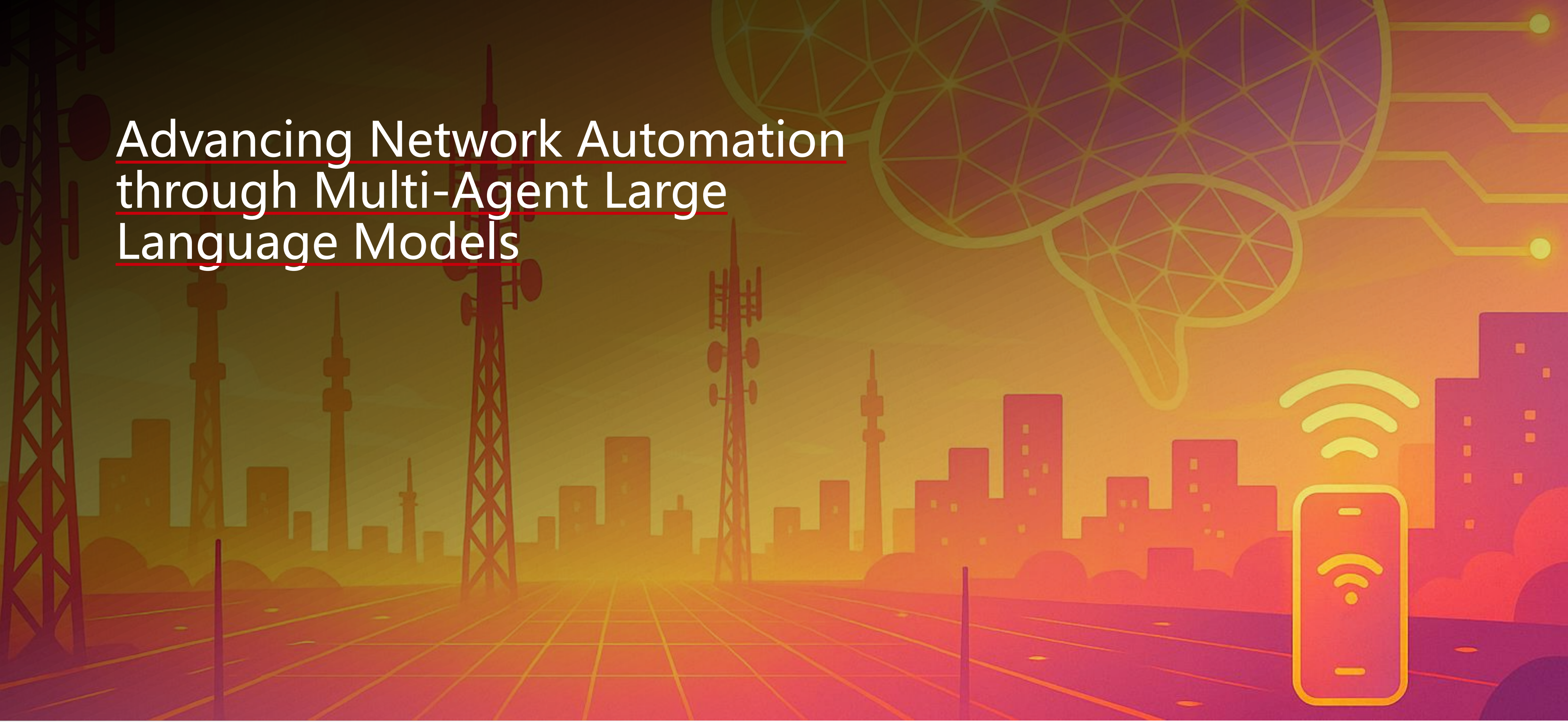
Advancing Network Automation through Multi-Agent Large Language Models



Dr. Nicola Piovesan
Senior Researcher

AI and Machine Learning in communication networks
AI4Good Global Summit 2025
July 10, 2025

HUAWEI

Introduction

- **Network optimization** is essential to ensure high performance in large-scale cellular networks, addressing two primary objectives:
  - **Enhancing UE performance**
    - Optimizing metrics like data rates, latency, coverage, and handover efficiency
  - **Minimizing network energy consumption**
    - Optimizing load distribution and activation of shutdown mechanisms
- Modern networks consist of **ten of thousands of BSs**, each operating multiple cells
  - Each cell has **several hundreds** of tunable **parameters**

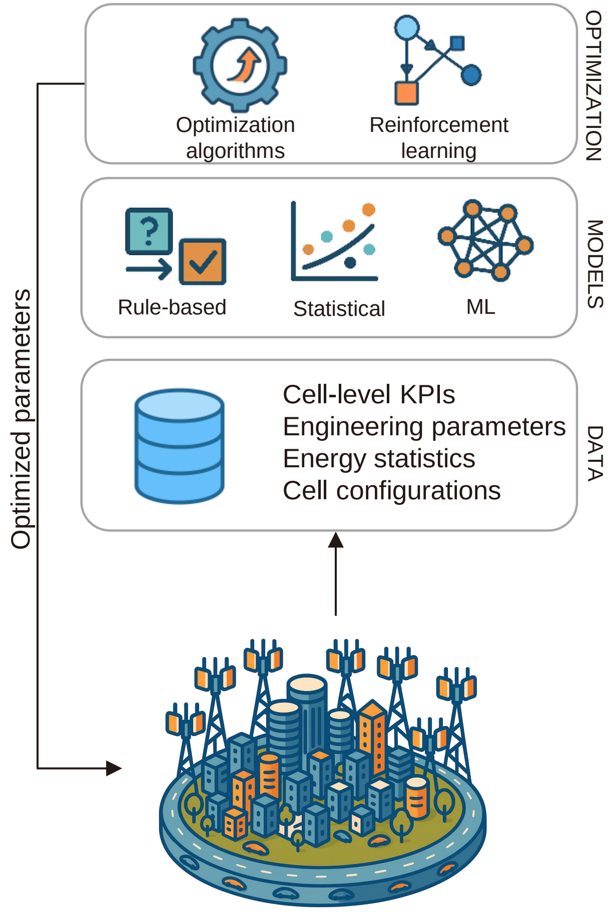
Data-driven modelling and optimization

- Accurate **modelling of the network characteristics and behaviors** is fundamental to **optimize the network performance** and energy efficiency

- **Data**
  - Data from the networks is analyzed and leveraged to create accurate models

- **Models**
  - Different network characteristics and behaviors are modelled using rule-based, mathematical or machine learning models

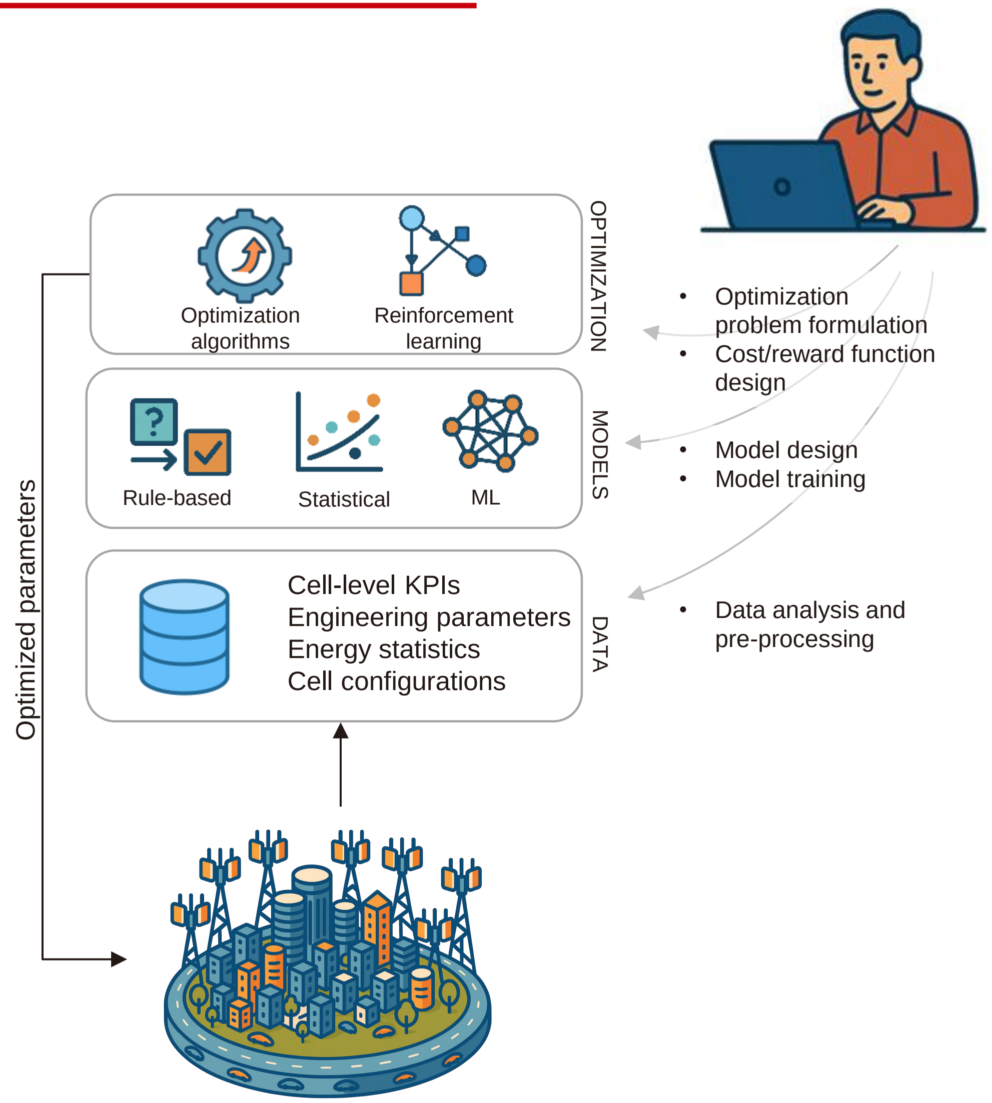
- **Optimization**
  - Optimizers or reinforcement learning agents leverage the models to optimize network parameters



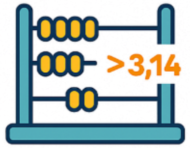
HUAWEI

Data-driven modelling of network behavior

- **Challenges:**
  - **High dependency on human intelligence (HI)** for model design, adaptation, and problem formulation
  - Different networks expose different data features; models often need **adaptation** or **re-engineering**
  - Deployment of new hardware or software introduces **new features and mechanisms** requiring continuous model updates
  - **Emerging problems** necessitate new formulations and solver selection



Limitations of LLMs



Computation limits

LLMs often fail at precise data manipulation and numerical reasoning

Domain Knowledge

LLMs, especially smaller models, lack deep expertise in telecom standards

Hallucinations

LLMs may produce plausible but incorrect outputs, with limited ability to self-detect errors

Concept-to-Execution

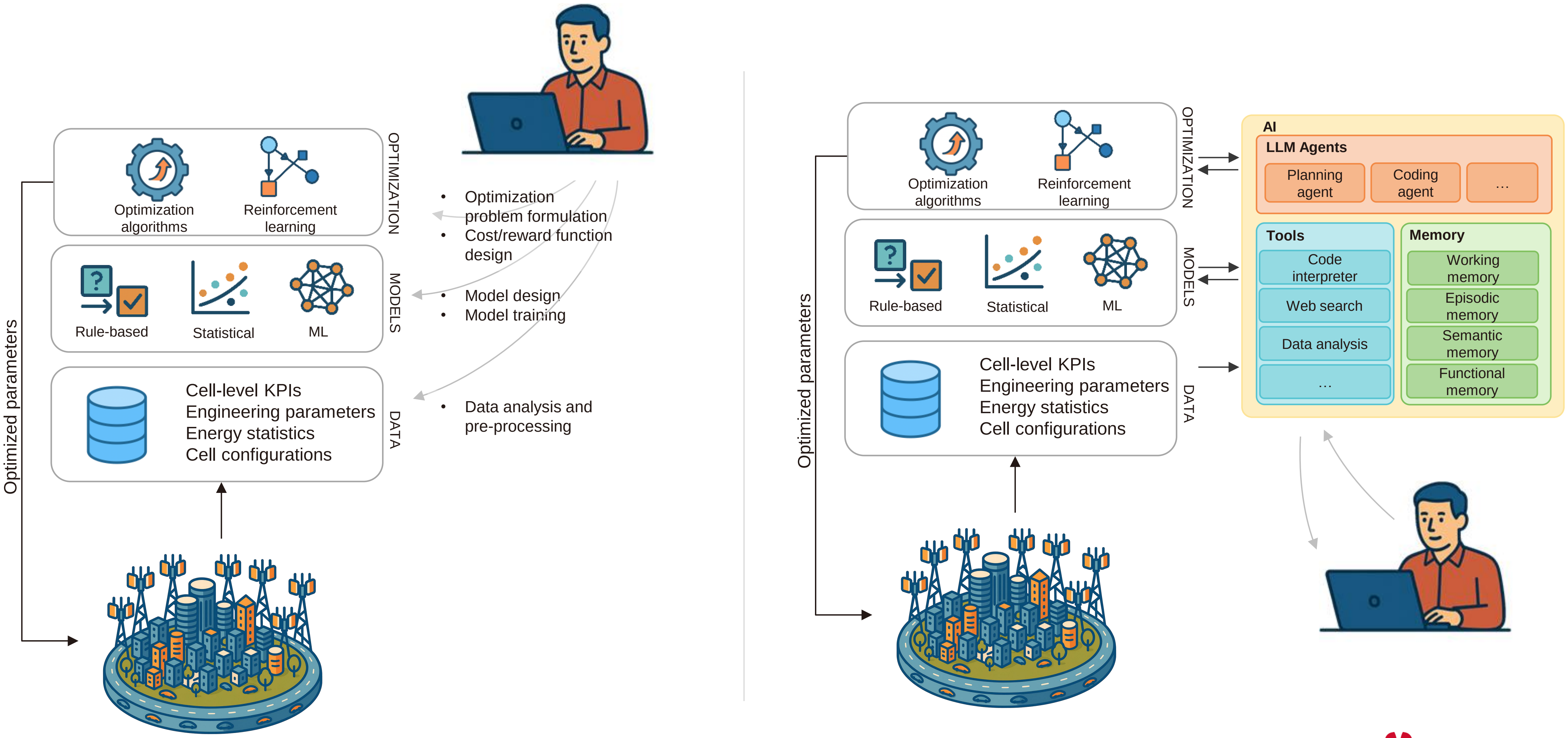
Even when understanding concepts, LLMs frequently mishandle practical application (e.g. unit errors in SINR calculations)

Planning

LLMs struggle with complex, multi-step planning, often losing track of context or dependencies in abstract, layered tasks

HUAWEI

From HI- to AI-based network optimization



HUAWEI

Multi-agent framework and tools example

Task:
What is the average downlink throughput of all active cells?

PLAN

Here is the plan of action that I will follow to solve the task:
Step 1: Load the data from the file './network-data/cell_kpi.csv' into a pandas DataFrame to understand the structure and identify the column that contains the downlink throughput and cell status.
Step 2: Filter the dataset to retain only rows where the cell status is active.
Step 3: Calculate the mean of the downlink throughput using the pandas DataFrame's mean() method.
Step 4: Round the mean fare value to two decimal places.
Step 5: Provide the formatted mean fare value as the final answer.

THINK

I need to load the data from the given file and inspect it to determine the column name for the downlink throughput and cell status.

TOOL

```
import pandas as pd

# Load data
df = pd.read_csv("./network-data/cell_kpi.csv“)
print(df.head())
```

THINK

The dataframe loaded as expected. The column which contains the downlink throughput is named 'DL_Throughput_Mbps', whereas the cell status is in column 'Cell_Status'. Now, I will filter for active cells and compute the average downlink throughput.

TOOL

```
# Filter active cells
active_cells = df[df['Cell_Status'] == 'Active’]

# Compute average DL throughput
avg_dl_throughput = round(active_cells['DL_Throughput_Mbps'].mean(), 2)
print(avg_dl_throughput)
```

THINK

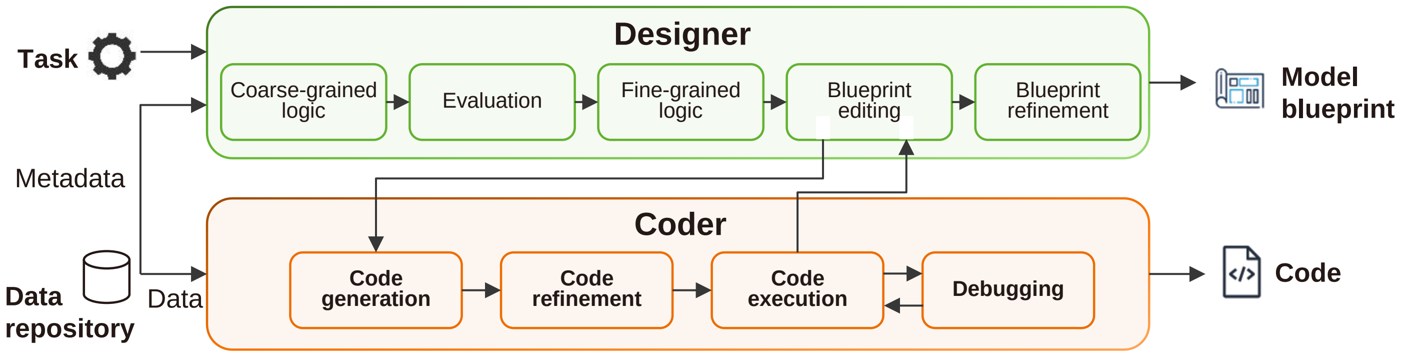
The task has been completed and the final answer has been calculated and formatted as required.

The average cell throughput for active cells is 34.65 Mbps.

Answer:
The average cell throughput for active cells is 34.65 Mbps.

LLM agents for network modelling

- To address LLM limitations in **crafting network models**, we propose a multi-agent framework with dedicated **Designer** and **Coder** agents



HUAWEI

LLM agents for network modelling - Designer

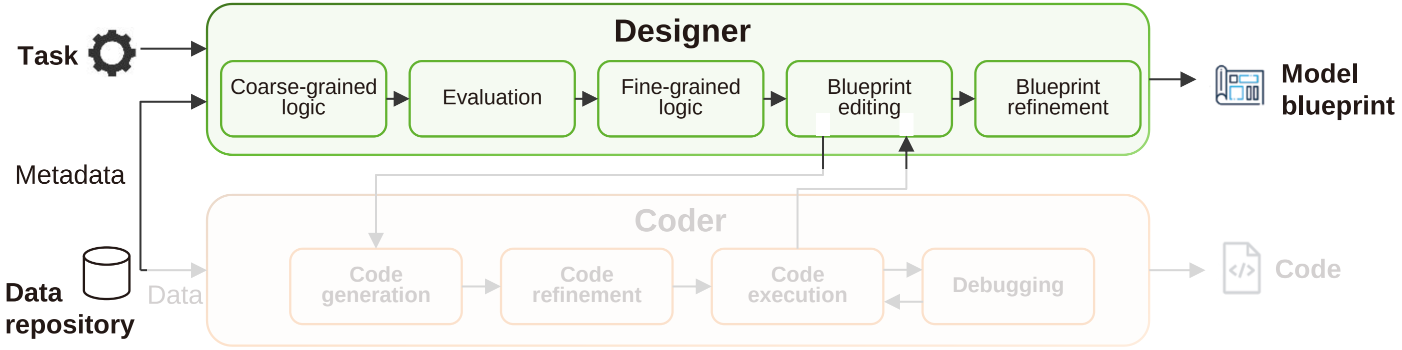
① **Coarse-grained logic**
Agent generates *N* independent high-level reflections consisting of conceptual understanding of the task without details

② **Evaluation**
The reflections are validated against the original task, addressing biases and challenges (e.g., data unavailability)

③ **Fine-grained logic**
Successful reflections are synthesized into a single high quality reflection, producing a comprehensive strategy with associated mathematical formulas and/or pseudo-code

④ **Blueprint editing**
Reflections are synthesized into an initial blueprint represented as a YAML file defining model steps, inputs, outputs, and logic for each step

⑤ **Blueprint refinement**
Blueprint accuracy is ensured by addressing issues (e.g., missing terms) and validating calculations for correctness and scale



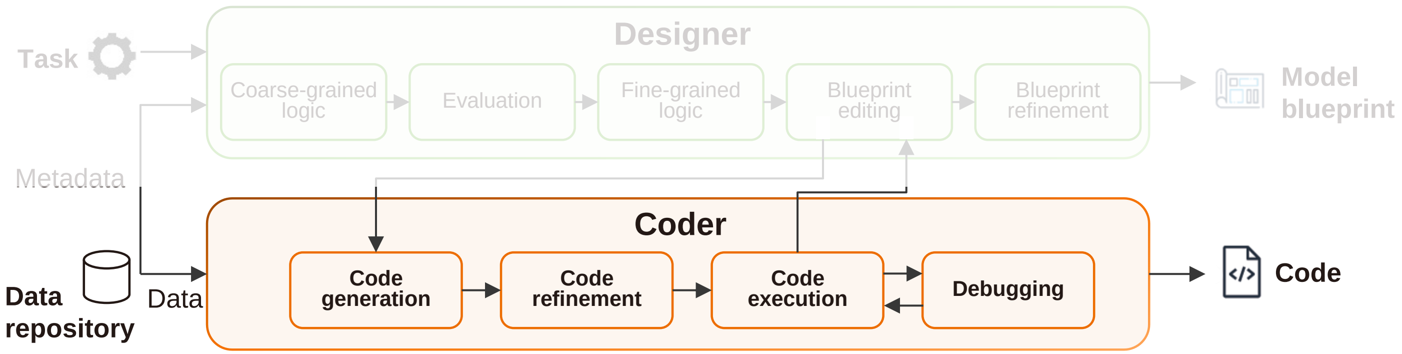
LLM agents for network modelling - Coder

① **Code generation**
Python code is implemented based on the logic of the blueprint

② **Code refinement**
The code is verified and improved, addressing common issues (e.g., appropriate scales, required file loading)

③ **Code execution**
The code is executed; if the execution is successful, the outcome is fed back to the Designer

④ **Code debugging**
If the code execution fails, error reports from the interpreter are evaluated and the code is iteratively refined until successful execution



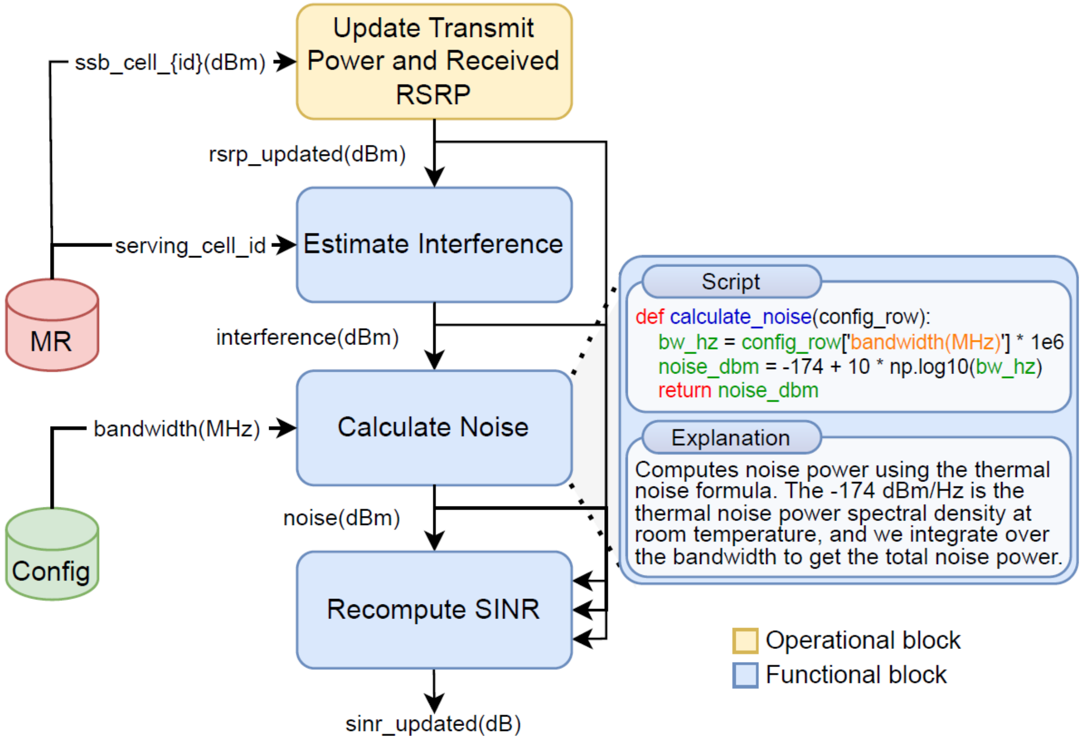
HUAWEI

Results - Tasks

- We consider a **small simulated network** with **19 tri-sector BSs** and evaluate four tasks:
  - **Power control**
    Effect of transmit power changes on UE SINR
  - **Energy saving**
    Effect of deactivating one BS on network energy consumption
  - **Energy saving vs SINR**
    Effect of deactivating one BS on network energy consumption and UEs SINR
  - **New BS deployment**
    Impact of deploying a new BS on UEs SINR

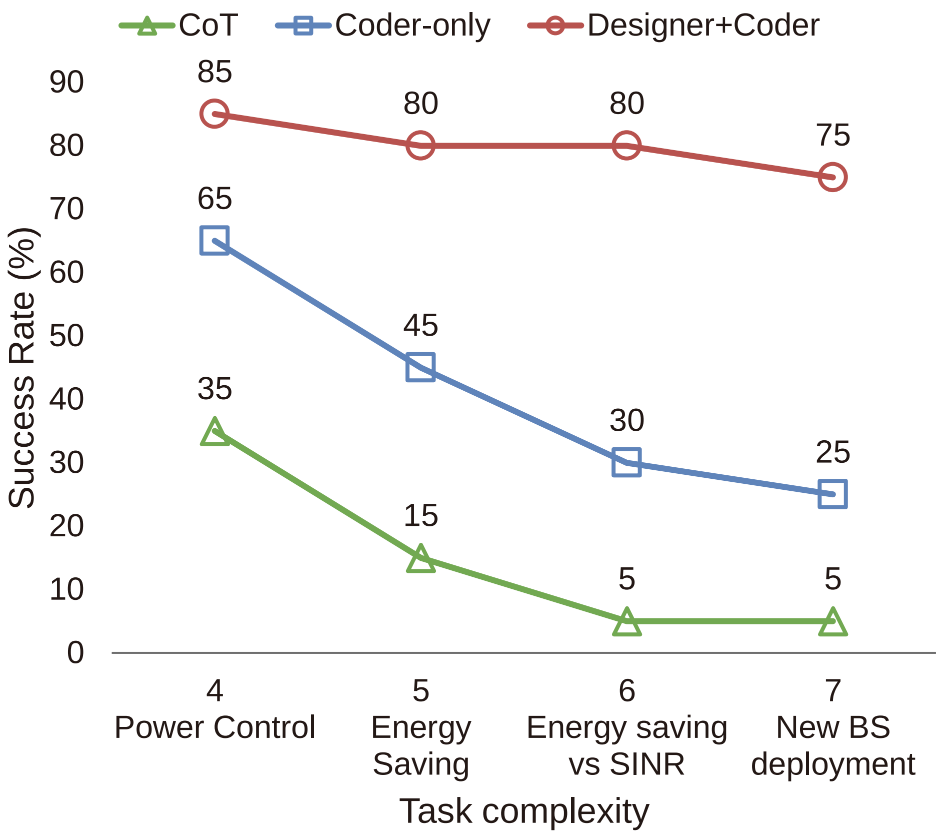
Results - Tasks

- We consider a **small simulated network** with **19 tri-sector BSs** and evaluate four tasks:
  - **Power control**
    Effect of transmit power changes on UE SINR
  - **Energy saving**
    Effect of deactivating one BS on network energy consumption
  - **Energy saving vs SINR**
    Effect of deactivating one BS on network energy consumption and UEs SINR
  - **New BS deployment**
    Impact of deploying a new BS on UEs SINR



Results - Performance

- We benchmark three methods:
  - **Chain-of-Thought (CoT)**: GPT-4o reasoning with direct code generation
  - **Coder-only**: Uses the Coder agent to implement the model design outputted by GPT-4o
  - **Designer+Coder** : Employs both Designer and Coder agents in our framework
- **Agentic framework achieves over 75% success rate**, outperforming other approaches, particularly as task complexity increases



[1] F. Ayed, A. Maatouk, N. Piovesan, A. De Domenico, M. Debbah, Z. Luo, "Hermes: A Large Language Model Framework on the Journey to Autonomous Networks", https://arxiv.org/abs/2411.06490

HUAWEI

Integration with SRCON 2.0

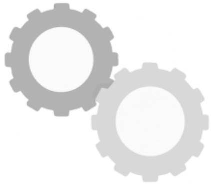
GTSLLM
Domain Reasoning Model

Intermediate to advanced network optimization engineer

- Analysis of poor quality issues in road tests
- Scenario-based problem identification/analysis
- Analysis of KPI issues
- Summary of problem identification
- User wireless quality analysis
- Network optimization analysis report

Base model → **Domain knowledge** → **Domain reasoning enhancement**

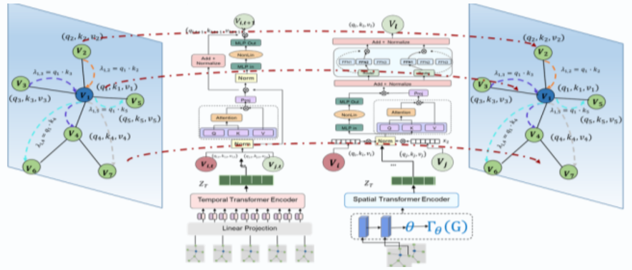
Action



Feedback

SRCON2.0
Wireless Network World Model

Complex wireless network perception/simulation/optimization capabilities



Wireless network mechanism / high-precision perception /... integration

Natural language interaction

Network Optimization Engineer

Integrate with the existing network system

Current network operation system

HUAWEI

Conclusions

1. **Vanilla LLMs are insufficient for autonomous network optimization** due to limitations in numerical reasoning, planning, and reliable execution
2. **Agentic frameworks, integrating multiple specialized LLM agents with access to external tools and memories, represent a promising architecture** for enabling AI-driven, self-optimizing networks
3. **Embedding telecom domain knowledge into LLMs and their tools is essential** to address the complexity and specificity of network operations and standards

Thank you.

nicola.piovesan@huawei.com

HUAWEI