Big Data for Little Cores

Combining Learning and Control for Mobile Energy Efficiency

Paper # 92

Abstract

Systems designed for mobile and Internet-of-things (IoT) must deliver performance to interactive applications (that monitor the environment and respond to users) and they must conserve resources to extend battery life. Meeting these conflicting goals is complicated by two issues: (1) hardware heterogeneity leading to complicated optimization spaces and (2) dynamic operating conditions including unpredictable changes in available resources and application behavior. Machine learning techniques handle complicated optimization spaces, but do not incorporate models of system dynamics; optimal control techniques provide formal guarantees of dynamic behavior, but cannot handle non-linear system models. In this paper, we propose a combination of learning and control techniques to meet mobile system's performance requirements in unpredictable environments. Specifically, we combine a hierarchical Bayesian model (HBM) with a lightweight control system (LCS). The HBM runs on a remote server, aggregating data from many separate mobile systems to learn customized models of performance/power tradeoffs. The LCS runs locally on the mobile system and tunes resource usage to meet performance requirements efficiently. We implement both the HBM and LCS and test their ability to learn and control ARM big.LITTLE systems. Compared to existing learning and control methods, our proposed combined system delivers more reliable performance – only XX% error compared to YY% for learning and ZZ% for control – and higher energy efficiency – achieving ZZ% of optimal and improving by XX% over learning and YY% over control. When used in a multi-application scenario these numbers improve: xx% error compared to yy% for learning and zz% for control and improvements of xx% and zz%, respectively, in energy efficiency.

1. Introducation

Mobile systems have clear requirements for correct operation: they must meet performance goals necessary for interacting with sensors and human users, but must also conserve energy to maximize battery life. To address these often conflicting requirements, hardware platforms have become increasingly diverse and complicated. Many such processors support, for example, different core types with different performance/power tradeoffs, which can be operated at a wide range of different speeds. Meeting performance requirements is further complicated by the dynamic nature of computing systems: application demands can vary widely as a function of input or application phase and multiple applications may compete for resources.

Thus, two central challenges arise to meeting performance requirements with minimal energy on mobile systems: (1) complexity and (2) dynamics. Each challenge has been addressed individually. First, machine learning approaches can identify the complicated, power/performance tradeoff spaces that arise on configurable, heterogeneous mobile systems []. Such approaches can learn complex models, identifying and avoiding local extrema that lead to inefficient resource usage. Second, optimal control theoretic techniques ensure performance is met with minimal energy by tuning resource allocation as applications run []. Control techniques provide a formal basis for reasoning about dynamics and can ensure performance requirements are met despite application, input, or workload fluctuations.

Learning and control techniques have complementary strengths and weaknesses. Learning approaches handle complexity, but have no established mechanism for managing system dynamics; more powerful learning methods also tend to incur higher computational cost making them ill-suited for runtime use on energy-constrained devices. Control approaches handle dynamics, but rely on linear models that are increasingly insufficient to capture the diversity of modern hardware.

We therefore propose combining learning and control techniques to manage both complexity and dynamics. Specifically, we use a hierarchical Bayesian model (HBM) to aggregate data across multiple devices and applications, creating accurate, customized models mapping application resource usage to performance and power. To mitigate overheads, the HBM runs on a remote server, allowing it to quickly collect data from a number of applications running on separate devices and amortize the cost of computing the models across all those devices. Once learned, the models are sent to individual devices where a lightweight control system (LCS) uses them to ensure that performance requirements are met with minimum energy even if the application changes phase, processes an unexpected input, or runs with other applications competing for resources. The control system is computationally efficient and provides formal guarantees that it will converge to the desired performance despite unpredictable system dynamics. These guarantees are a product of the combined learning and control framework. The accuracy of the learner in the face of system complexity is essential for guaranteeing performance in highly dynamic systems **TODO: 1**.

While control and learning frameworks exist, the key to combining them is creating an interface between the learning and control systems. Specifically, learning frameworks for resource management map configurations (*e.g.*,resource allocations) into estimated performance and power. These mappings are discrete and non-linear, capturing the behavior of the underlying system. Controllers, in contrast, work with continuous linear models. Therefore, our proposed combination of learning and control requires an interface to convert the discrete non-linear learned models into continuous linear models. We address this challenge by forming the lower convex hull of points on the learned power/performance tradeoff space. Interpolating between these points gives us a piecewise linear function that is appropriate for control models, yet still captures the significant behavior of the underlying system. This interface allows us to combine the approaches studied in this paper, and we believe it is sufficiently general to apply to other combinations of learning and control as well.

To evaluate our approach, we implement the HBM in Matlab and run it on an x86 server. We implement the LCS in C and evaluate it on ODROID XU3 development boards featuring Samsung Exynos 5 Octa processors based on the ARM big.LITTLE architecture. We run 20 different benchmarks to test the HBM's ability to learn application specific models and the LCS's ability to deliver performance with near minimal energy consumption. We compare to published learning and control methods in a variety of settings. While many applications have inherent dynamics (*i.e.*,different processing phases), we explicitly test the ability to adapt to the unknown by running each application with other, random applications. We evaluate both the ability to deliver requested performance and the energy efficiency and find that the proposed approach:

- **Delivers Better Performance:** We quantify the ability to meet performance goals by calculating the error between the desired and delivered performance. In a single application setting, our approach achieves an average error of XXX%, compared to YYY% for existing learning methods and ZZZ% for existing control approaches (See Section **??**). In a multi-application setting, our approach avchieves an average error of YYY%, compared to WWW% and ZZZ% (See Section **??**).
- **Requires Lower Energy on Average:**
- **Performs Far Better in the Worst Case:**
- **Successfully Adapts to Dynamics: TODO: 1**

In summary, this paper makes the following contributions:

- Proposing the combination of a hierarchical Bayesian learning with a lightweight control system to meet the twin challenges of addressing complexity and dynamics to deliver performance with minimal energy on mobile systems.
- Demonstrating and implementing an interface for combining discrete learned models of resource usage with continuous control models of resource dynamics.
- Evaluating the implementation and comparison to existing, independent learning and control techniques.