Research Seminar

Towards Edge-Native Foundation Models

Speaker: Prof. Song GUO, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology

Date: March 19, 2024 (Tue)

Time: 16:00 (HKT)

Venue: Room 308, Chow Yei Ching Building, HKU

POSTPONED

Abstract

Foundation Models like GPT, LLaMA, and DALL-E have been transformative in AI, demonstrating remarkable versatility across tasks. Yet, the full potential of edge computing—with its inherent benefits in cost, latency, and privacy—remains untapped for deploying these models. In this talk, we unveil the concept of Edge-native Foundation Models, an innovative approach that harnesses the power of distributed, diverse, and collaborative edge computing resources. We introduce a user-friendly, live FM-as-a-service system, allowing seamless access to Foundation Model services without the burdens of expensive deployment or intricate management. Furthermore, we present a novel, environment-responsive adaptation strategy for Edge-native FMs, enabling rapid tuning to meet the dynamic demands of edge environments. Crucially, our methodology emphasizes a commitment to ethical standards and regulatory compliance for AI governance at the edge. We envisage a future where Foundation Models are conceived, nurtured, and utilized within the edge ecosystem. Our empirical findings suggest that Edge-native Foundation Models could level the AI playing field, disrupt the centralization of data processing, and offer a viable, scalable architecture for AI's evolution.

About the speaker:

Song Guo is a full professor in the Department of Computer Science and Engineering at Hong Kong University of Science and Technology. Prof. Guo made fundamental and pioneering contributions to the development of edge AI and cloud-edge computing. He published many papers in top venues and received over a dozen Best Paper Awards from IEEE/ACM conferences, journals and technical committees. He is the recipient of 2024 Edward J. McCluskey Technical Achievement Award, Gold Medal in 2023 Geneva Inventions Expo, and Intellectual Property Ambassador Award in 2020 Hong Kong Social Enterprise Competition. Prof. Guo is a Fellow of the Canadian Academy of Engineering, Member of Academia Europaea, and Fellow of the IEEE.