Predictive Thermal Management for Energy-efficient Execution of Concurrent Applications on Heterogeneous Multi-cores

INTRODUCTION

**Eduardo Wachter - e.weber-wachter@soton.ac.uk**

Why?
- Avoid Thermal peaks, Thermal hotspots and Thermal cycling → **more reliability**
- Temperature also influences Power → **Potential power savings**

Contributions:
- An **accurate temperature prediction** algorithm for any frequency set of the system.
- A **Runtime manager** that predicts the temperature for the next period while avoiding temperature thresholds

METHODOLOGY

- A training data set to **classify the best regression** model is created offline
- Log the temperature, frequency and power consumption for memory and big cluster. The data is collected on the big and LITTLE cluster of the chip
    - Data collected at a rate of 1 Hz will later be used for 1 Hz temperature prediction. Temperature will be predicted for the next second
- The regression model that provides **minimum error on predicting temperature** is then used to feed the **runtime manager (RTM)**
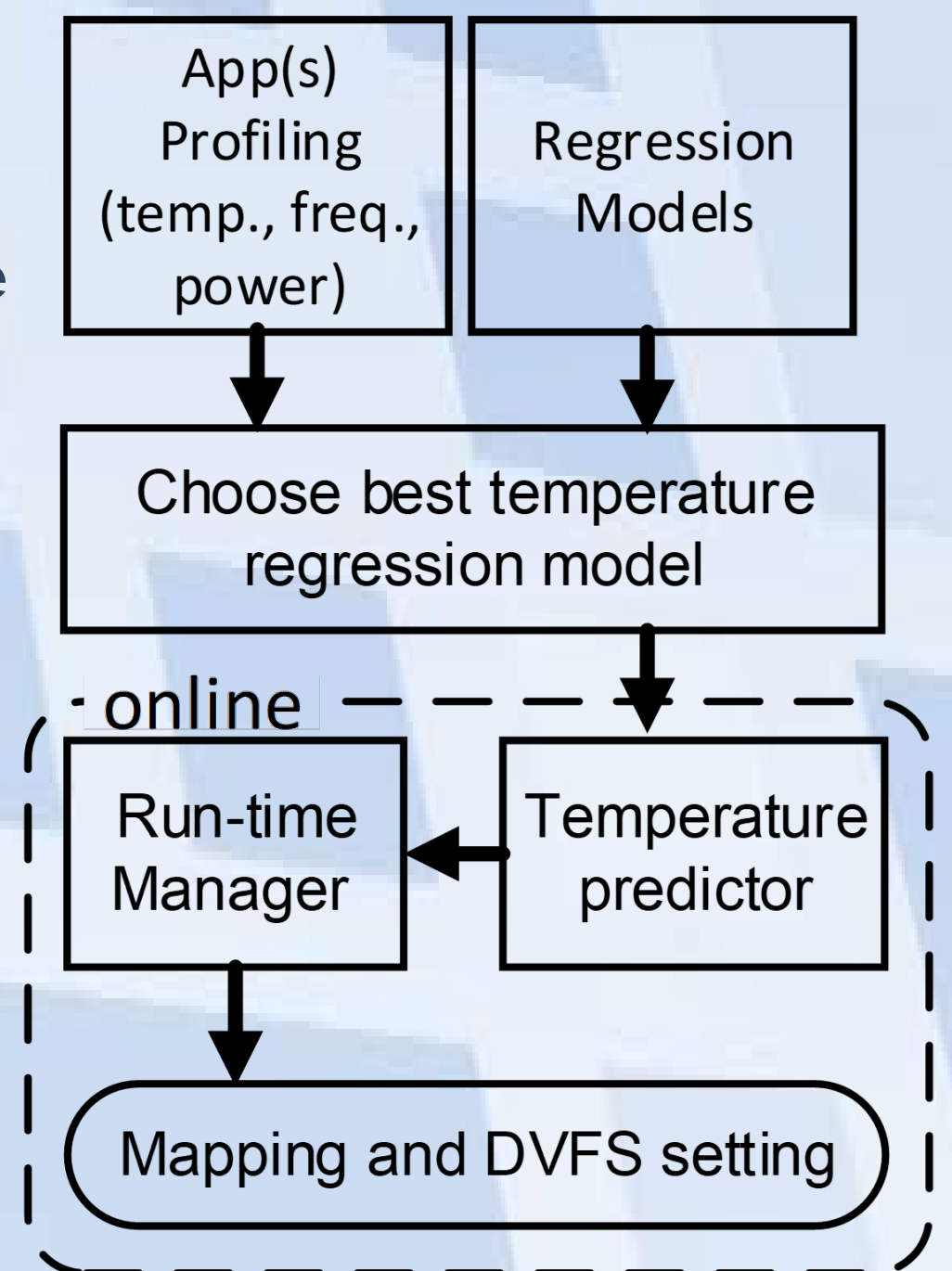
PROPOSAL

- The runtime manager uses the temperature predictor when setting the frequency and mapping of the tasks
- Temperature predictor Regression Model is based on power and temperature measurements and predictions:

$$\tilde{T}(t) = constant + \alpha T(t-1) + \beta T(t-2) + \gamma \tilde{P}_t^{big} + \theta P_t^{mem}$$

~ means prediction; (t-1) is measured past interval; $\alpha, \beta, \gamma, \theta$ and constant are calculated by the regression model

- **Runtime Manager** chooses a given frequency:
    - Temperature is predicted for this frequency
    - If temperature is greater than threshold, sets a lower frequency to satisfy requirements

Flowchart:
- App(s) Profiling (temp., freq., power)
- Regression Models
- Choose best temperature regression model
- online: Run-time Manager ← Temperature predictor
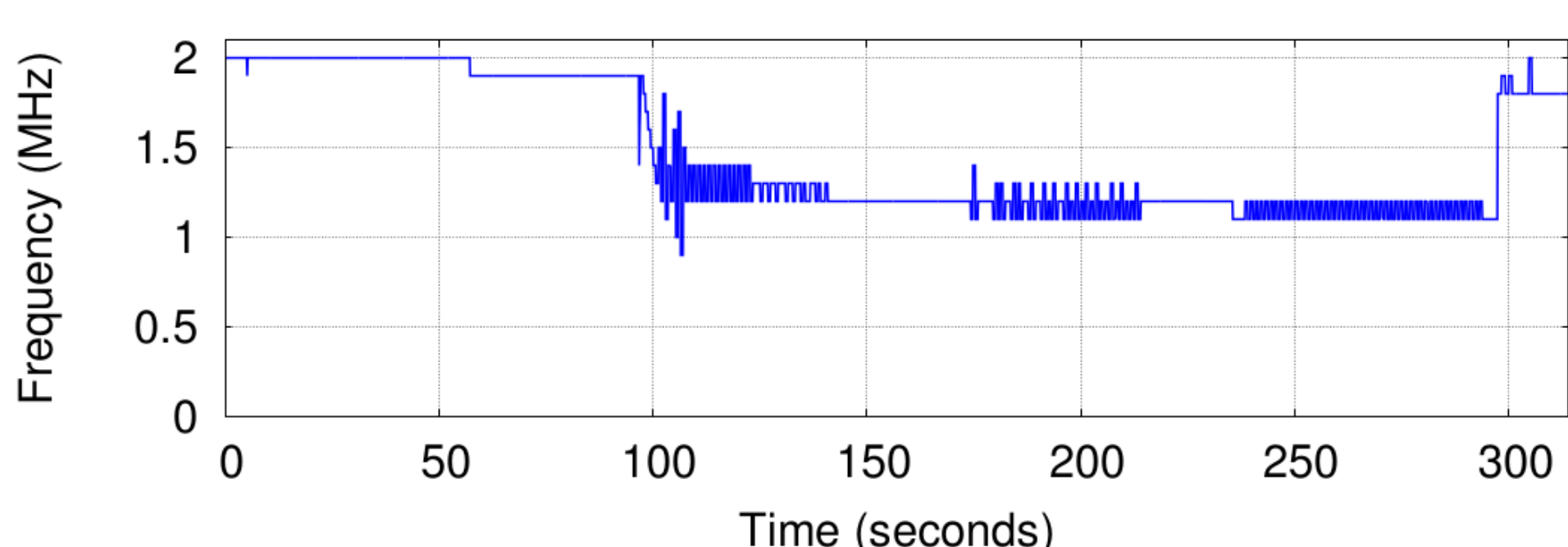- Mapping and DVFS setting

RESULTS

- Validation on the ODROID-XU3 using PARSEC and SPLASH benchmarks
- Regression model **average error** is 1.19° C on runtime frequency changing
- One, two and three applications are executing concurrently
- APPs mapped according past metric
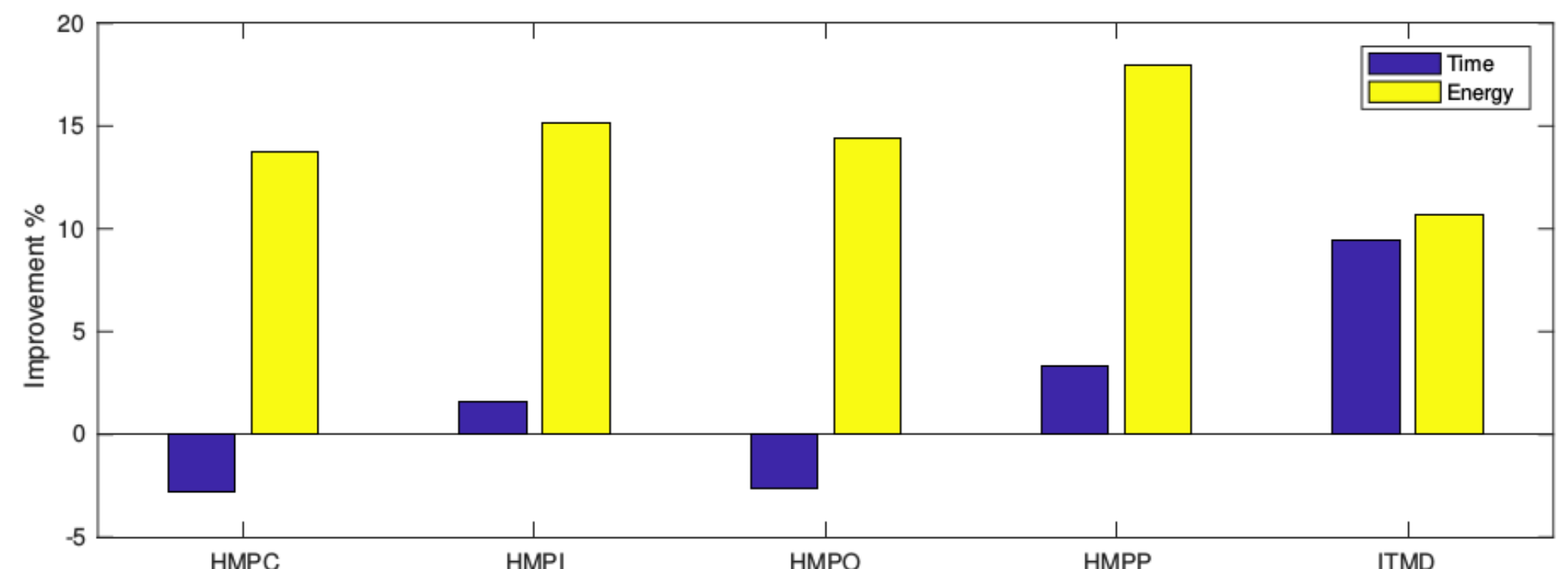- Temperature threshold is set to 90° C

without prediction
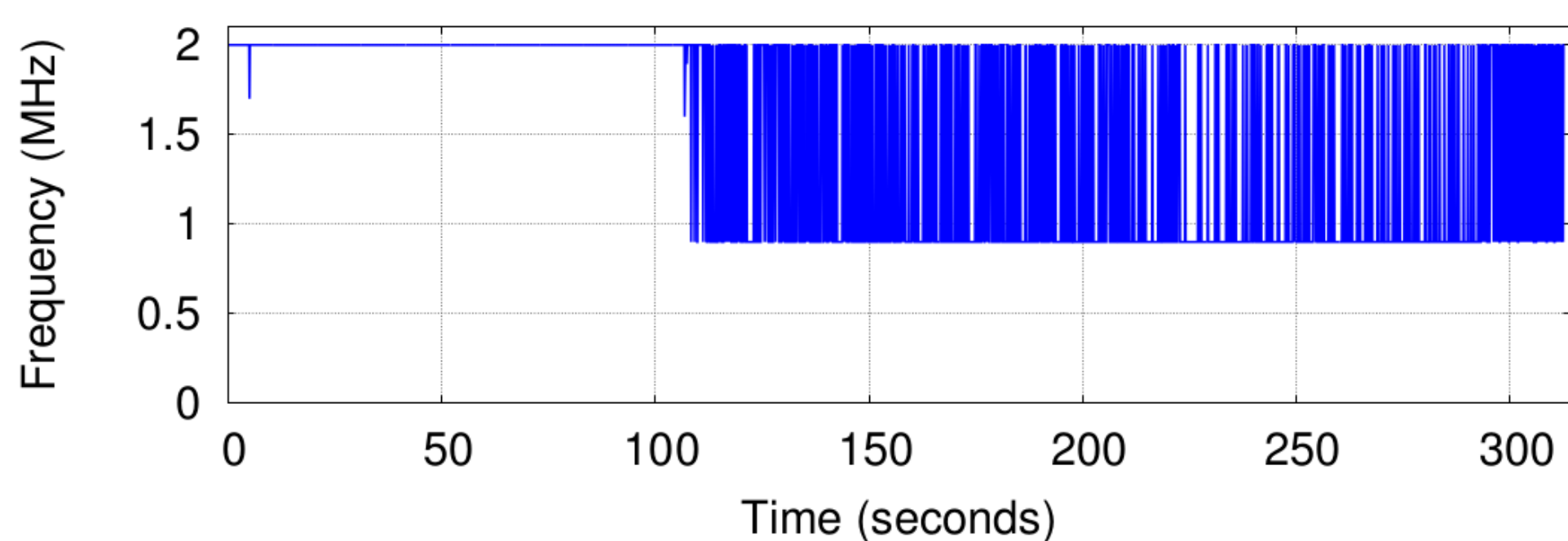
with prediction

Improvement for 2 concurrent apps of PARSEC

Improvement for 3 concurrent apps of PARSEC