


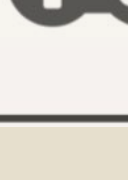


IoT Sensor Gym: Training Autonomous IoT Devices with Deep Reinforcement Learning

Abdulmajid Murad,¹ Frank Alexander Kraemer,¹ Kerstin Bach,¹ Gavin Tylor²

¹Norwegian University of Science and Technology, NTNU, Norway. ²US Naval Academy, USA

Games	→	IoT
Player 	→	IoT Device Agent
World 	→	Weather and other relevant conditions.
Reward 	→	Data measurements
Control 	→	Duty cycle, sensing frequency,...

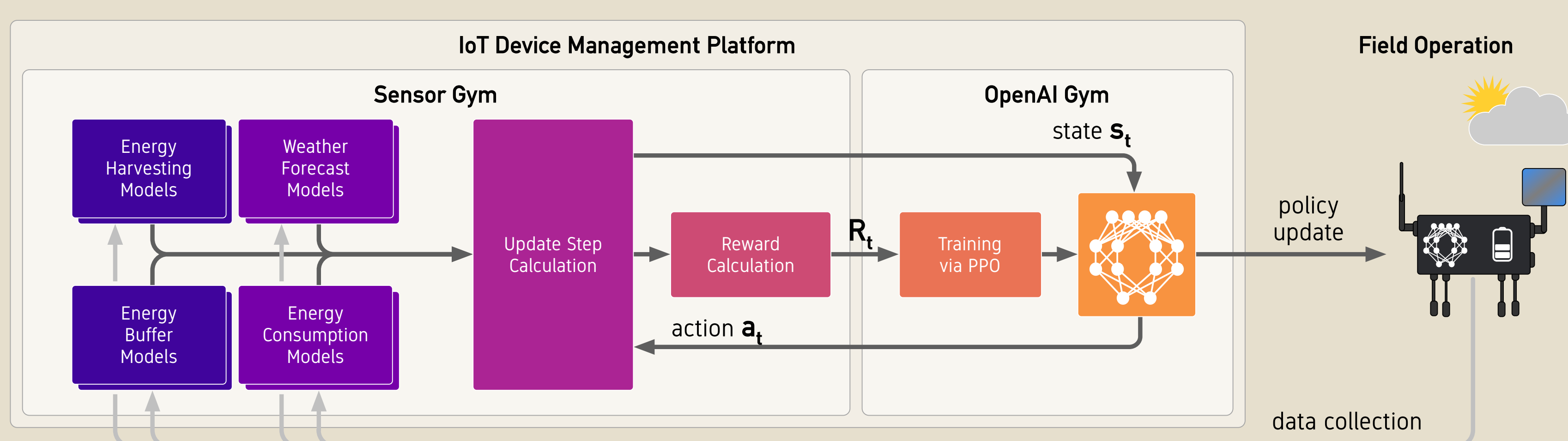


Left: We apply deep reinforcement learning to the domain of IoT, inspired by the analogy between gaming and IoT. **Right:** IoT sensor devices in changing weather and climate conditions.

Controlling IoT Sensors Like Games

Current IoT solutions often rely on manual, static configurations or fined-tuned algorithms that fail to suit all nodes of large-scale IoT systems.

Deep reinforcement learning proved to be effective in training autonomous agents, for instance to play games. Agents are able to make complex decisions and learn in non-stationary environments. This enables scalable solutions for heterogeneous devices to act optimally in dynamic and non-stationary environments.

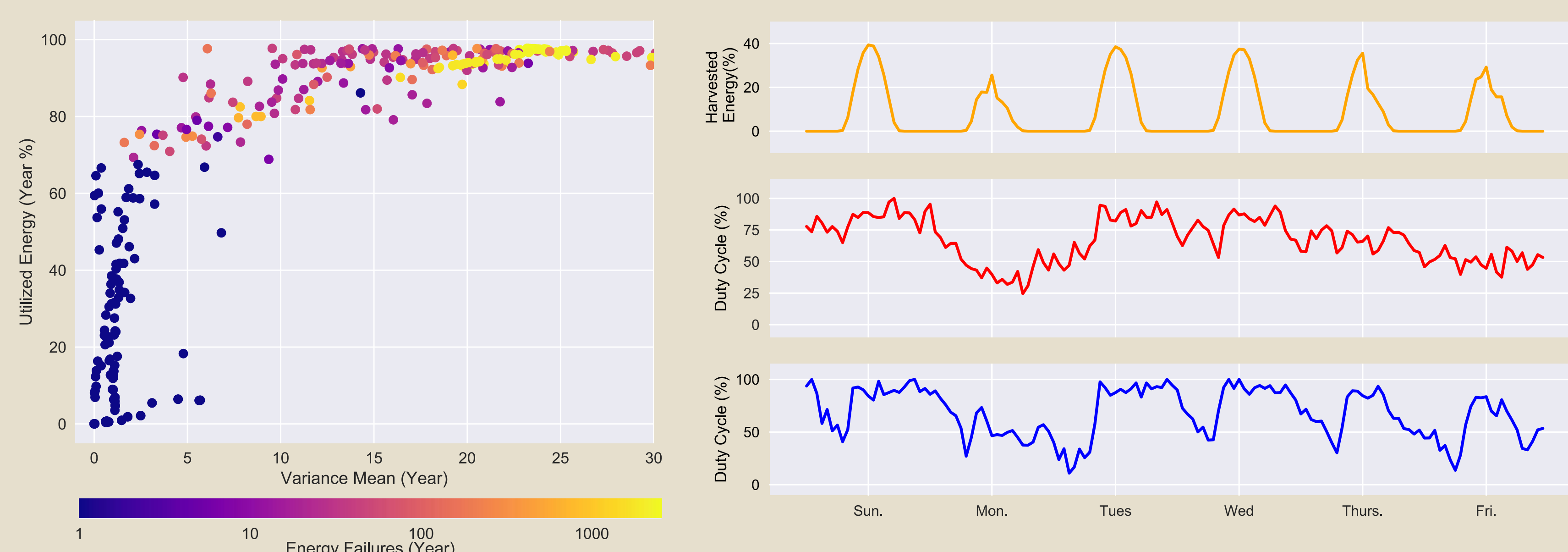


The architecture of the IoT Sensor Gym framework and the process of training and deploying RL agents into sensor devices.

A Framework for Training IoT Devices

We built the IoT Sensor Gym as an extension to the OpenAI Gym framework. Sensor Gym provides an environment specific to constrained IoT devices, with an emphasis on their energy budget.

IoT devices are simulated using a variety of models which can be combined and configured to match various use cases.



Left: Performance of more than 300 different agents over a whole year. The agents have different hyper-parameters that model the tradeoff between application goals.

Right: Solar energy intake and resulting duty cycle of two selected agents over six days.

Example: Duty-Cycle Optimization

We trained RL agents using the PPO algorithm to control solar-powered IoT nodes autonomously [1] and designed a reward function reflecting the application goals of an IoT system:

- Utilizing as much of the incoming solar energy as possible.
- Operating without failing by depleting its buffer.
- Having a duty cycle with low variance to ensure steady data collection.

References

[1] A. Murad, F.A. Kraemer, K. Bach, and G. Taylor. Autonomous Management of Energy-Harvesting IoT Nodes Using Deep Reinforcement Learning. In SASO 2019, IEEE.

[2] IoT Sensor Gym source code. <https://github.com/Abdulmajid-Murad/IoT-Sensor-Gym>.