

2025 CCF 国际AIOps挑战赛赛制介绍

聂晓辉

中国科学院计算机网络信息中心

主办单位：中国计算机学会（CCF）

承办单位：中国计算机学会互联网专委会、中国科学院计算机网络信息中心、中国移动研究院、清华大学

协办单位：华为2012实验室、阿里云、中兴通讯、中国移动九天团队、南开大学、西安电子科技大学、清华大学计算机科学与技术系、神州灵云

目录 CONTENTS

第一章 比赛背景

第二章 赛题解读

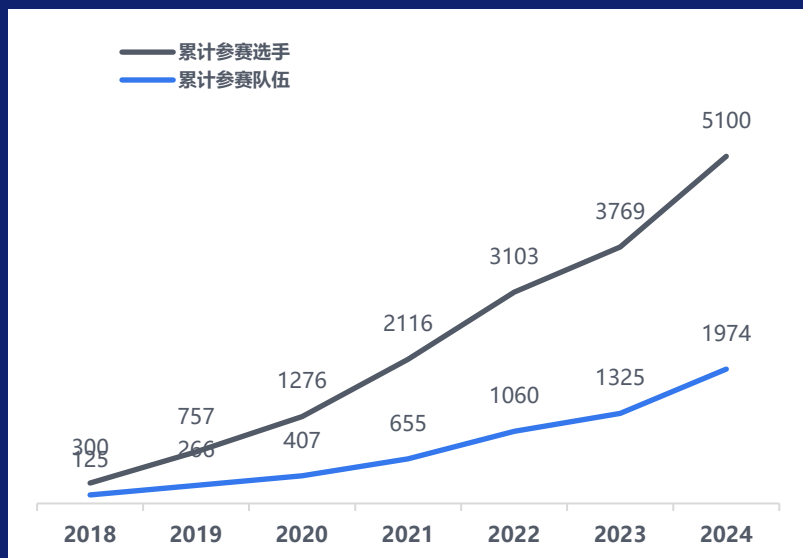
第三章 赛程说明

第四章 参考文献

CCF国际AIOps挑战赛竞赛背景

AIOPS | 2025 CCF国际AIOps挑战赛
2025 CCF International AIOps Challenge

CCF国际AIOps挑战赛（China Computer Federation International AIOps Challenge）由清华大学联合中国计算机学会（CCF）自2018年共同发起。旨在借助社区力量，运用人工智能算法解决各类运维难题。迄今为止已经成功举办七届，吸引了大量AIOps从业者和关注者，随着赛事规模和影响力的不断扩大，形成了运维行业著名的专业赛事品牌，并逐渐发展为智能运维领域最具影响力的专业赛事。



1900+ 参赛队伍

1000+ 参与单位

5000+ 参赛选手

4000+ 社区活跃成员

20,000+ 公众号关注数

150,000+ 在线参与人数

（截止目前 2018-2024 届累计）

CCF国际AIOps挑战赛7年历程 (2018-2024)

AIOPS | 2025 CCF国际AIOps挑战赛
2025 CCF International AIOps Challenge

2018

- KPI异常检测 (125支队伍)
- 来自五家互联网公司的监控指标数据

2019

- 多维监测指标的异常定位 (141支队伍 457名选手)
- 来自两家公司的多维检测指标数据

2020

- 微服务应用系统故障发现和根因定位 (141支队伍 519名选手)
- 来自一家运营商的多种指标、调用链; 复赛提供云资源、在线评测

2021

- 云环境下商业银行应用系统的故障实时检测与根因定位(248支队伍 800+名选手)
- 来自银行的两个系统; 发布指标、调用链和日志数据; 初赛和复赛提供建行云、在线评测

2022

- 微服务架构下电商系统的故障识别与分类 (300+支队伍 900+名选手)
- 故障分类新赛题; 来自搭建的微服务电商系统, 面向k8s云原生架构; 多层级、丰富的故障类型

2023

- 开放式赛题 (265+支队伍 700+名选手)
- 基于日志、调用链、指标的一种或多种模态数据的故障检测、定位、根因分析、影响分析等, 以及大语言模型在运维领域的应用

2024

- 基于检索增强的运维知识问答 (649支队伍 1331名选手)
- 采用大模型检索增强 (RAG) 技术, 基于5G领域运维技术文档, 探索如何结合领域私有技术文档进行高效私域知识问答



2025（第八届）CCF国际AIOps挑战赛由中国计算机学会（CCF）主办，中国计算机学会互联网专委会、中国科学院计算机网络信息中心、中国移动研究院和清华大学承办，华为2012实验室、阿里云、中兴通讯、中国移动九天团队、南开大学软件学院、西安电子科技大学和清华大学计算机科学与技术系等单位协办。本次大赛采用“双赛道”机制，分别设置云服务领域的“基于大模型智能体的微服务根因定位”赛题和无线领域的“基于时间序列的无线网络因果预测与变更决策”赛题，旨在吸引更广泛的参赛群体，进一步激发智能运维技术的创新活力。

<https://challenge.aiops.cn/home/competition>



扫码参赛

<div><div>2025 CCF 国际AIOps 挑战赛INTERNATIONAL AIOPS CHALLENGE</div><div>赛道一 基于大模型智能体的微服务根因定位</div></div>	<div>赛道一: 基于大模型智能体的微服务根因定位</div> <div>53支队伍</div> <div>本赛题旨在评估大模型在根因定位任务中的能力。参赛者将基于微服务系统数据, ...</div> <div>🕒 2025/05/31-2025/09/14</div>	<div>所有赛道总奖金:</div> <div>¥ 140000</div>
<div><div>2025 CCF 国际AIOps 挑战赛INTERNATIONAL AIOPS CHALLENGE</div><div>赛道二 基于时间序列的无线网络因果预测与变更决策</div></div>	<div>赛道二: 基于时间序列的无线网络因果预测与变更决策</div> <div>22支队伍</div> <div>本赛题的核心任务是基于无线网络真实时间序列数据的多任务学习</div> <div>🕒 2025/05/31-2025/09/14</div>	<div>所有赛道总奖金:</div> <div>¥ 140000</div>

赛道一：基于大模型智能体的微服务根因定位

赛题链接：<https://challenge.aiops.cn/home/competition/1920410697896845344>



在大规模微服务系统中，服务依赖复杂、监控数据多样，一旦发生故障，如何快速准确地定位根因组件及其诱因，是保障系统稳定运行的关键。**微服务系统中的根因定位问题。** 面对服务依赖复杂、监控数据多样的挑战，传统方法难以高效准确定位故障根因。随着**大语言模型（LLM）**的发展及其在理解、推理和规划方面的强大能力，基于LLM驱动的智能体和多智能体系统为微服务根因定位提供了新思路。**参赛者需基于比赛提供的微服务系统数据，构建一个融合监控指标、日志和调用链数据的大模型根因定位系统，实现自然语言理解、智能分析、故障场景思维链生成，并自动识别异常的根因组件与故障类型，推动智能体技术在实际运维场景中的应用。**



赛题一：基于大模型智能体的微服务根因定位

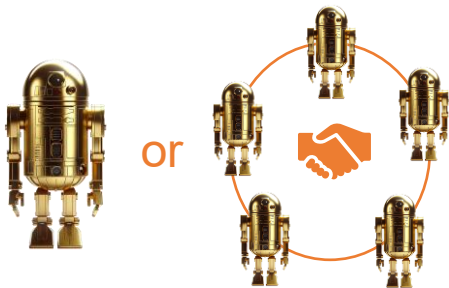
比赛任务

输入

The system experienced an anomaly from 2025-04-30T19:10:15Z to 2025-04-30T19:40:15Z. Please infer the possible causes.

触发定位

基于大模型智能体根因定位系统



单智能体

多智能体协同

可解释的根因分析报告

```
{
  "component": "checkoutservice",
  "reason": "disk IO overload",
  "time": "2025-04-21 12:18:00",
  "reasoning_trace": "reasoning_trace": [
    {
      "step": 1,
      "action": "LoadMetrics(checkoutservice)",
      "observation": "disk_read_latency spike observed at 12:18"
    },
    {
      "step": 2,
      "action": "TraceAnalysis('frontend -> checkoutservice')",
      "observation": "checkoutservice appears multiple times in self-loop spans"
    },
    {
      "step": 3,
      "action": "LogSearch(checkoutservice)",
      "observation": "IOError found in 3 log entries"
    }
  ]
}
```

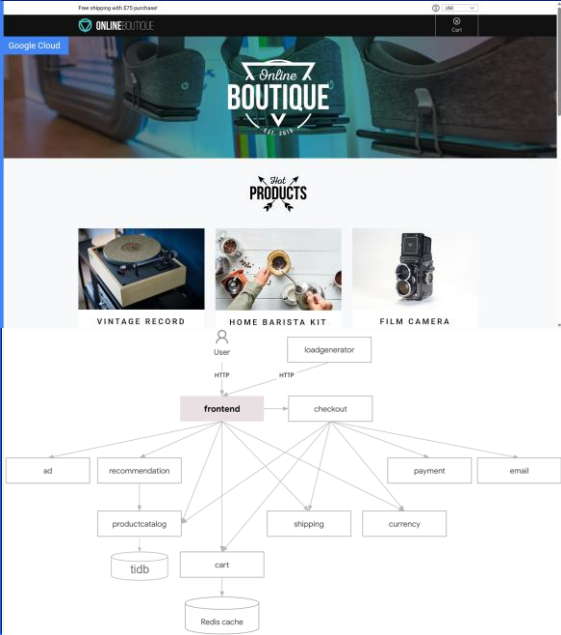
根因组件

故障原因

故障推理过程

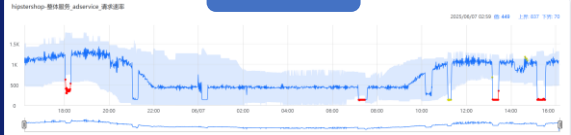
输出报告

微服务系统



采集

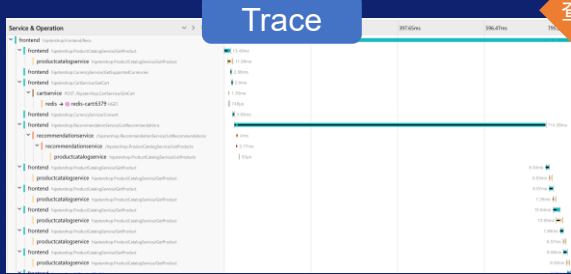
Metric



Log

```
time = 2025-04-30T19:10:15Z
[{"timestamp": "2025-04-30T19:10:15Z", "severity": "info", "source": "checkoutservice", "message": "conversion request successful", "type": "info"}, {"timestamp": "2025-04-30T19:10:15Z", "severity": "error", "source": "checkoutservice", "message": "disk IO error", "type": "error"}]
```

Trace



查询分析

赛题一：评分规则

维度	权重（初始设定）	说明
组件准确率 LA	0.40	是否准确识别出根因组件（component 字段）
原因准确率 TA	0.40	是否准确识别出根因类型/诱因（reason 字段）
推理效率 Efficiency	0.10	推理路径是否紧凑，步数越少得分越高（根据平均推理步数 APL 评估）
推理可解释性 Explainability	0.10	推理链条是否覆盖关键证据，逻辑完整性是否良好（基于结构化轨迹评估）

选手得分 **Final Score = (0.40 × LA + 0.40 × TA + 0.1 × Efficiency + 0.1 × Explainability) × 100**

- LA = 正确预测的组件数量 / 总评估样本数量
- TA = 正确根因数量 / 总评估样本数量
- Efficiency = $\exp(-(APL-5)/5)$, APL为所有正确结果的平均推理步数，最大值为1，步数越多分数越小
- Explainability = 选手推理过程中观察到的正确关键信息数 / 总的关键证据点数量

详细规则见：<https://challenge.aiops.cn/home/competition/1920410697896845344>

APL	Efficiency 得分
4步	1.22 → 截断为 1.00
5步	1.00
10步	0.37
15步	0.13
20步	0.05

国际AIOps挑战赛

去登录

概览数据排行榜规则

大赛报名
创建队伍
参考范例
赛事环境
评分标准
用例说明
技术要求与实现建议
FAQ

维度	权重（初始设定）	说明
组件准确率 LA	0.40	是否准确识别出根因组件（component 字段）
原因准确率 TA	0.40	是否准确识别出根因类型/诱因（reason 字段）
推理效率 Efficiency	0.10	推理路径是否紧凑，步数越少得分越高（根据平均推理步数 APL 评估）
推理可解释性 Explainability	0.10	推理链条是否覆盖关键证据，逻辑完整性是否良好（基于结构化轨迹评估）

注：各评分维度的权重可能根据比赛进展情况进行动态调整，如有调整将通过赛事公告进行说明。

赛题一：一个简单的例子

组委会提供 - 参考答案

```
{
  "component": "checkoutservice",
  "reason": "disk IO overload",
  "time": "2025-04-21 12:18:00",
  "reasoning_trace": "reasoning_trace": [
    {
      "step": 1,
      "action": "LoadMetrics(checkoutservice)",
      "observation": "disk read latency spike observed at 12:18"
    },
    {
      "step": 2,
      "action": "TraceAnalysis('frontend -> checkoutservice')",
      "observation": "checkoutservice appears multiple times in self-loop spans"
    },
    {
      "step": 3,
      "action": "LogSearch(checkoutservice)",
      "observation": "IOError found in 3 log entries"
    }
  ]
}
```

选手提交答案 - 评测示例1 <答案正确>

```
{
  "component": "checkoutservice",
  "reason": "disk IO overload",
  "time": "2025-04-21 12:18:00",
  "reasoning_trace": [
    {
      "step": 1,
      "action": "QueryMetric(checkoutservice)",
      "observation": "I/O latency peak at 12:18"
    },
    {
      "step": 2,
      "action": "TraceCheck",
      "observation": "checkoutservice is on a self-loop chain"
    },
    {
      "step": 3,
      "action": "LogInspection",
      "observation": "multiple IOError records"
    }
  ]
}
```

评分解释:

- ✓ 根因组件正确 → LA = 0.40;
- ✓ 故障类型正确 → TA = 0.40;
- ✓ Reasoning_trace 共 3 步, 推理紧凑 → Efficiency = 1.00 × 0.10 = 0.10;
- ✓ 命中 3 个关键证据点 (指标、日志、调用链) → Explainability = 0.10;
- 💡 总分: 0.40 + 0.40 + 0.10 + 0.10 = 1.00

赛题一：数据与比赛环境

1.

比赛数据开源，总计提供**400个故障数据**，**分两阶段给出**，**每次200个**，故障类型包括：Service、POD、Node的 JVM、数据库、网络、资源、变更（代码、配置错误）等故障类型；
2.

中国科学院计算机网络信息中心将为赛事提供服务器算力及大模型接口支持，报名组队成功后填写 <https://www.wjx.top/vm/tUFzNJC.aspx>，主办方将主动联系各参赛团队协调资源分配事宜。

2025 CCF
国际AIOps
挑战赛
AIOps CHALLENGE

赛道一
基于大模型智能体的微服务根因定位

☆赛道一: 基于大模型智能体的微服务根因定位

本赛题旨在评估大模型在根因定位任务中的能力。参赛者将基于微服务系统...
© 2025/05/31 - 2025/09/14
主办单位: 中国计算机学会 (CCF)

所有赛道总奖金:
¥ 1400000

参加比赛

参赛人数
72

开始

报名截止
2025-06-26

结束

① 赛题

② 数据

③ 排行榜

④ 规则

招募

本次比赛使用由主办方构建的微服务环境生成故障数据。具体包括指标、日志、trace等数据。具体数据格式可参考以下样例数据：
https://www.aiops.cn/gitlab/aiops-live-benchmark/aiopschallengedata2025/-/tree/main/sample/ref_type=heads
初赛所用数据将分两个阶段提供：
第一阶段（6月27日）（详细地址待后续更新）
第二阶段（7月18日）（详细地址待后续更新）

总体概述

本数据集包含系统采集的三类关键数据：监控指标（Metric）、分布式调用链（Trace）和容器日志（Log）。数据均以 Parquet 格式存储，以便高效读取与分析。整体结构如下：

log-parquet

log_filebeat-server_2025-05-27_00-00-00.parquet

log_filebeat-server_2025-05-27_01-00-00.parquet

log_filebeat-server_2025-05-27_02-00-00.parquet

metric-parquet

app

pod

pod_adservice-0_2025-05-27.parquet

pod_ms_hipstershop_2025-05-27.parquet

service

service_adservice_2025-05-27.parquet

infra

infra_node

infra_node_cpu_usage_rate_2025-05-27.parquet

infra_node_disk_read_bytes_total_2025-05-27.parquet

infra_pod

infra_pod_pod_cpu_usage_2025-05-27.parquet

infra_pod_pod_fs_reads_bytes_2025-05-27.parquet

infra_tldb

infra_tldb_block_cache_size_2025-05-27.parquet

infra_tldb_connection_count_2025-05-27.parquet

other

infra_pd_abnormal_region_count_2025-05-27.parquet

infra_pd_leader_count_2025-05-27.parquet

infra_tikv_available_size_2025-05-27.parquet

trace-parquet

trace_jaeger-span_2025-05-01_00-59-00.parquet

trace_jaeger-span_2025-05-01_01-59-00.parquet

容器日志数据

POD、Service、系统的应用指标

NODE、POD的机器指标

TIDB数据库监控指标

Trace



中國科技雲

China Science & Technology Cloud

中國科技雲大模型API接口使用手册

帮助中心

Home

中國科技雲

大模型API接口

使用手册

索引

Powered by YouTiao & Tencent, thank you!

* 1. 团队名称（注：填写问卷前，请在挑战赛官网完成“确认组队”，方可申请，单人参赛也需要点击确认。发放资源前，会核实组队完成信息。）

* 2. 团队ID

* 3. 队长姓名

* 4. 队长手机号

* 5. E-mail

* 6. 队长微信号

挑战赛参赛选手：**8C32GB机器 + 固定额度DeepSeek模型TOKEN**

答案提交次数：**最多350次，平均1天10次提交答案评测机会**

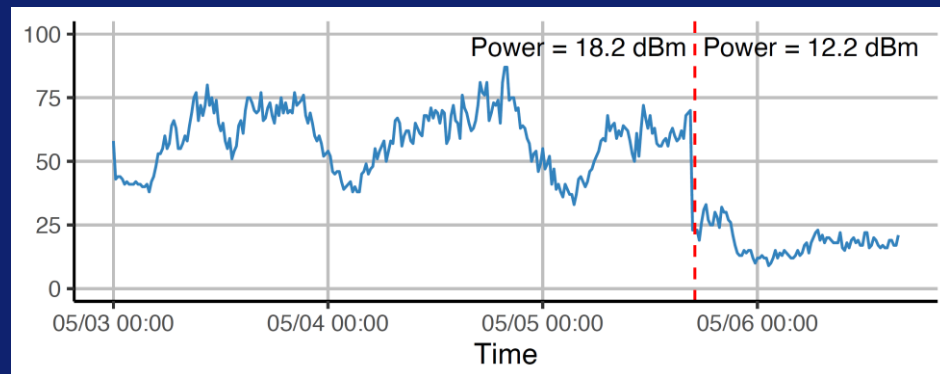
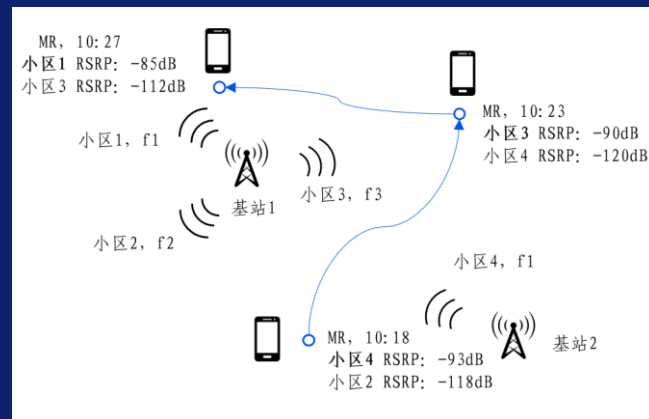
赛道二：基于时间序列的无线网络因果预测与变更决策

赛题链接：<https://challenge.aiops.cn/home/competition/1920775077574500373>



随着4G运维向纵深和精细层面发展、5G全面商用，无线网络通过多样的参数和大量的网络设备承载了无数连接，提供了超密集的覆盖，并产生了海量时间序列数据。为了保障无线网络稳定健康运行、减轻运维人员工作负担，需要更智能的运维技术应对日益增长的用户流量和网络设备数量。围绕无线网络的容量规划场景，本赛道要求参赛队伍完成如下由易到难的三个任务：

1. **时序预测：**预测平时KPI的变化趋势；
2. **因果预测：**预测变更（即调整参考信号功率）后KPI的变化趋势；
3. **变更决策：**对于期望的网络状态（实际数据中变更之后一天内的KPI），给出变更方案（实际数据中参考信号功率新的取值）。



任务一：时序预测

输入

```
ChangeInfo(  
  cell_id="cell-1",  
  start_time="2025-01-02 00:00:00"  
)
```

历史数据

```
{  
  "cell-1": {  
    "KPI": {  
      "rrc_connmax": [1, 2, 3],  
      "date_time": [  
        "2025-01-01 23:15:00",  
        "2025-01-01 23:30:00",  
        "2025-01-01 23:45:00"  
      ]  
    },  
    "NRM": {  
      "ReferenceSignalPower": [9.0],  
      "time": ["2025-01-01 00:00:00"]  
    }  
  }  
}
```

多时间序列
预测模型

输出

给定时刻之后KPI的预测值

```
{  
  "rrc_connmax": [4, 5, 6],  
  "date_time": [  
    "2025-01-02 00:00:00",  
    "2025-01-02 00:15:00",  
    "2025-01-02 00:30:00"  
  ]  
}
```

任务二：因果预测

输入

```
ChangeInfo(  
  cell_id="cell-1",  
  start_time="2025-01-02 00:00:00",  
  change={"ReferenceSignalPower": 10.0}  
)
```

历史数据

```
{  
  "cell-1": {  
    "KPI": {  
      "rrc_connmax": [1, 2, 3],  
      "date_time": [  
        "2025-01-01 23:15:00",  
        "2025-01-01 23:30:00",  
        "2025-01-01 23:45:00"  
      ]  
    },  
    "NRM": {  
      "ReferenceSignalPower": [9.0],  
      "time": [  
        "2025-01-01 00:00:00"  
      ]  
    }  
  }  
}
```

多时间序列
预测模型

输出

给定时刻之后KPI的预测值

```
{  
  "rrc_connmax": [5, 6, 7],  
  "date_time": [  
    "2025-01-02 00:00:00",  
    "2025-01-02 00:15:00",  
    "2025-01-02 00:30:00"  
  ]  
}
```


赛道二：基于时间序列的无线网络因果预测与变更决策

任务三：变更决策

输入

变更小区: "cell-1"

变更时刻: "2025-01-02 00:00:00"

历史数据

```
{
  "cell-1": {
    "KPI": {
      "rrc_connmax": [1, 2, 3],
      "date_time": [
        "2025-01-01 23:15:00",
        "2025-01-01 23:30:00",
        "2025-01-01 23:45:00"
      ]
    },
    "NRM": {
      "ReferenceSignalPower": [9.0],
      "time": [
        "2025-01-01 00:00:00"
      ]
    }
  }
}
```

变更目标/变更后的KPI

```
{
  "rrc_connmax": [5, 6, 7],
  "date_time": [
    "2025-01-02 00:00:00",
    "2025-01-02 00:15:00",
    "2025-01-02 00:30:00"
  ]
}
```

多时间序列
预测模型

输出

{"ReferenceSignalPower": 10.0}

赛题二：评估规则

对于模型输出的结果，采用SMAPE（对称平均绝对百分比误差，Symmetric Mean Absolute Percentage Error）进行准确度评分。

$$\text{SMAPE} = \frac{1}{n} \sum_{i=1}^n \frac{2|\hat{y}_i - y_i|}{|\hat{y}_i| + |y_i|}$$

$$\text{Score}(\tau) = \frac{\max(\tau - \text{SMAPE}, 0)}{\tau} \times 100\%$$

1. 初赛评分榜分数：

- **Score = 0.6 × 时序预测Score(1) + 0.4 × 因果预测Score(1)**

2. 评审会榜：

- **Score = 0.5 × 因果预测Score(1) + 0.5 × 变更决策Score(0.25)**

赛题二：数据与比赛环境

1. 真实移动无线网络数据，训练集包含约2万个无线小区，3周的21GB数据，数据闭源
2. 初赛开始后，选手自行在[中国移动开放创新平台（九天·毕昇）](#)报名组队即可使用平台提供的计算资源访问数据、完成算法。平台报名截止到7月11日。

① 概览

数据

排行榜

规则

初赛

样例数据及对应目录结构参加[此链接](#)。

字段的中英文名称对照参见提供的工具包中[aiops_challenge_2025/config/目录](#)下的yaml文件。

更多数据描述参见[中国移动开放创新平台（九天·毕昇）](#)的赛题说明-数据集部分。

```
<data_dir-1>
|- <cell_id-1>
| |- KPI.csv, 性能数据
| |- NRM.csv, 参数
| |- EP.csv, 工程参数
| |- PMRelation.csv, 小区对性能数据
| |- MR.csv, 按小区对粒度汇总的测量报告
| - .....
|- cells.csv, 每行一个<cell_id>用于检索
<data_dir-2>
|- .....
.....
```

中国移动
China Mobile

九天人工智能

九天·毕昇

体验版

首页

学习

比赛

求职

教学

科研

九天·毕昇

一站式人工智能学习和实战平台

基于中移九天深度学习平台，为AI学习者提供充沛的GPU算力、丰富的数据和学习实战资源，服务课程学习、比赛打榜、工作求职等全流程场景，并面向高校提供在线教学、科研开发的一站式解决方案

* 资源套餐

请选择

CPU 8核 内存 32G (0.1算力豆/6分钟)

NVIDIA V100 虚拟化 CPU 2核 内存 16G (0.5算力豆/6分钟)

剩余可用存储

第三章 赛程介绍

关键时间点

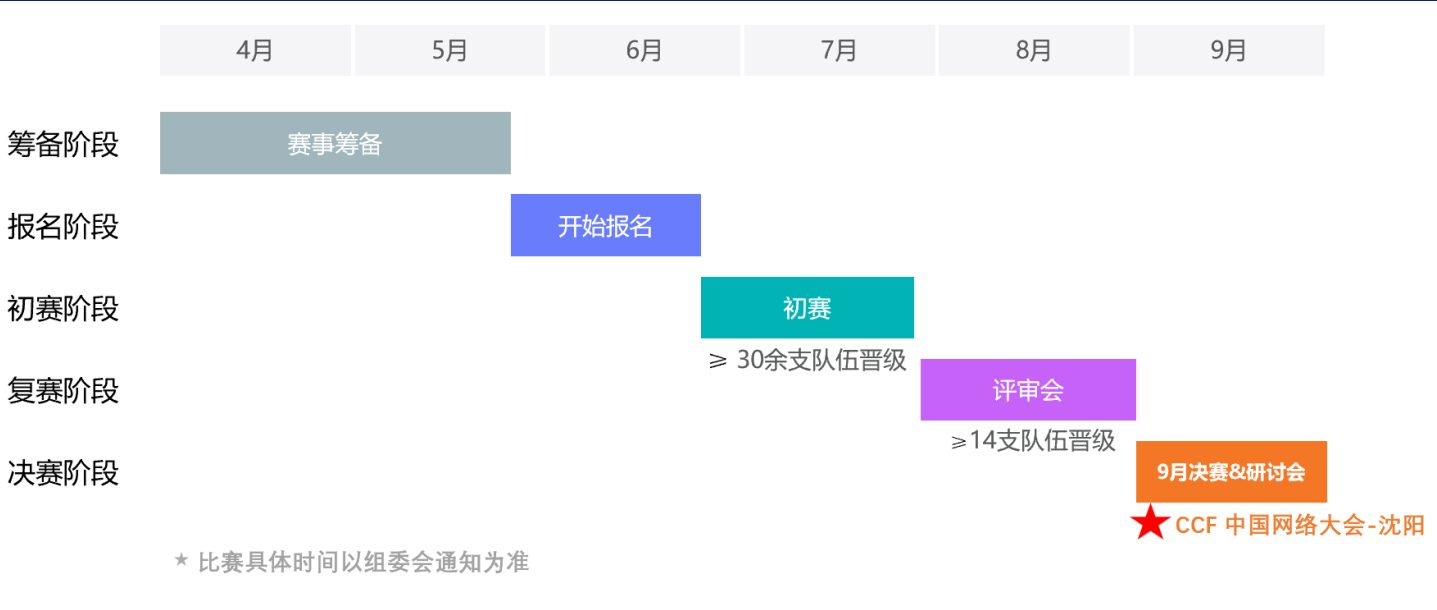
2025年06月26日 18:00:00 **报名截止**

2025年06月27日 10:00:00 **初赛开始**

2025年08月01日 18:00:00 **初赛结束**

2025年08月10日-9月1日 **评审会**

2025年09月12-14日 **决赛&研讨会**



报名方式

大赛面向全社会开放，高校，科研机构，行业从业人员等均可参与

报名前

- <https://challenge.aiops.cn/home/competition>注册账号
- 提交身份认证审核，通过后绑定手机号和填写账户资料，之后点击“参加比赛”



报名后

- 组建团队（用户名或链接邀请），组建团队完成后需由队长确认组队完成
- 参赛队伍人数不限，可以单人组队，也可以自由组队。但资源分配按照最多5人次分配
- 两个赛道均可参加，同一个赛道每名选手限参加1支参赛队伍
- 选手确认组队完成即可申请机器资源

报名截止时间约为 2025 年 6 月 26日

答疑

微信群后续会进行赛事通知和技术答疑，建议队伍里每个成员都能入群，加社区小助手即可邀约进群



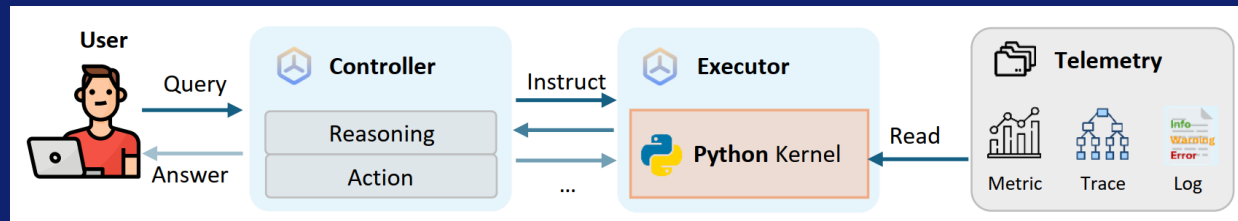
社区小助手

注：本届参赛选手将在OpenAI Ops社区9-11群中进行交流

第四章 参考文献

赛道一：OpenRCA (ICLR 25)

- 提出任务：输入自然语言查询 + 日志、指标、调用链 → 输出结构化根因（时间+组件+原因）
- RCA-agent设计：将大模型与Python代码执行结合，分工进行分析与推理
 - Controller 负责决策与推理
 - Executor 负责生成并执行代码
- 遵循常规排障流程：指标分析 → 调用链定位 → 日志验证



Query

"Within the time range of March 6, 2021, from 23:00 to 23:30, there was one reported failure in the system. Your task is to determine the root cause component."

Telemetry (of the whole bank system)

Metric

```
MG01-CPU_Load: [1.12, 1.04, 1.15, ...]
IG01-MEM_Free: [79.0, 80.0, 79.0, ...]
Mysql02,FS_Capacity: [17.6891, 17.6912, ...]
Tomcat04,PROC_Count: [1.0, 3.0, 3.0, 2.0, ...]
Redis01,Bandwidth_Util: [0.0, 0.4, 0.2, 0.0, ...]
```

Trace

```
IG01->MG01->MG01
      |->IG01->Tomcat02
      |   |->Tomcat03->Mysql01
      |   |->Tomcat04->Mysql02
      |   |->Tomcat01->Tomcat01->Tomcat01
```

Log

```
Tomcat03:
[Info] CMS-concurrent-sweep-start
[Info] CMS-concurrent-sweep:1.017/1.023 secs
...
Tomcat04:
[Info] CMS-concurrent-reset-start
[Info] Time:user=1.02,sys=0.00,real=1.02 secs
...
```

Reasoning Chain (Thought [T] and Action [A]) summarized version

- [T]: None
- [A]: Read files in '2021_03_06/metric/' and obtain all valid KPI names in 'kpi_name' fields.
- [T]: Valid KPIs identified: CPU_Load, MEM_free, ... We should perform anomaly detection first.
- [A]: Extract data between 23:00 to 23:30 and calculate P95 thresholds for these KPIs.
- [T]: Data extraction success. The thresholds are ... We should continue anomaly detection.
- [A]: Filter the anomalous data points that exceed the thresholds in all KPIs.
- [T]: Anomalies identified: Redis01, Mysql02, etc. We should confirm the true faults then.
- [A]: Filter isolated data points and sudden spikes, which is likely to be the false positive.
- [T]: Faults confirmed: Mysql01, Mysql02, Tomcat01. We should identify the root cause now.
- [A]: Read files in '2021_03_06/traces/' to extract traces involving these fault components.
- [T]: Several corresponding traces are extracted. We should analyze their dependency next.
- [A]: Sort the traces by occurrences and print the most common call chain in the tree structure.
- [T]: The most common chain is ...->Tomcat01->Tomcat01. Need log analysis on Tomcat01.
- [A]: Read files in '2021_03_06/logs/' and extract logs generated by Tomcat01.
- [T]: Multiple logs of Tomcat01 show TCP connection fluctuation and memory allocation failure.
- [A]: We can stop reasoning and make a conclusion now.

Conclusion: Tomcat01 faces network packet loss and potential memory leaks.

Structured Prediction

```
{
  "root cause 1": {
    "component": "tomcat01",
    "reason": "network packet loss"
  }
}
```

Ground-Truth Label

```
{
  "root cause 1": {
    "time": "23:18:00",
    "component": "tomcat01",
    "reason": "network latency"
  }
}
```

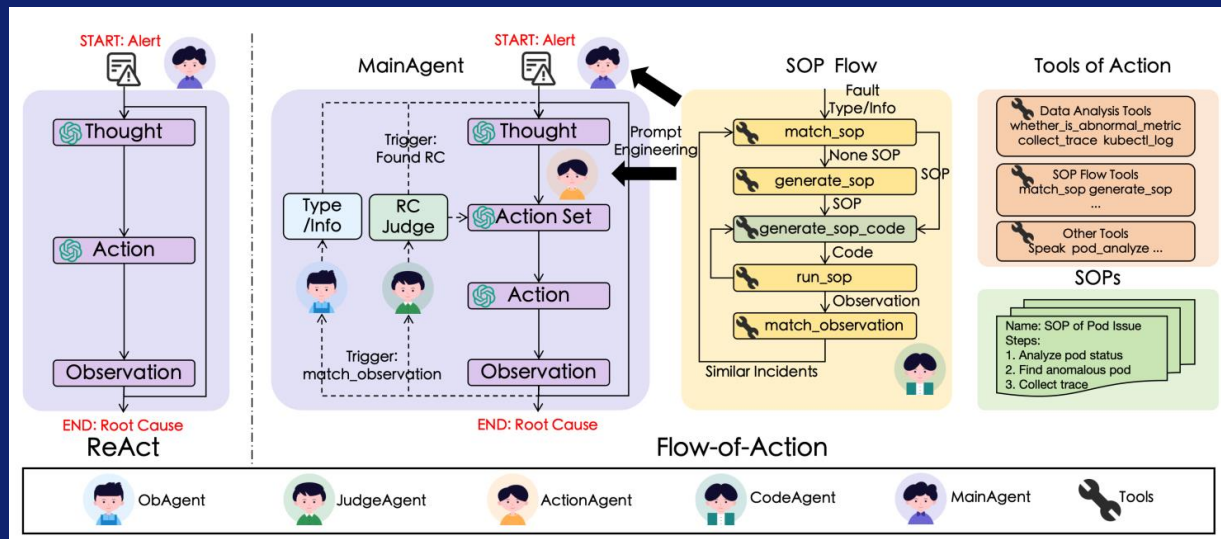
Evaluation Result

```
{
  "correct element": "component",
  "wrong element": "None",
  "ignored element": "reason",
  "evaluation result": true
}
```

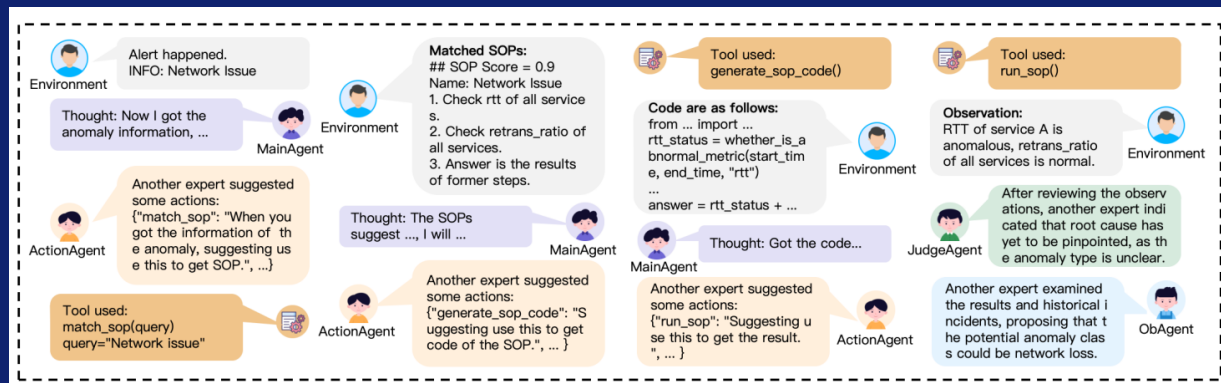
Xu J, Zhang Q, Zhong Z, et al. OpenRCA: Can Large Language Models Locate the Root Cause of Software Failures?[C]//Proceedings of the Thirteenth International Conference on Learning Representations (ICLR 2025). 2025: Singapore EXPO, April 24–28, 2025.

赛道一： Flow-of-Actions(WWW 25)

- 引入“标准操作流程 SOP”机制,用专家经验总结的流程模板约束模型推理路径
- 多智能体协作框架设计
 - MainAgent 负责全局控制
 - ObAgent 自动提取故障信息
 - JudgeAgent 判断是否已找到根因
 - CodeAgent 负责将 SOP 转换为代码并执行分析
- SOP可自动生成并转为可执行代码，解决流程缺失问题

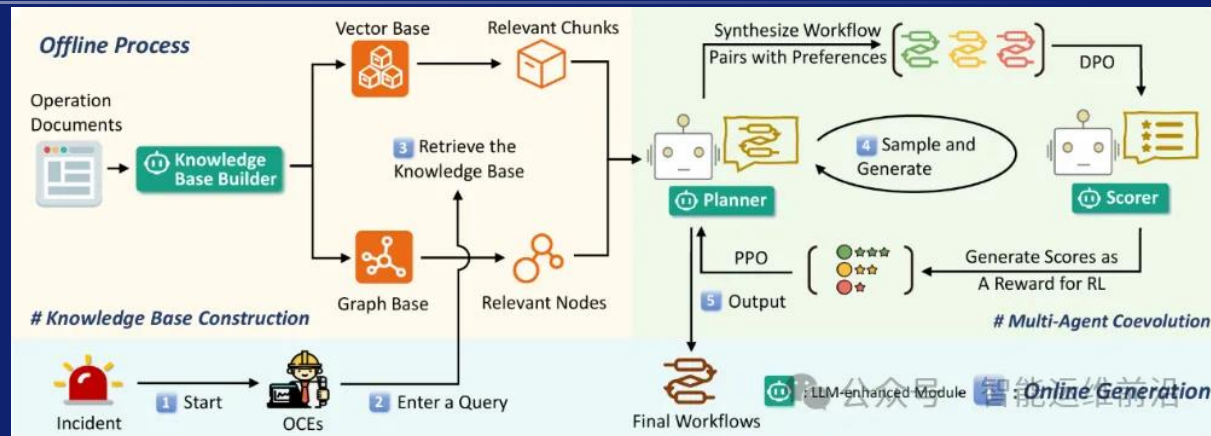


Pei C, Wang Z, Liu F, et al. Flow-of-Action: SOP Enhanced LLM-Based Multi-Agent System for Root Cause Analysis[C]//Proceedings of the ACM Web Conference 2025 (WWW '25). 2025: Sydney, NSW, Australia, April 28 – May 2, 2025.

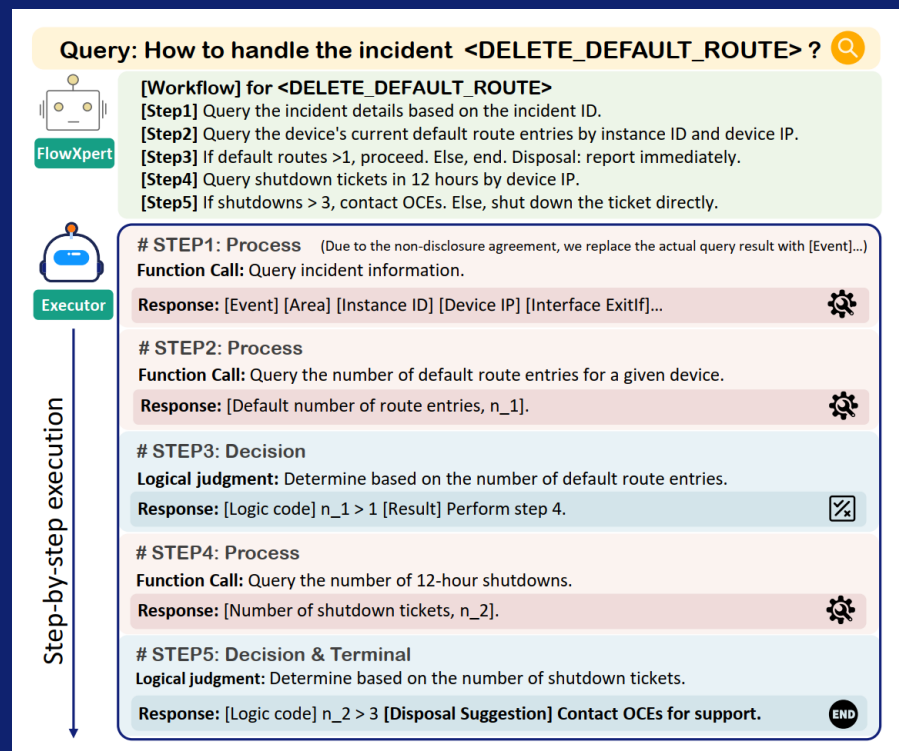


赛道一： FlowExpert(KDD 25)

- 从运维文档中抽取事件节点，构建具备上下文语义与依赖结构的知识图谱，为推理提供精确支撑。
- 采用“思维分解”策略驱动多轮推：将自然语言查询拆解为多个子问题（指标、链路、日志等），由大模型逐步引导工具协同完成。
- 结合链路依赖图进行组件可达性过滤。

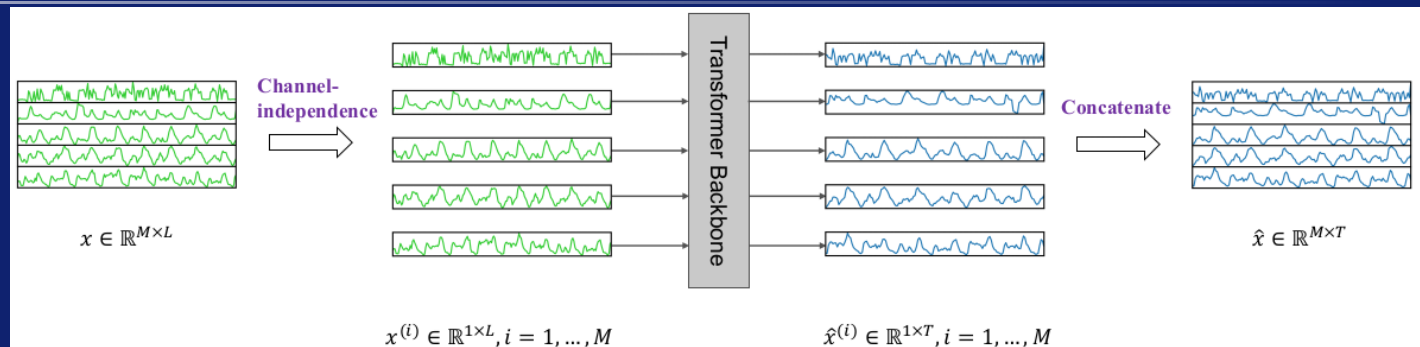


Shi B, Luo Y, Wang J, et al. FlowXpert: Expertizing Troubleshooting Workflow Orchestration with Knowledge Base and Multi-Agent Coevolution[C]//Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2025: Toronto, Canada, August 3–7, 2025.

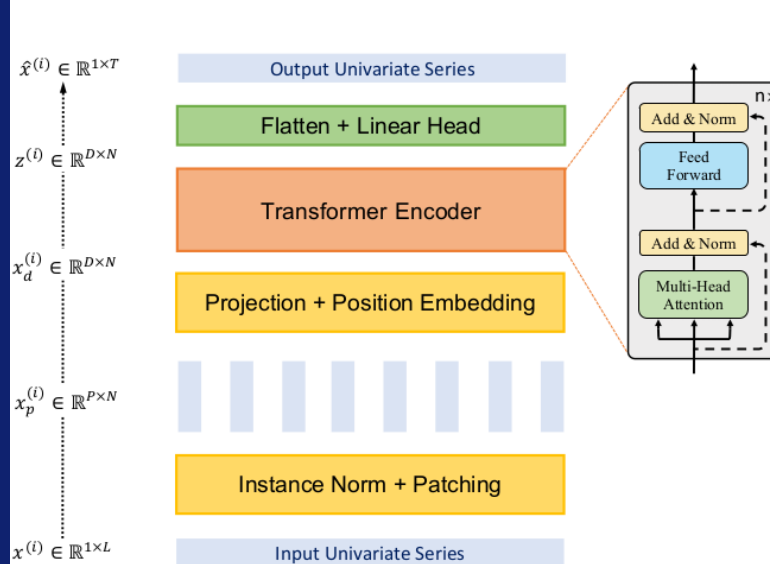


赛道二： PatchTST (ICLR 23)

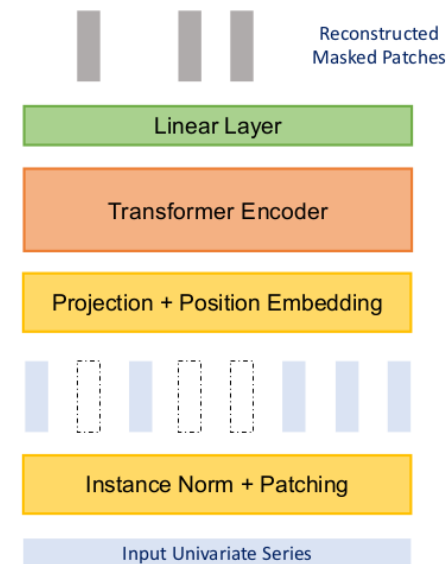
- 信道独立：分别处理各时间序列
- 切片：以子序列为编码基础，而非单个值
- 参数共享：各时间序列共享Transformer模块



(a) PatchTST Model Overview



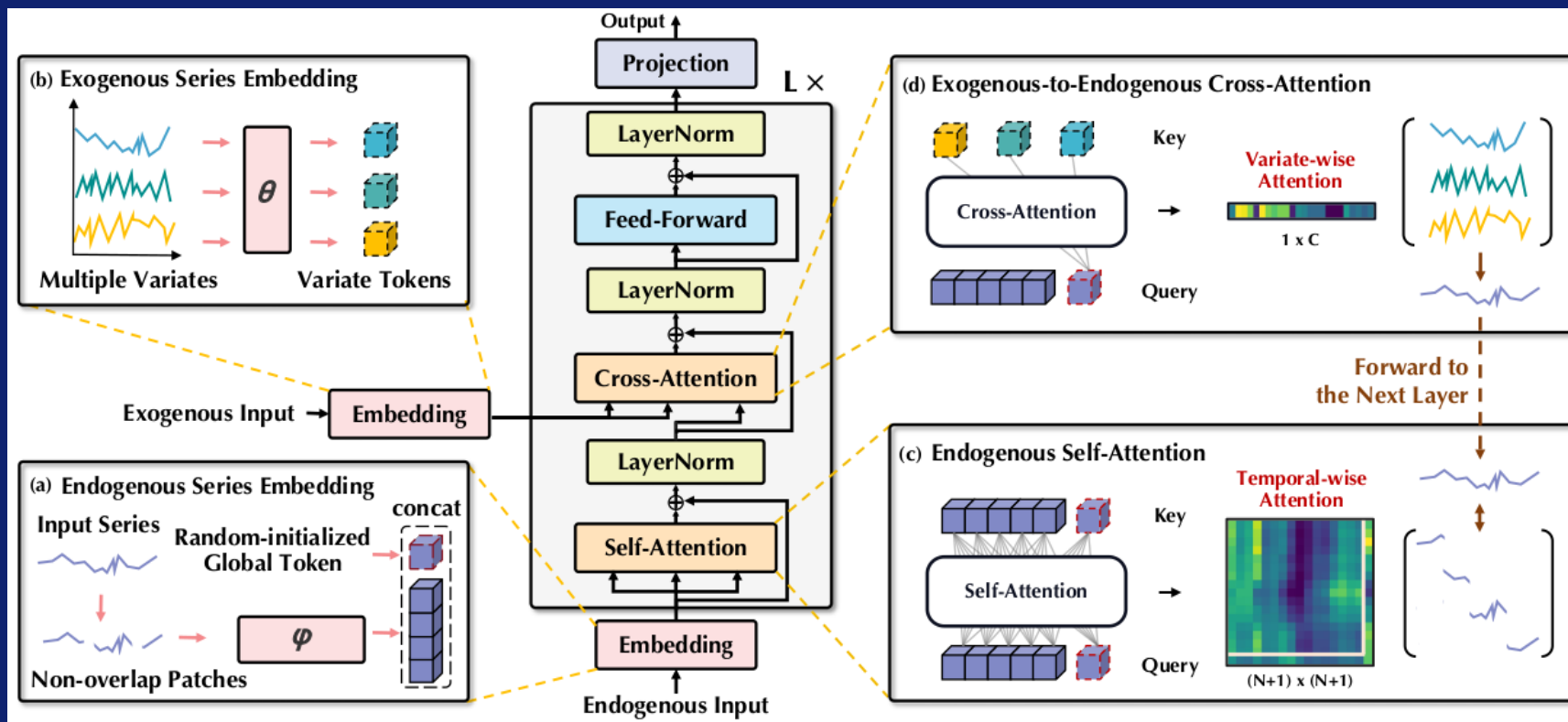
(b) Transformer Backbone (Supervised)



(c) Transformer Backbone (Self-supervised)

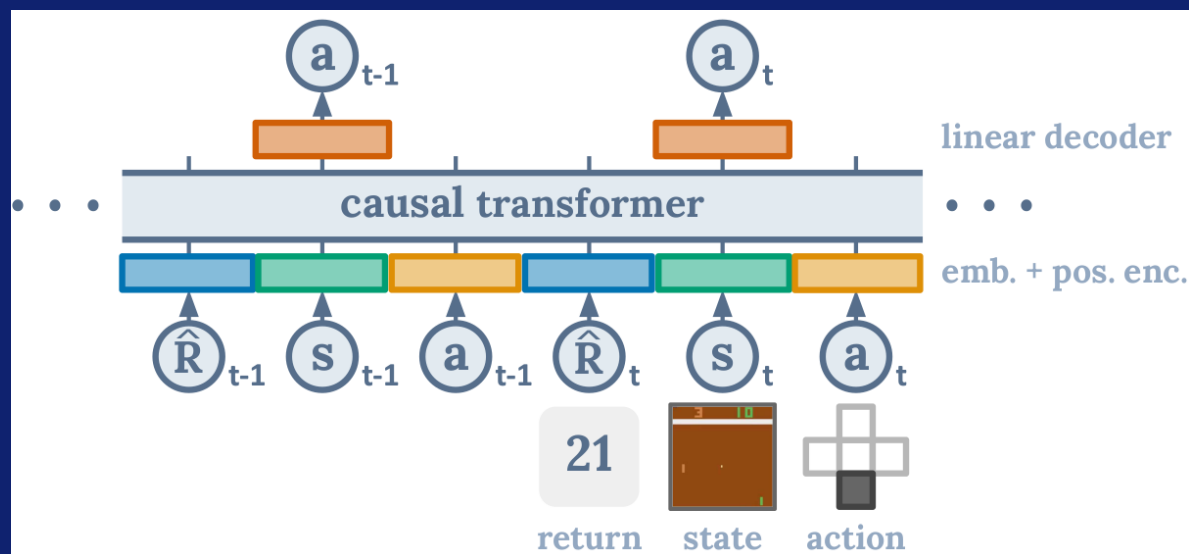
赛道二： TimeXer (NIPS 24)

- 独立编码外源变量
- 内源变量 (KPI等指标) 与外源变量计算互注意力



赛道二： DecisionTransformer (NIPS 21)

- 把序贯决策问题建模为奖励、状态、动作序列
- 应用Transformer拟合下一步的动作



THANKS



扫码参赛



扫码入群

主办单位：中国计算机学会（CCF）

承办单位：中国计算机学会互联网专委会、中国科学院计算机网络信息中心、中国移动研究院、清华大学

协办单位：华为2012实验室、阿里云、中兴通讯、中国移动九天团队、南开大学、西安电子科技大学、清华大学计算机科学与技术系、神州灵云