RUNTIME MONITORING OF MACHINE LEARNING-BASED SCHEDULING ALGORITHMS TOWARD ROBUST DOMAIN-SPECIFIC SYSTEMS-ON-CHIP

WARF: P250031US01

Inventors: Umit Ogras, Ahmet Goksoy, Alish Kanani

Machine learning (ML) algorithms are being rapidly adopted to perform dynamic resource management tasks in heterogeneous system-on-chips. For example, ML-based task schedulers can typically make quick, high-quality decisions at runtime. However, like any ML model, offline-trained policies for scheduling decisions depend critically on the representativeness of the training data. Hence, ML model performance may diminish or even catastrophically fail under unknown workloads, especially new applications. Therefore, there is a need for improved monitoring of ML- based automated schedulers.

The Invention

UW-Madison researchers have developed a novel framework to continuously monitor the system to detect unforeseen scenarios using a gradient-based generalization metric called coherence. The proposed framework accurately determines whether the current policy generalizes to new inputs. If not, it incrementally trains the ML scheduler to ensure the robustness of the task-scheduling decisions.

Tech Fields
- Information Technology : Computing methods, software & machine learning
- Information Technology : Hardware

For current licensing status, please contact Michael Carey at mcarey@warf.org or 608-960-9867