Scalable Load Management using Reinforcement Learning

Oak Ridge National Laboratory (ORNL)
Helia Zandi, R&D Staff
zandih@ornl.gov
Project Summary

Timeline:
Start date: 10/9/2018
Planned end date: 9/30/2021

Key Milestones
1. Develop, formulate, and test the RL algorithm in simulation; 9/30/2019
2. Test the scalability of the developed load management system; 9/30/2020
3. Compare RL performance with a golden standard optimization technique; 6/30/2021

Budget:
Total Project $ to Date:
• DOE: $1,824 K
• Cost Share: $0
Total Project $:
• DOE: $2,100 K
• Cost Share: $0

Key Partners:
| Southern Company | University of Tennessee, Knoxville |

Project Outcome:
- A scalable load management system that can be deployed by utilities on grid existing infrastructure
- A learning-based optimization algorithm that can be applied to existing homes and new construction with minimal effort and minimal additional devices
Challenge: The Effect of Emerging Technologies on the Electric Grid

- Over **75 billion** connected devices predicted to be in use by 2025
- There is an immense opportunity for a management system which can **control and coordinate the power use** of these devices
- **41%** of the energy consumption in the United States is from buildings
- Advanced sensing and controls have the potential to save energy in buildings up to **40%**

What does the electric grid need?
Approach: Project Overview

Design, develop, and field evaluate a scalable and cost-effective load management system using Reinforcement Learning (RL)

Project Objectives

**Objective 1:** Develop Reinforcement Learning-based optimization and control methods for understanding energy use patterns and for load scheduling

**Objective 2:** Develop a scalable load management system to access flexibility in loads

**Objective 3:** Perform field validation of the software framework and demonstrate benefits of running RL-based optimization and control in residential buildings
Load Management System Software Architecture

- Hierarchical cloud-based multi-agent system (MAS) architecture
- Critical data are identified
Approach: Control Development Approach

Model Development
1. Single-zone single-family house
2. Two-zone single-family house
3. Single-stage HVAC
4. 2-zone HVAC
5. Water heater

Simulation Testbed
1. Single-zone building, single-stage HVAC simulation testbed
2. 2-zone building, 2-zone HVAC simulation testbed

Algorithm Development
Developed and evaluate two different model-free deep reinforcement learning approach:
1. Deep Q-Network (DQN)
2. Deep Deterministic Policy Gradient (DDPG)
Approach: Reinforcement Learning

- Reinforcement learning (RL) method is a type of machine learning method that optimizes the decision-making strategy of an agent within an unknown environment.

- RL uses a Markov Decision Process (MDP).

- There are mainly two types of RL method:
  - Value-based RL method: estimates Q-value of a state-action pair
  - Policy-based RL method: generate probability of all feasible action for current state

---

**Value-based RL method**: estimates Q-value of a state-action pair.

**Policy-based RL method**: generate probability of all feasible action for current state.

---


- Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Evan McKee, Yan Du, Bori Cui, Teja Kuruganti, Helia Zandi “Evaluating the Adaptability of Reinforcement Learning based HVAC Control for Residential Houses” MDPI Sustainability as part of the Special Issue Building and Urban Energy Prediction-Big Data Analysis and Sustainable 2020, [https://doi.org/10.3390/su12187727](https://doi.org/10.3390/su12187727)
Approach: HVAC Optimization for 2-Zone Single Family Building

Building model

Neural network-based Q-learning

RL Model Development Based on DQN

Progress: Validation-1: with the Same House as a Simulation

- 74°F (23.33°C) - 75°F (23.89°C) - 73°F (22.77°C)
- 72°F (22.22°C) - 73°F (22.77°C) - 71°F (21.67°C)
- 75°F (23.89°C) - Comfort violation - 73°F (22.77°C)
- 73°F (22.77°C) - 71°F (21.67°C)

- 74°F (23.33°C) - 75°F (23.89°C) - 73°F (22.77°C)
- 72°F (22.22°C) - 73°F (22.77°C) - 71°F (21.67°C)
- 75°F (23.89°C) - Comfort violation - 73°F (22.77°C)
- 73°F (22.77°C) - 71°F (21.67°C)
Day wise cost comparison of operating HVAC with cooling set point (baseline) and pre-trained RL model

<table>
<thead>
<tr>
<th>Cooling Set Point</th>
<th>73 °F</th>
<th>74 °F</th>
<th>75 °F</th>
<th>(74, 75, 73) °F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (kWh)</td>
<td>368.01</td>
<td>333.49</td>
<td>301.37</td>
<td>334.84</td>
</tr>
<tr>
<td>Cost ($)</td>
<td>56.94</td>
<td>51.91</td>
<td>46.74</td>
<td>51.69</td>
</tr>
<tr>
<td>MOC (min)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Baseline</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (kWh)</td>
<td>430.89</td>
<td>390.96</td>
<td>352.72</td>
<td>393.97</td>
</tr>
<tr>
<td>Cost ($)</td>
<td>40.10</td>
<td>35.46</td>
<td>31.58</td>
<td>36.09</td>
</tr>
<tr>
<td>MOC (min)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>79</td>
</tr>
</tbody>
</table>

| % Cost reduction  | 29.57%| 31.68%| 32.43%| 30.17%          |

Comparison of the cost of operation, energy consumption and MOC for a fixed cooling setpoint baseline and pre-trained
Adaptability Testing

Generating unique building models
Validating with synthetic houses

Trained Model

Randomized parameters

Unique Model

Trained parameters

Uniform distribution ±20% of trained value

74°F (23.33°C)
72°F (22.22°C)
73°F (22.77°C)
75°F (23.89°C)
71°F (21.67°C)
73°F (22.77°C)
75°F (23.89°C)
71°F (21.67°C)

Outside Temperature
Indoor Air Temperature
Cool on
schedule LT
schedule UT

Pre-cooling events

Minutes
28000
29000
30000
31000
32000

Temperature (degC)

Comfort violation

75°F (23.89°C)
73°F (22.77°C)
73°F (22.77°C)
71°F (21.67°C)

Validation with synthetic house and real house

Validation with synthetic house

- Deployed with Model Pre-Trained on the:
  - Deployed House Data
  - Synthetic house data
- Demonstrated an average saving of 20%

Validation with real house

<table>
<thead>
<tr>
<th>Cooling Set Point</th>
<th>73 °F</th>
<th>74 °F</th>
<th>75 °F</th>
<th>(74, 75, 73) °F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy (kWh)</td>
<td>330.15</td>
<td>297.72</td>
<td>269.20</td>
<td>299.61</td>
</tr>
<tr>
<td>Cost ($)</td>
<td>50.94</td>
<td>46.00</td>
<td>41.81</td>
<td>45.98</td>
</tr>
<tr>
<td>MOC (min)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>97</td>
</tr>
</tbody>
</table>

| **RL**            |       |       |       |                 |
| Energy (kWh)      | 390.75| 352.36| 317.71| 358.73          |
| Cost ($)          | 35.17 | 30.94 | 27.41 | 31.57           |
| MOC (min)         | 0     | 0     | 0     | 69              |

% Cost reduction:
- Baseline: 30.9%
- RL: 32.72%
- EnergySaving: 34.44%
- CostSaving: 31.33%
**Approach: WH Optimization**

**Focus:** Control a hybrid water heater using demand response (DR) commands for minimum electricity cost under a time-of-use (TOU) electricity pricing policy.

---

**State (s)**
- Water temperatures
- Electricity prices
- Hot water usages

**Action (a)**
- Load up
- Normal operation
- Shed

**Reward (r)**
- (-)Electricity cost

---

Progress: RL Training

RL agent #1 (RL #1)
- node temperatures
- Electricity prices of next 30 minutes

RL agent #2 (RL #2)
- node temperatures
- Electricity prices of next 30 minutes
- Hot water usage volumes of next 30 minutes

RL agent #3 (RL #3)
- node temperatures
- Electricity prices of next hour
- Hot water usage volumes of next hour

RL agent #4 (RL #4)
- node temperatures
- Electricity prices of next two hours
- Hot water usage volumes of next two hours

Operation strategy | Look Ahead | Electricity Cost | Element Usage | Average COP<sub>HP</sub>
--- | --- | --- | --- | ---
Baseline | N/A | $1.36 | 75 min | 4.88
RL agent #2 | 30 min | $1.21 | 89 min | 5.25
RL agent #3 | 1 hour | $1.03 | 82 min | 5.47
RL agent #4 | 2 hours | $0.91 | 53 min | 5.15
MPC #2 | 30 min | $1.31 | 157 min | 5.76
MPC #3 | 1 hour | $1.04 | 104 min | 5.70
MPC #4 | 2 hours | $0.94 | 84 min | 5.61
Day-ahead* optimization | 5 days | $0.81 | 47 min | 5.48
Baseline vs RL

Baseline

RL
## Progress: WH Deployment Results

<table>
<thead>
<tr>
<th>Operation strategy</th>
<th>March 27</th>
<th>March 28</th>
<th>March 29</th>
<th>March 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>RL</td>
<td>$0.07</td>
<td>$0.14</td>
<td>$0.09</td>
<td>$0.17</td>
</tr>
<tr>
<td>Baseline</td>
<td>$0.11</td>
<td>$0.23</td>
<td>$0.11</td>
<td>$0.26</td>
</tr>
</tbody>
</table>

### Cost
- In all days, the RL cost less than the baseline
- RL saved in the range of 18% to 39%
- RL achieved more savings when hot water demand is higher

### Comfort
- The upper tank is reserved for users by the manufacturer and are not available for any DR command
- In both cases, the upper tank temperature never went below the comfort range
Impact

- Developed and demonstrated a learning-based load management system
  - Addressing the need for a scalable Grid-interactive Efficient Buildings (GEB) platform
  - Delivering towards BTO strategy for GEB - seamless connectivity between different devices and utilization of optimization and learning algorithm for optimal scheduling of residential loads.

- The data-driven RL algorithm addresses the need for:
  - Control algorithms that are self-aware and self-calibrating
  - Improve energy efficiency, reducing peak demand, and improving comfort.

- This project contributes towards the BTO emerging technology goal of reducing U.S. building portfolio’s carbon footprint in half by 2035

- RL-based algorithm demonstrated an average demand cost savings of 25% while maintaining the occupant comfort
Stakeholder Engagement

- Weekly meetings
- ORNL has presented the project, findings, and lessons learned to other national laboratories, professional societies, workshops, conferences, seminars, industry representatives.
- Team members are active in professional societies
- Published 14 Journal and conference papers, 1 journal papers under review, and more in process.
- ORNL submits Quarterly Progress Report (QPR) to DOE
Remaining Project Work

- Improve the performance of the developed algorithms for both WH and HVAC optimization
- Enhance the RL implementation with deploying RL from scratch
- Draft a technical report which contains the results of the system integration, data collection, data analysis, algorithms design, lessons learned, and field testing
Thank you

ORNL
Helia Zandi, R&D Staff
zandih@ornl.gov

ORNL’s Building Technologies Research and Integration Center (BTRIC) has supported DOE BTO since 1993. BTRIC is comprised of 50,000+ ft² of lab facilities conducting RD&D to support the DOE mission to equitably transition America to a carbon pollution-free electricity sector by 2035 and carbon free economy by 2050.

Scientific and Economic Results
238 publications in FY20
125 industry partners
27 university partners
10 R&D 100 awards
42 active CRADAs

BTRIC is a DOE-Designated National User Facility
REFERENCE SLIDES
List of Publications

- for HVAC Energy Control", IEEE BigData20 Industry & Government
List of Publications


- Kuldeep Kurte, Jeffrey Munk, Kadir Amasyali, Olivera Kotevska, Robert Smith, Helia Zandi “Electricity aware deep reinforcement learning based intelligent HVAC control” ACM 1st International Workshop on Reinforcement Learning for Energ, https://doi.org/10.1145/3427773.3427866


- Olivera Kotevska, Jeffrey Munk, Kuldeep Kurte, Kadir Amasyali, Robert Smith, Helia Zandi "Methodology for Interpretable Reinforcement Learning Model
Project Budget

**Project Budget:** $700K (FY19), $700K (FY20), $700K (FY21)

**Variances:** N/A

**Cost to Date:** $1,824K

**Additional Funding:** None

### Budget History

<table>
<thead>
<tr>
<th></th>
<th>FY2019– FY 2020 (past)</th>
<th>FY 2021 (current)</th>
<th>FY 2022 (planned)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOE</strong></td>
<td>$1,400K</td>
<td>$700K</td>
<td>$0</td>
</tr>
<tr>
<td><strong>Cost-share</strong></td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
</tbody>
</table>


## Project Plan and Schedule

### Project Schedule

<table>
<thead>
<tr>
<th>Task</th>
<th>Q1 (Oct-Dec)</th>
<th>Q2 (Jan-Mar)</th>
<th>Q3 (Apr-Jun)</th>
<th>Q4 (Jul-Sep)</th>
<th>Q1 (Oct-Dec)</th>
<th>Q2 (Jan-Mar)</th>
<th>Q3 (Apr-Jun)</th>
<th>Q4 (Jul-Sep)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1 Milestone: Draft software architecture specification for the load management system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2 Milestone: Develop field evaluation plan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3 Milestone: Implement software application for load management system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4 Milestone: Define and formulate the initial RL algorithm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1 Milestone: Initial field validation at Yarnell Station</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2 Milestone: Try various RL strategies and algorithm improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3 Milestone: Draft a data collection &amp; analysis report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4 Milestone: Scalability testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1 Milestone: Improve RL algorithm performance &amp; its cost savings and comfort in simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2 Milestone: Field test the improved methodology at Yarnell station house</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3 Milestone: Draft a data collection &amp; analysis report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Completed Work

- **FY2019**
  - Q1 Milestone: Initial field validation at Yarnell Station
  - Q2 Milestone: Try various RL strategies and algorithm improvement
  - Q3 Milestone: Draft a data collection & analysis report
  - Q4 Milestone: Scalability testing
- **FY2020**
  - Q1 Milestone: Improve RL algorithm performance & its cost savings and comfort in simulation
  - Q2 Milestone: Field test the improved methodology at Yarnell station house
  - Q3 Milestone: Draft a data collection & analysis report
- **FY2021**
  - Q4 Milestone: Improve RL performance and draft a technical report

### Active Task (in progress work)

- **Milestone/Deliverable (Originally Planned) use for missed**
- **Milestone/Deliverable (Actual) use when met on time**