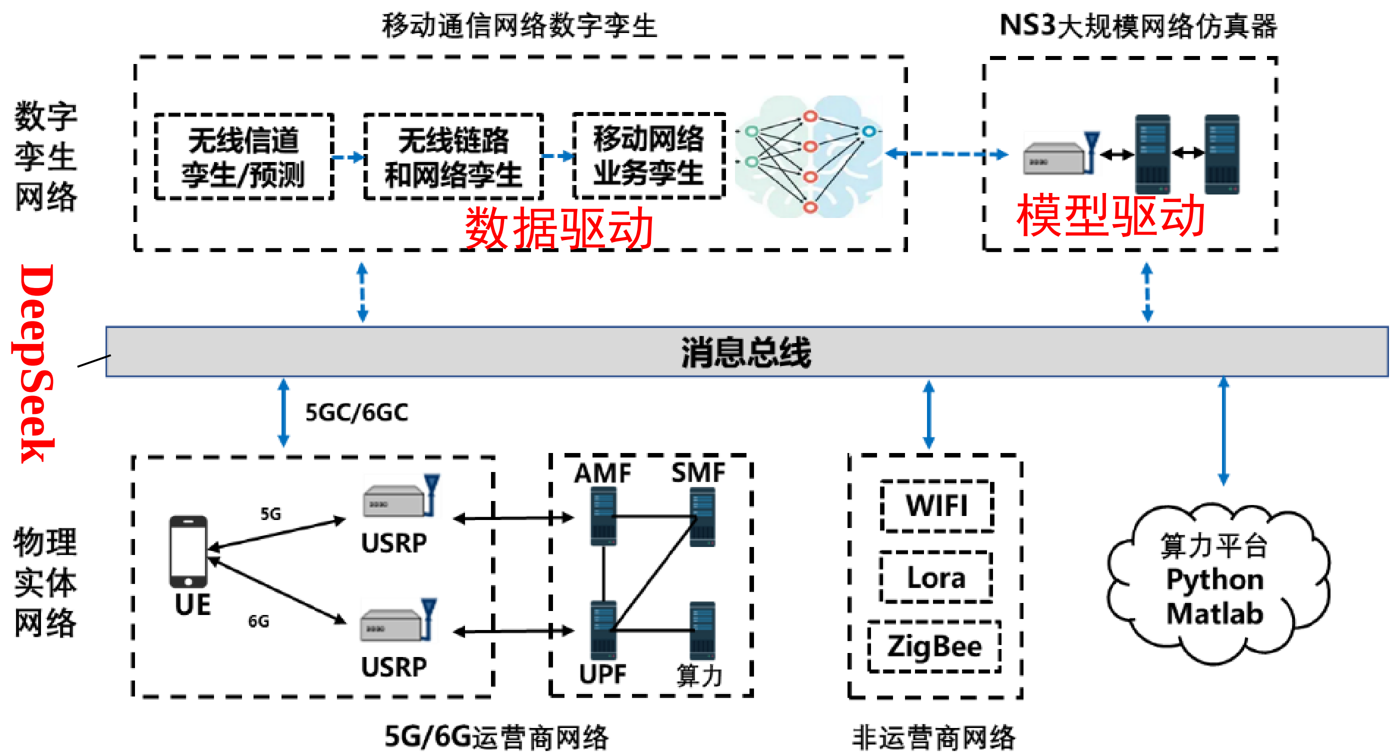


- Full-chain: channel, link/network, and services 全链条：包括信道、链路/网络和业务
- Self-evolving: 自演进: DTN不可能包罗万象，需要动态增加新功能，性能自我提升
- Elastic 可伸缩：DTN软件上可拆减组装，适应不同需求

可伸缩数字孪生网络系统设计和实现

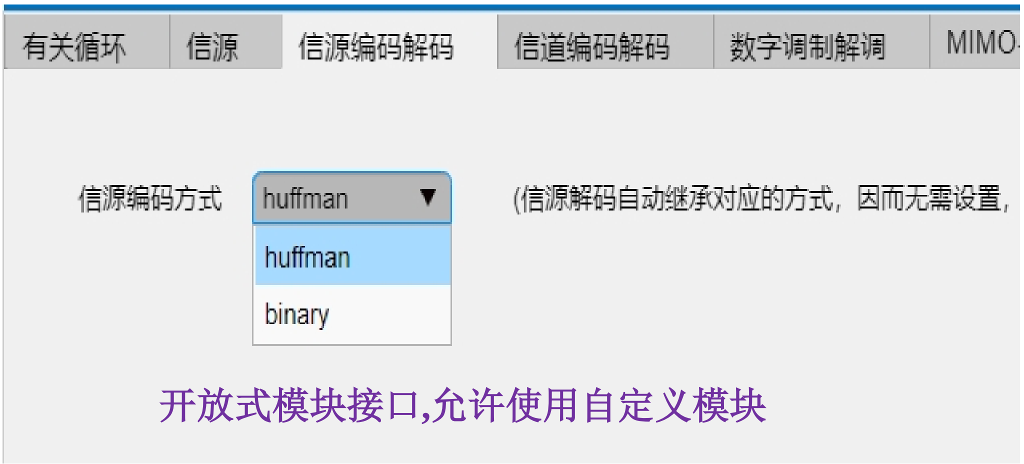
- 思路：通过软件系统实现数字孪生网络，达到可用
- 意义：对内，在架构上增加新功能和新场景；对外，自由选择搭配生成不同大小/复杂度的DTN以适应个性化需求（如通信环境和算力）- 可伸缩
- 技术挑战：系统架构、协议和具体实现机制设计

- 技术核心：
  - 可插拔的总线结构，消息格式定义；
  - 数据-模型双驱动：既有基于数据驱动的（神经网络）孪生，又有基于模型驱动的（仿真，**物理定理/数学公式**）孪生
  - **数据集+神经网络 (Python code)**

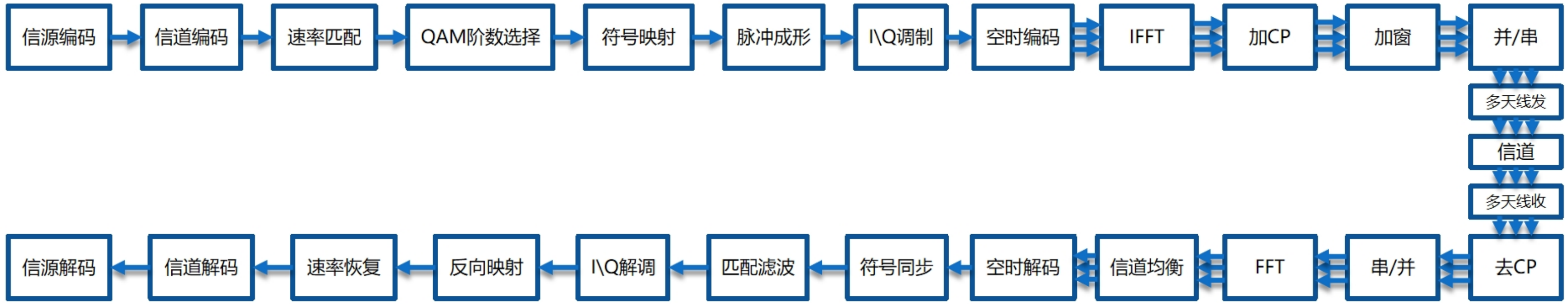


使用Matlab仿真通信链路

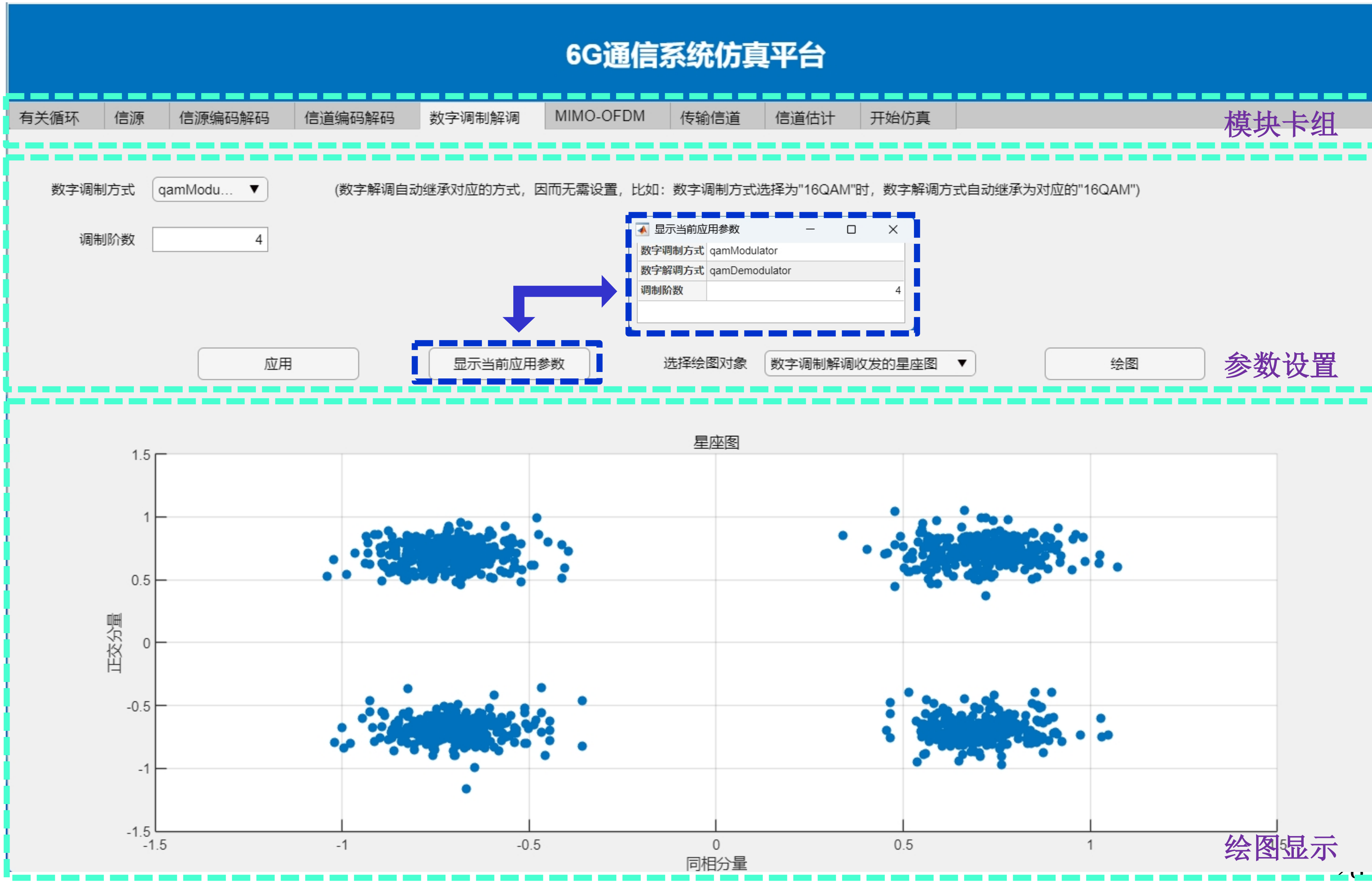
- ❑ **基本特点**
  - ❑ 开放式模型接口和GUI演示模块
  - ❑ 基于模型驱动的通信物理层仿真
- ❑ **平台使用**
  - ❑ 在GUI中依次设置各模块参数,点击开始仿真后即可绘制所需曲线和图像



- ❑ **通信物理层仿真模块**



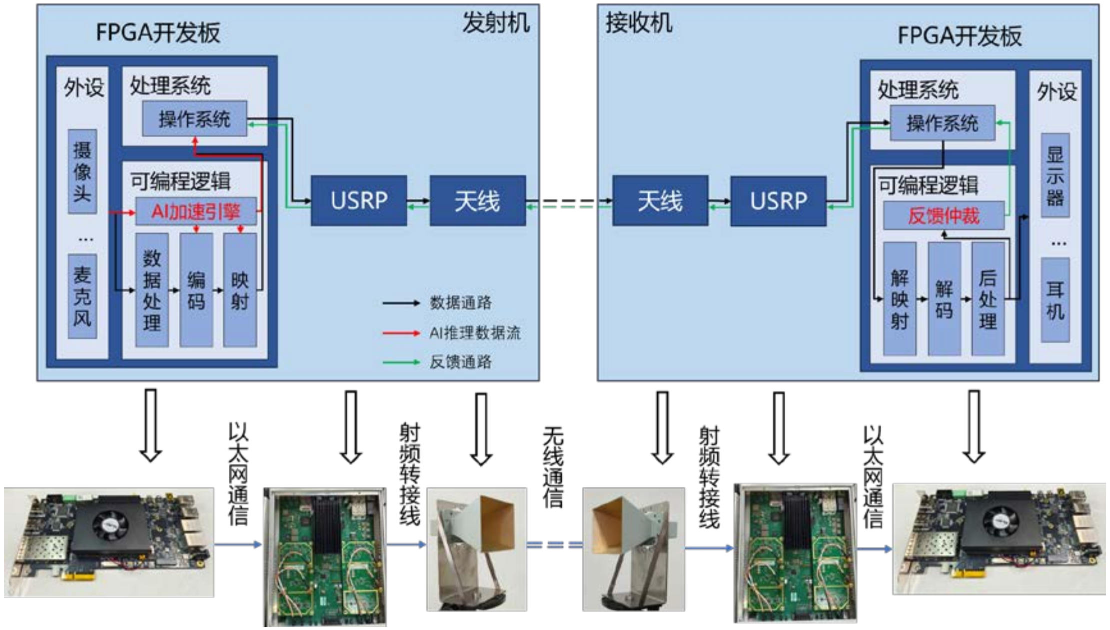
使用Matlab仿真通信链路

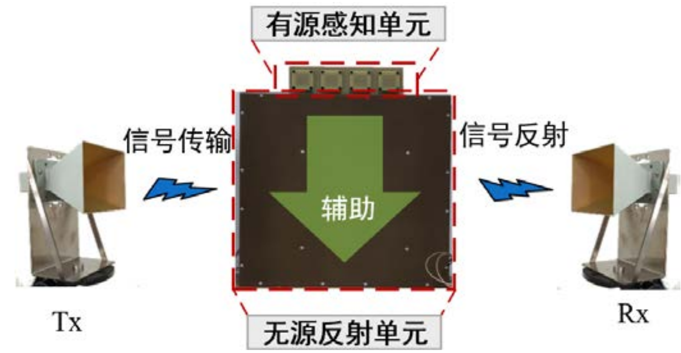


基于NS-3仿真器的模型驱动DTN

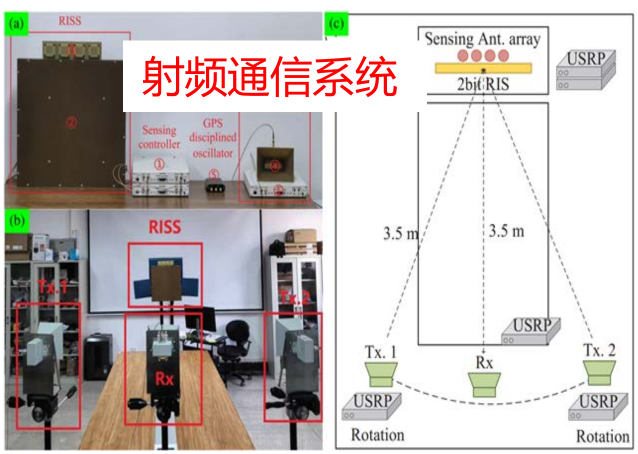
- 可仿真2000+节点
- 异构网络（WiFi/ZigBee/LoRaWAN/4G/5G)
- 可对实物节点控制
- 自组织网络路由、物联网能耗算法等

从真实物理系统采集数据





带感知单元的
IRS智能反射面



异构网关

思考

- 什么是数字孪生?
- 大模型在数字孪生扮演什么样的角色?
- 你认为在未来6G中，数字孪生能够实现吗？为什么?

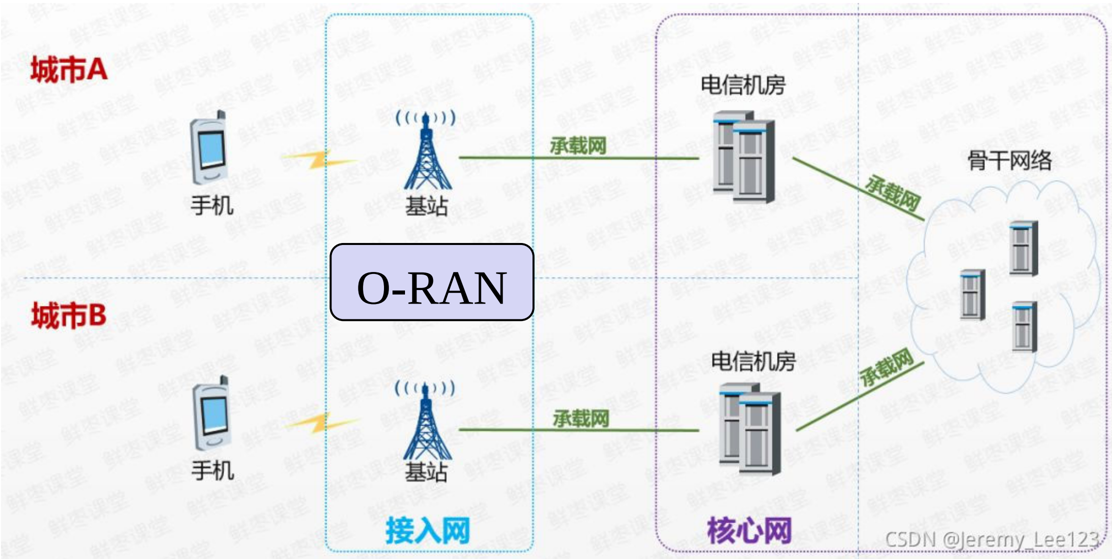
Agenda

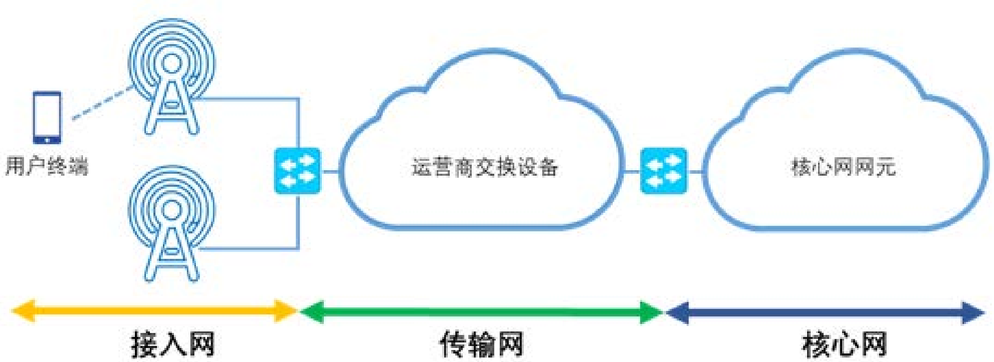
- 背景和驱动

- 大模型赋能语义通信
- 大模型驱动数字孪生网络
- 大模型增强O-RAN

- 总结

传统无线接入网——RAN

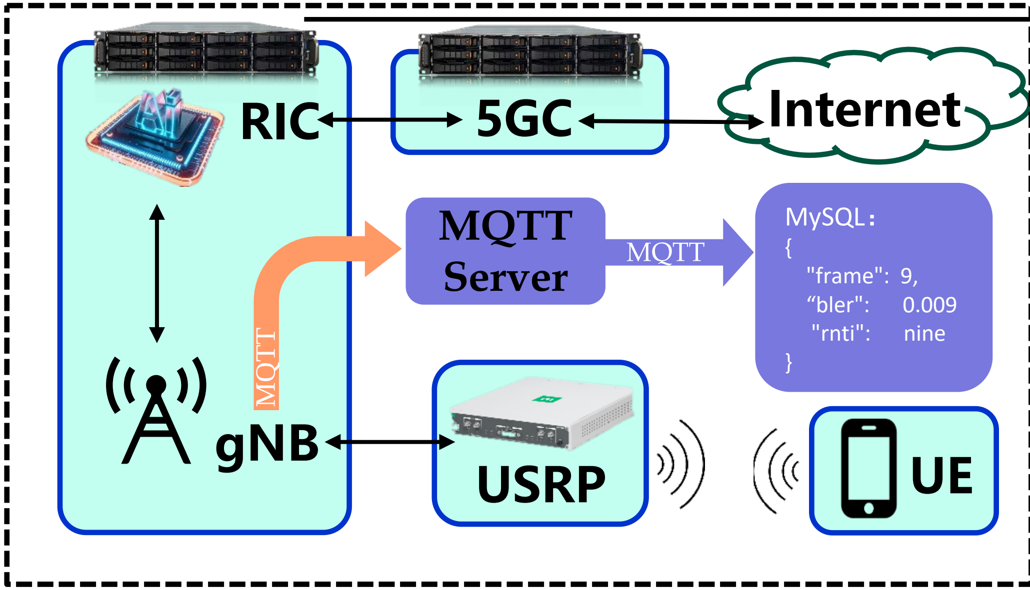




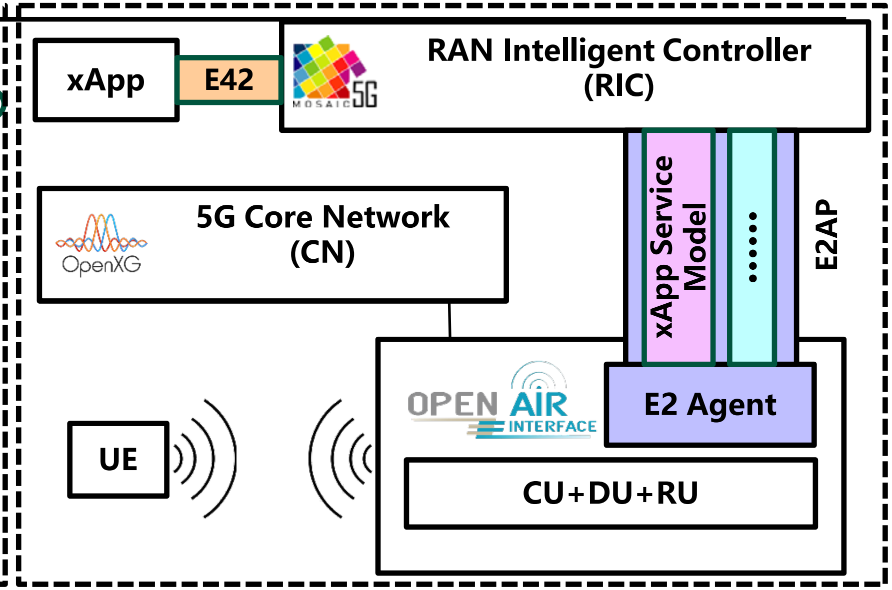
开放无线接入网（O-RAN，Open Radio Access Network）是一种基于开放性、智能化和虚拟化理念的新型无线接入网架构。其核心目标是打破传统RAN设备"端到端"封闭的部署模式，通过开放接口、模块化设计和多厂商互通，提升网络的灵活性和可扩展性。O-RAN架构中引入了虚拟化的基带单元（vDU、vCU）、分布式与集中式处理模式，以及由AI驱动的智能控制器（RIC，RAN Intelligent Controller），从而支持网络的按需优化和智能资源调度。

基于O-RAN架构的5G系统搭建

❑ 系统物理架构 Physical Architecture



❑ 系统软件架构 Software Architecture



❑ 基于O-RAN架构的5G系统 5G Mobile System Enabled by O-RAN

5G Base Station (OAI 5G):

- CU+DU+RU
- 多个UE接入
  Multiple UE access
- 通过USRP无线信道
  USRP wireless channel

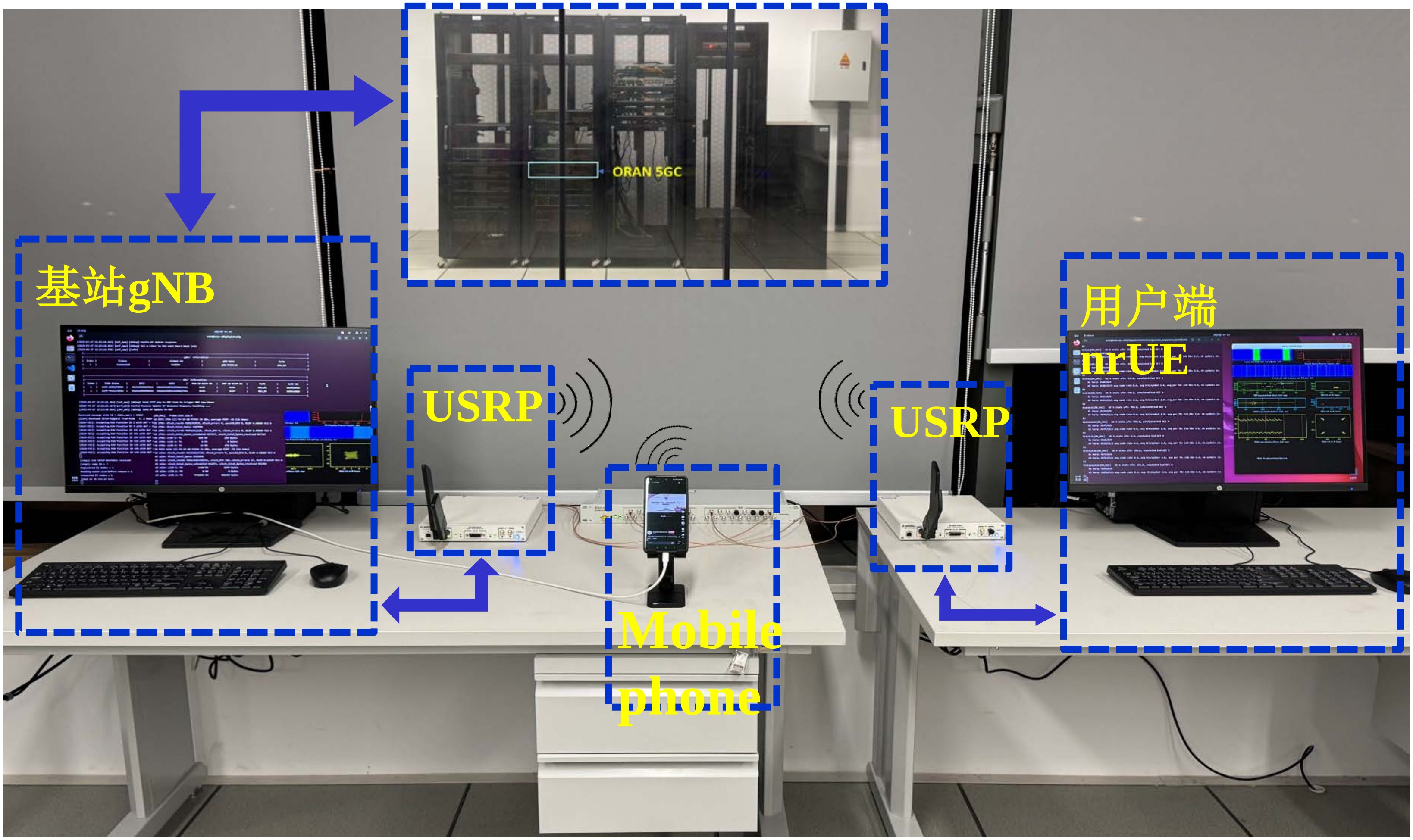
FlexRIC

RAN Intelligent Controller (FlexRIC):

- 通过E2接口与5G系统连接
  Connects to 5G systems via the E2 interface
- 支持自定义xApp开发， Supports custom xApp development
- 基站数据采集， Collects base station data
- 下发控制指令， Issues control commands

 34

ORAN 5G 实物系统 Prototype

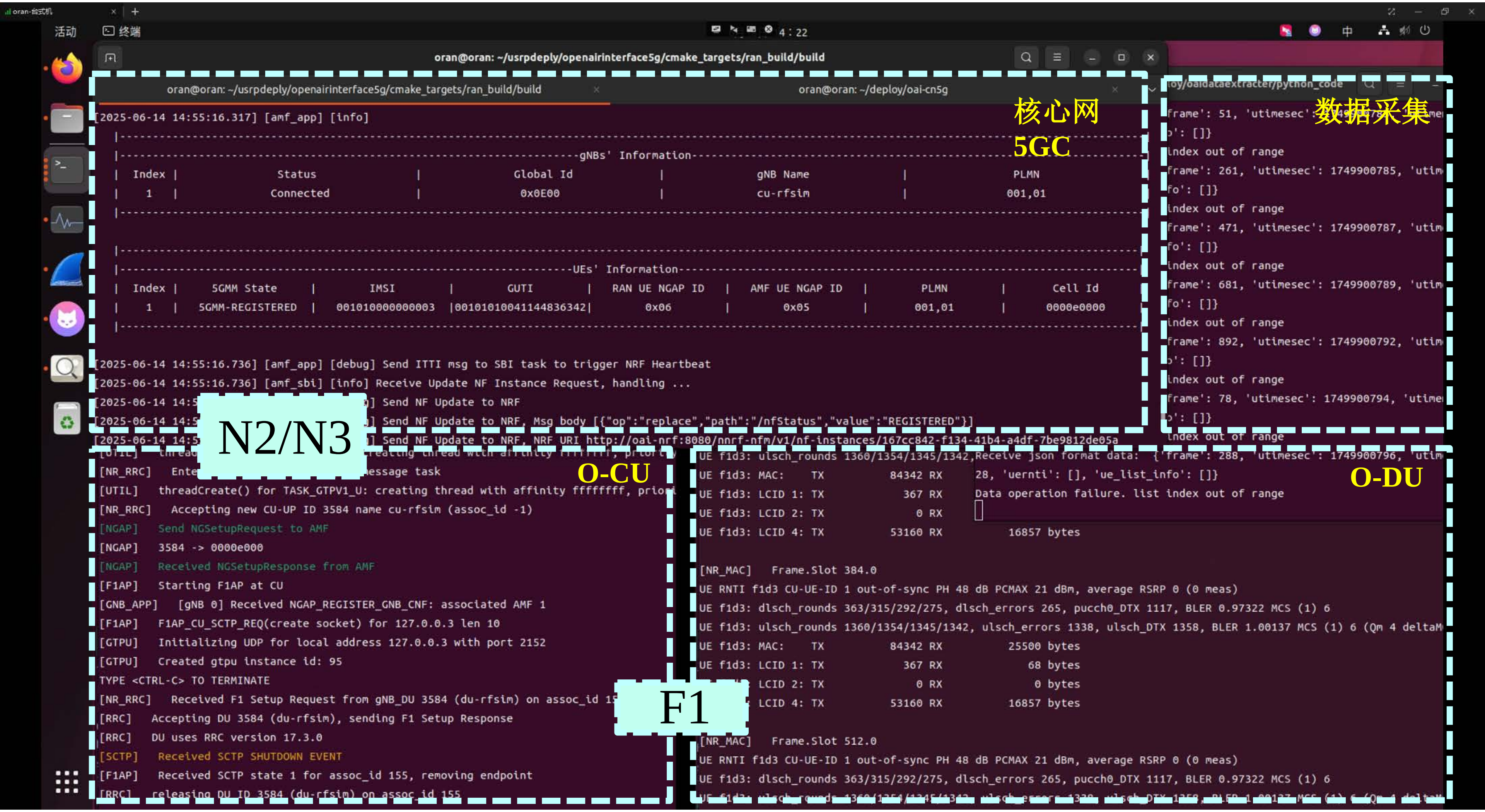


基于xAPP的MCS(调制编码)控制系统

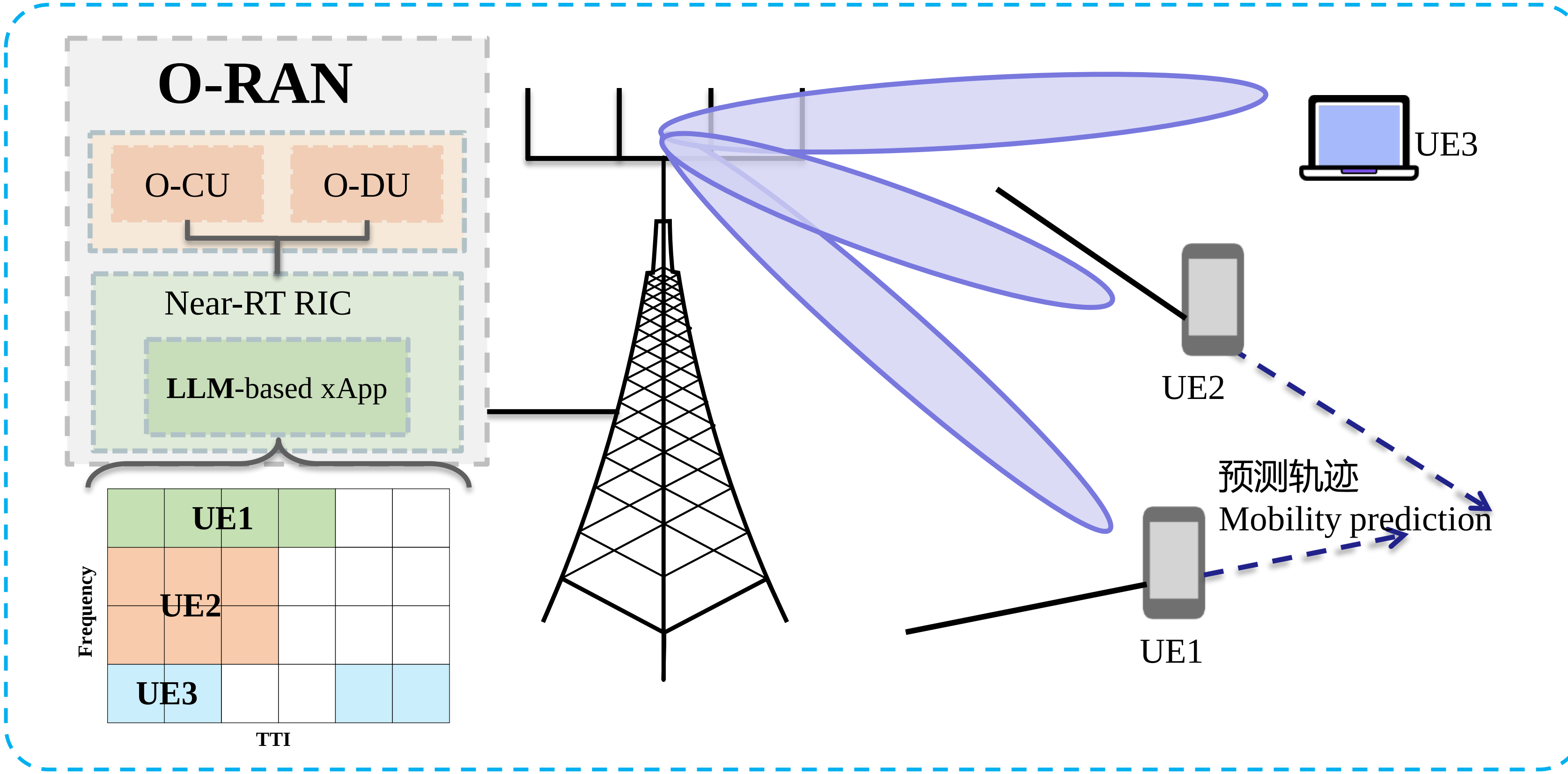
基于O-RAN架构的CU、DU分离部署5G系统搭建

Functional Split Options 2

- CU+DU+RU

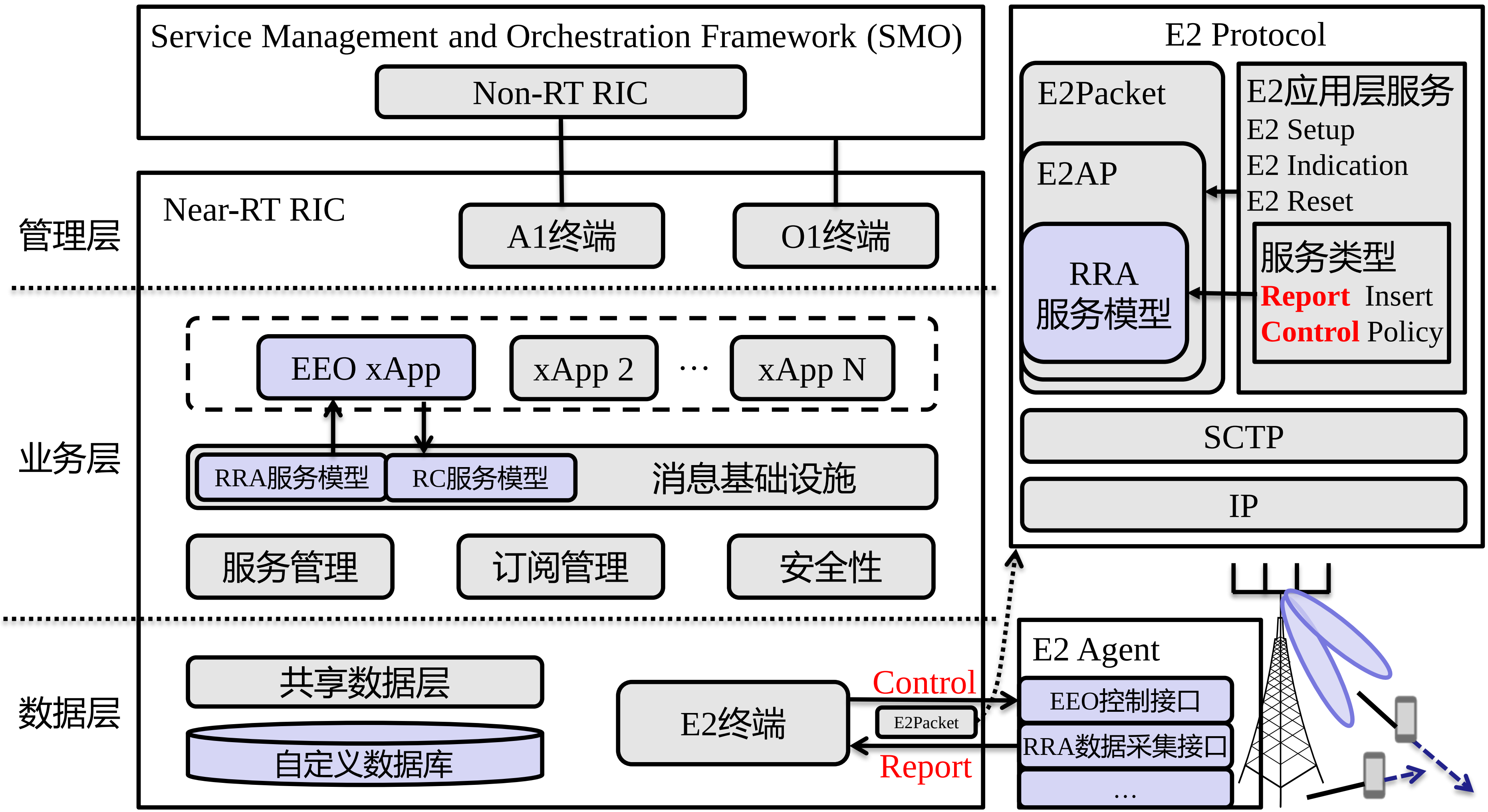


O-RAN 架构下AI辅助资源分配 – LLM-enabled Dynamic Resource Allocation Aided by xAPP

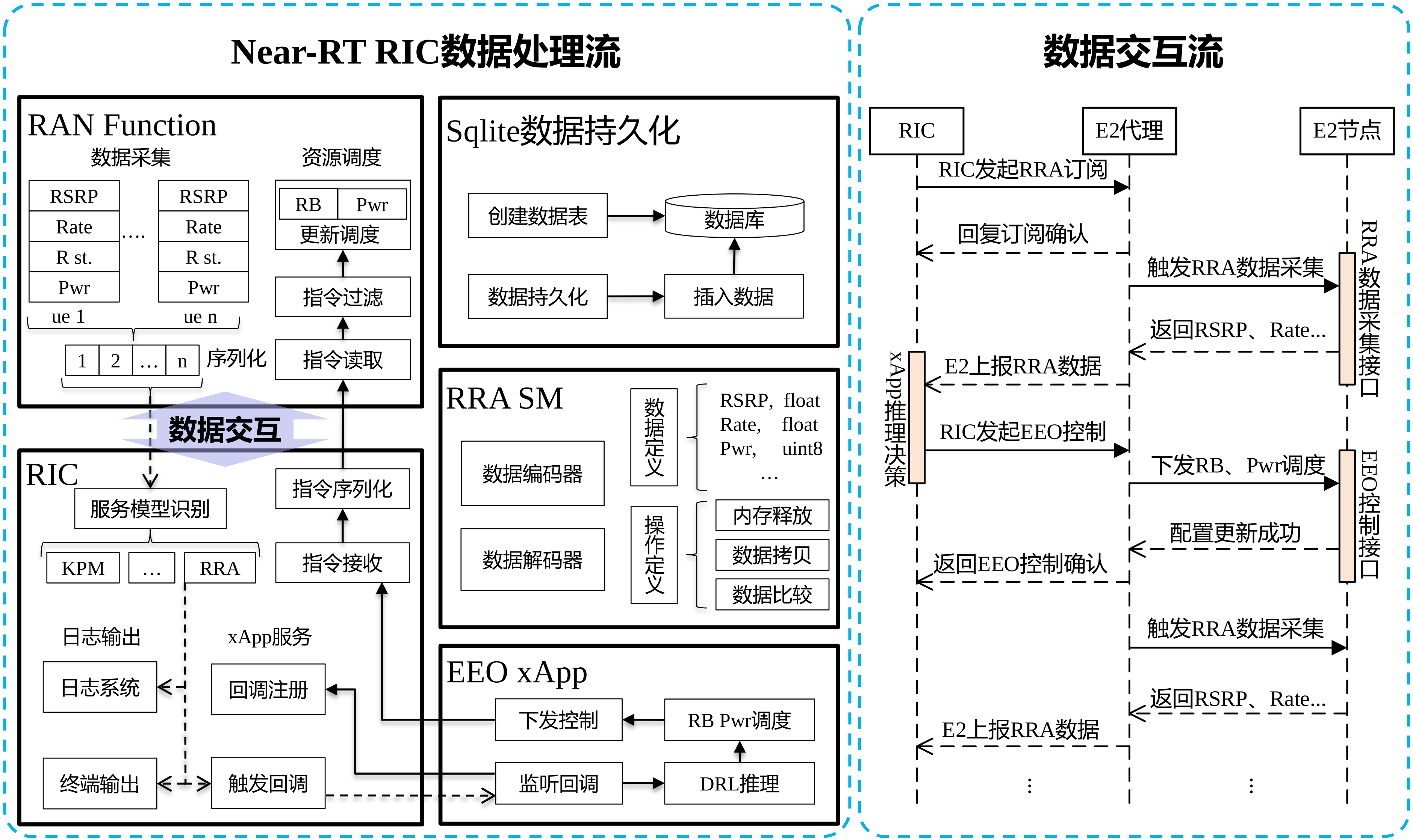


- 用户移动模型建模，卡尔曼滤波对位置信息去噪，**提升目标感知精度**
- LLM预测用户下一时刻位置信息，**避免**因实时信息决策造成**策略滞后**
- 深度强化学习综合调度资源块分配和功率分配，**提升系统能效**

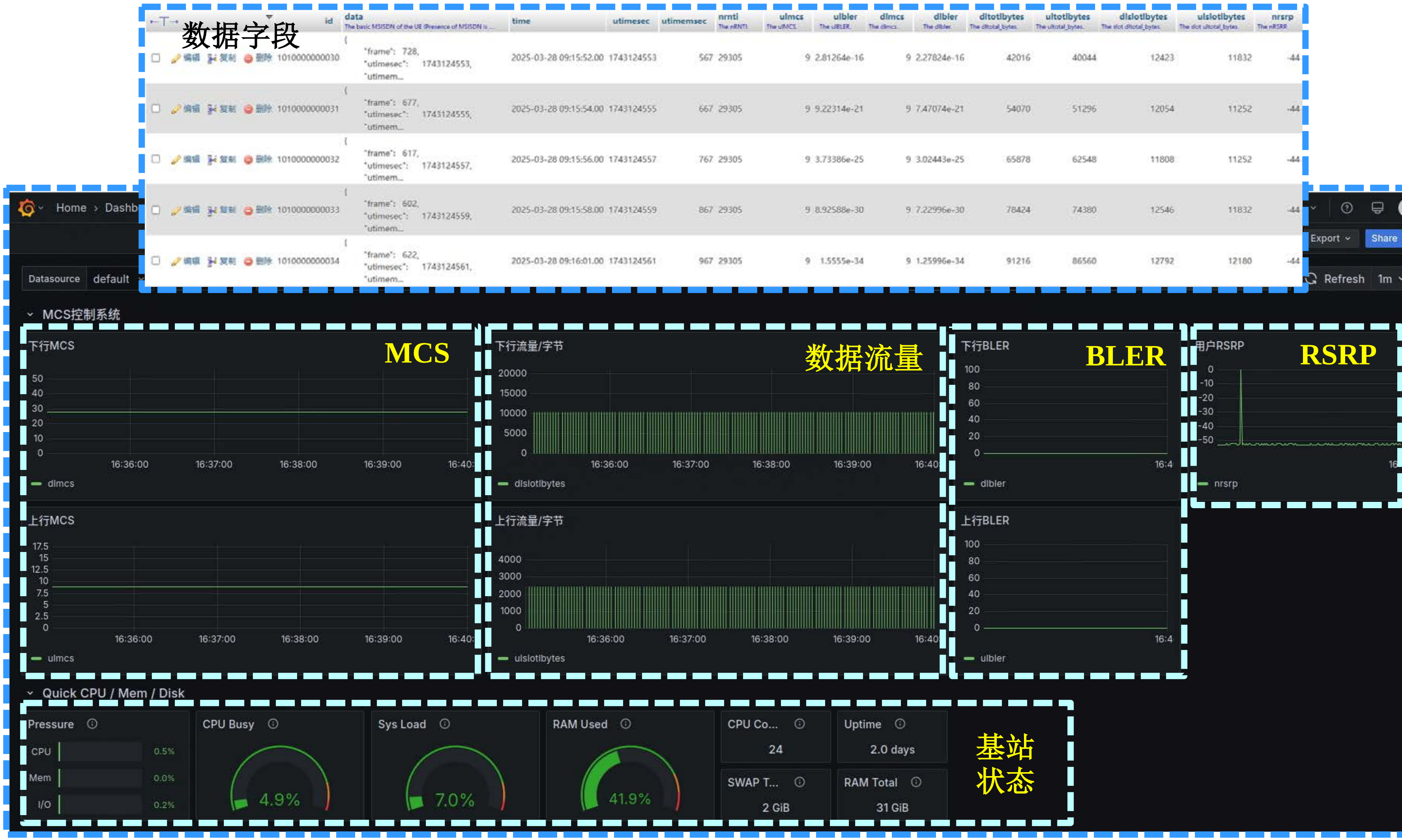
O-RAN xApp赋能能效优化



O-RAN 数据流 Dataflow



基于O-RAN架构的5G数据收集系统



思考

- RAN、O-RAN、AI-RAN的区别?
- 大模型是如何引入到RAN架构中的?

Summary

- Native AI in 6G is promising to support full spectrum, all scenarios and all services （6G内生智能势在必行）
- Generative AI is a key enabling tool for native AI in 6G. （生成式AI是重要的使能工具+传统模型）

https://ninelab.org.cn/

微信公众号：**南苏智慧网络实验室**

Thanks for your attention!