Contribution ID: 8 Type: 海报展示

基于大语言模型的动态高精度场景驱动人因可靠性数据采集方法研究

Abstract

Human reliability data is critical for advancing reliability analysis methodologies. However, existing data collection methods lack dynamic features and suffer from coarse granularity, with a heavy reliance on expert knowledge, making the process time-consuming and labor-intensive. To address these challenges, this paper proposes a scenario-driven, automated human reliability data collection paradigm based on large language models (LLMs). Our approach focuses on the workload indicators behind human errors. First, we collected operator workload data in a high-temperature gas-cooled reactor (HTGR) control room environment, using the NASA-TLX questionnaire under different scenarios. Subsequently, we fine-tuned the large language model Qwen-7B based on the collected real-world data. Experimental results demonstrate that the proposed method can simulate the cognitive workload of operators in real-time within the explored scenarios. Compared to traditional HUNTER methods, our approach offers significant advantages and outperforms existing commercial large language models in prediction accuracy, providing more accurate, flexible, and scalable workload estimates.

Keywords

Automated Human Reliability Data Collection; Large Language Models; Dynamic Load Prediction; High-Temperature Gas-Cooled Reactor (HTGR)

摘要

人因数据对于推进可靠性分析方法具有重要意义。然而，现有的数据收集方法缺乏动态特征，且颗粒度较粗糙，大量依赖专家知识，过程费时费力。为了解决这些挑战，本文提出了一种基于大语言模型的场景驱动的自动化人因可靠性数据收集研究范式。我们的方法聚焦于人因失误背后的工作负荷指标，首先，我们基于 NASA-TLX 问卷，在高温气冷堆 1：1 主控室环境收集了操作员在不同场景下的工作负荷。接着，基于所收集的真实数据微调大语言模型 qwen-7B。实验结果表明，所提出的方法能够实时模拟在探讨场景下的操作员的认知负荷，相比于传统的 HUNTER 方法具有使用优势，且预测精度优于现有的商业大语言模型，能够提供更加准确、灵活和可扩展的工作负荷估计。

关键词

人因可靠性数据自动收集；大语言模型；动态负荷预测；高温气冷堆

Authors: Prof. 王, 海涛; Prof. 梁, 金刚; 贾, 倩倩; 童, 节娟; 萧, 星宇 (清华大学)

Presenter: 萧, 星宇 (清华大学)

Session Classification: 海报展示

Track Classification: 02 海报展示: 海报展示