PhD Forum Abstract: Scheduling Tasks on Intermittently Powered Systems

Bashima Islam
bashima@cs.unc.edu
University of North Carolina at Chapel Hill

ABSTRACT

The recent development of extremely low-power computing devices and efficient energy harvesters led to the creation of computing systems that are powered by intermittently available harvested energy, e.g., solar, piezoelectric, and radio-frequency (RF). Such computing systems go through power-on and off phases due to the lack of adequate harvesting energy. These systems are known as Intermittent Computing Systems. While existing works on intermittent computing systems concentrate preliminary on the lower level goals, e.g., execution progress and memory consistency [1, 3, 5, 6, 9], the potential of such systems under timing constraints is yet to be explored. Some applications of intermittent systems with timing constraints include monitoring wildlife, health, infrastructure and environmental conditions, pedestrian safety, indoor localization and occupancy detection. In this work, we schedule tasks on intermittent systems where tasks may have timing constraints. We focus on the timely-response of intermittent systems by (1) developing unified frameworks that integrate harvesting and real-time systems, and (2) engineering machine learning algorithms for timely execution of the important portion of a task via imprecise scheduling.

KEYWORDS

Intermittent Systems, Real-Time Systems, Scheduling, Intermittent Learning, Deep Neural Network

1 SCHEDULING COMPUTATIONAL AND HARVESTING TASKS ON INTERMITTENT SYSTEMS

The sporadic nature of harvestable energy and the mutually exclusive computing and charging cycles of intermittently-powered systems pose a unique and challenging real-time scheduling problem where the existing real-time algorithms fails due to the lack of interruption in execution time. This mutual exclusion is introduced by storage-based energy harvesting system where a capacitor is used to store the harvested energy when the harvestable energy rate (supply) is less than the energy consumption rate (demand). Though other cases where harvestable energy rate is higher than consumed energy rate exists, in this work, we focus on the demand > supply as many intermittent systems follows it [3–5, 9].

To address this, we propose to schedule not only the computing cycles but also the energy harvesting cycles of an intermittent system by considering the dynamics of the environment and the pattern of harvestable energy. We propose an offline and an online scheduling algorithm for scheduling real-time tasks on intermittent systems. This offline algorithms knows the pattern of harvestable energy as a priori and first schedules the harvesting tasks. Then it schedules the computing tasks in the time-slots where no energy is available to harvest while maintaining the energy and timing constraints. Next it iteratively replaces unnecessary harvesting tasks with computing tasks without creating any energy scarcity.

Many IoT tasks demand for an online scheduling approach where the decision needs to be made on the go. In such algorithms, the harvestable energy is not known a priori. We propose an online threshold-based scheduling algorithm that avoids situations where no energy is available to harvest, and the harvested energy is not sufficient to run the system. In this algorithm, the harvestable energy us predicted at the beginning of each time-slot. First, we schedule all the computing jobs using As Late As Possible (ALAP) scheduling algorithm [8]. Next, we check the available harvestable energy at each time-slot and compute a job earlier than scheduled when the available harvestable energy is below a threshold.

Figure 1: Comparison of the performance of OEC scheduler with a system with ALAP scheduler and a system without any scheduler.

To evaluate the proposed algorithms, we conduct simulation as well as trace-based (solar and RF energy source) and real-life experiments with solar energy sources. Our results shows that our proposed offline scheduling algorithm has 92% similar performance as an optimal scheduler and our online algorithm schedules 8%–22% more jobs than earliest deadline first (EDF) [7], rate monotonic (RM) [7], and as late as possible (ALAP) scheduling algorithms. We deployed solar-powered batteryless systems where four intermittent applications are executed in the TI-MSP430FR5994 microcontroller and demonstrate that the system with our proposed online scheduler misses 63% less deadline than a non-realtime system and 8% less deadline than the system with a baseline (as late as possible) scheduler in Figure 1.

2 UTILIZING IMPRECISE COMPUTING FOR SCHEDULING TASKS ON INTERMITTENT SYSTEMS

Many real-world workloads are imprecise, i.e., they are error-tolerant and requires partial execution to achieve desired outcome. For example, deep neural network inference, where often all the layers are not needed for classification and 3D scene reconstruction, where not all stereo images are required to generate a continuous 3D scene. Though a task can be scheduled when the power is persistent it
may suffer from time scarcity when power is intermittent. However, using imprecise computing we might be able to finish the task within deadline. We illustrate this using an example in Figure 2.

We propose an energy-aware and outcome-aware soft real-time imprecise task scheduling framework for intermittent systems that executes such computationally demanding tasks. First step of our approach is to study the tasks to identify redundant computation that are not always required to be executed. We focus on (1) identifying redundant data points to avoid processing them and (2) identifying required processing of a data point to avoid unnecessary computation. We observe one example task for each type of redundancy reduction.

First, we identify unnecessary data points in a 3D scene reconstruction application by exploiting the linear and angular motion of the dual-camera system. We select the image-pairs that generate a 3D point cloud with minimal overlap with the existing point cloud required to achieve the desired quality of the combined point cloud. Next, we reduce unnecessary processing in deep neural network inference by leveraging the semantic diversity of input data and the layer-dependent expressiveness of deep features to enable early termination of a DNN inference task based on the quality of the input data. We propose a layer-aware loss function to improve the accuracy of a cluster-based, semi-supervised inference algorithm that uses an intermediate layer of a DNN as the representation of the input examples.

To enable scheduling imprecise tasks on intermittently-powered systems, we study the energy harvesting pattern and model the predictability of an energy harvester used in a specific application scenario. We observed that energy generated by harvester is bursty, i.e., the state of energy generation is maintained over a short period. This property enables us to obtain a probabilistic model of the energy harvesting pattern—which can be characterized by a single parameter, called 𝜂-factor, given for a particular harvester used in a particular application scenario. The introduction of 𝜂-factor abstracts away the unpredictability of an energy harvester and enables the development of scheduling algorithms that make informed decisions based on the predicted energy over a short period in the future. As prediction of short time intermittence, e.g., cloud coverage and occlusions can be quite difficult to predict, instead of predicting the intermittence, we determine the probability of continuous non-intermittence.

Based on these models, we propose an imprecise-computing based online scheduling algorithm that improves the deadline-aware execution of complex tasks, e.g., DNN inference, running on the intermittently-powered systems. This algorithm leverages 𝜂 of the energy source, along with the properties of the input data to adapt the execution of real-time DNN tasks.

We implement the imprecise scheduler for DNN tasks on a TI MSP430FR5994 microcontroller and evaluate its performance using two standard datasets (MNIST and ESC-10) as well as in six real-world acoustic event detection experiments with two energy sources (solar and RF). We achieved 14.09%–26.10% reduction in execution time using early termination. Moreover, the scheduler schedules 10.93%–26.92% more jobs within the deadline than the EDF scheduling algorithm. Furthermore, it gains up to 9.29% higher inference accuracy than the imprecise variant of EDF while scheduling same number of jobs.

3 CHALLENGES AND FUTURE WORKS

Time Keeping Challenges. One of the major challenges of timely execution in intermittent systems is to incorporate intermittent clocks to the system. However, the recent development of Remanence timekeepers [2] that uses RC circuits for timekeeping requirement has high resolution of up to 1 ms. These clocks can keep time up to 100s which is sufficient for intermittent systems.

Future Works. In the first part of this work, we have only looked at energy harvesting systems where the demand is greater than the supply. Next we want to observe systems where the relation between the demand and the supply is dynamic and can vary with time. Moreover, the early termination of DNN proposed in Section 2 only works for convolutional and fully connected neural networks. In future works, we want to extend this for other types of neural networks, e.g., recurrent neural networks.

REFERENCES