Malwal Shreyas Ravindra, 2026, 14:5
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Reinforcement Learning For Energy Management In Smart Homes: A Review Of Algorithms, Environments And Iot-Integrated Control

Malwal Shreyas Ravindra[1]Prof. R. V. Varhade[2]Prof. M. C. Kolse[3]Prof. S. K. Said[4]
Department of Artificial Intelligence and Data Science
Jaihind College of Engineering
Pune, Maharashtra, India

Abstract- Rising electricity demand and the spread of smart home technology create a need for intelligent, sustainable residential energy management. Traditional approaches rely on fixed schedules and predefined rules that respond poorly to changing prices, demand, weather and user requirements. Reinforcement Learning (RL) offers a different approach in which an agent learns optimal control decisions by interacting with the home environment and receiving feedback. This paper reviews RL for smart-home energy management with the goals of reducing consumption and cost while maintaining comfort. It introduces Markov decision processes and the main RL algorithms, from tabular Q-learning to deep Q-networks, proximal policy optimization and soft actor-critic, then surveys applications to demand response, appliance scheduling, heating, ventilation and air conditioning control, electric-vehicle charging, rooftop solar and battery storage. Simulation environments for training and benchmarking agents are discussed. A five-layer architecture integrating IoT sensing and smart meters, state construction with forecasting, an RL decision engine, device control and a user interface with comfort preferences is proposed, together with research gaps, objectives, a comparison of algorithms and challenges such as safety, sample efficiency, comfort modelling and transfer between homes. The review concludes that RL can adaptively optimize household energy use when combined with safety constraints and user feedback.

***Index Terms*—Reinforcement learning, smart home, energy management, demand response, deep Q-network, HVAC control, EV charging, IoT**

I. INTRODUCTION

Household electricity demand is growing with air conditioners, water heaters, electric vehicles and appliances, and peak demand stresses the grid. Time-of-use tariffs, rooftop solar and home batteries give households the opportunity to save money by shifting consumption, but managing all these decisions manually is impractical.

Rule-based home automation, such as "run the washing machine at 2 pm", cannot adapt to changing prices, weather and occupant behaviour. Reinforcement learning, described by Sutton and Barto [1], enables an agent to learn control policies from experience by maximizing long-term reward. Vázquez-Canteli and Nagy [2] reviewed RL for demand response and found it increasingly used for building energy control.

This paper reviews RL algorithms and applications in smart-home energy management and proposes an IoT-integrated architecture.

II. RELATED WORK

Mohsenian-Rad *et al.* [3] proposed game-theoretic demand-side management to schedule household consumption and reduce peak-to-average ratios, an optimization baseline against which RL methods are compared.

Wei *et al.* [4] applied deep RL to building HVAC control and reduced energy cost while keeping temperatures within comfort bounds. Mocanu *et al.* [5] used deep policy-gradient and Q-learning methods for online building energy optimization at the level of individual devices, showing substantial cost savings.

Wan *et al.* [6] developed model-free deep RL for real-time electric-vehicle charging scheduling under

© 2026Malwal Shreyas Ravindra. his is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

Sr. No.	Ref.	Year	Methodology	Findings	Research Gap
1	[3]	2010	Game-theoretic demand-side management	Reduced peak-to-average ratio	Needs known demand
2	[9]	2015	Deep Q-network	Human-level control from experience	Discrete actions
3	[4]	2017	Deep RL for HVAC	Lower cost within comfort bounds	Simulated building
4	[5]	2019	Online deep RL building optimization	Device-level cost savings	Simulation-based
5	[6]	2019	Model-free RL EV charging	Lower charging cost	Single vehicle
6	[2]	2019	Review of RL for demand response	Mapped algorithms and models	Review only
7	[8]	2019	CityLearn environment	Standard RL benchmark	Simplified models
8	[7]	2020	Deep RL in power systems overview	Broad application survey	Few real deployments

uncertain prices and arrival times, minimizing charging cost while meeting energy needs. Zhang *et al.* [7] reviewed deep RL across power system applications, including demand response and energy management.

For benchmarking, Vázquez-Canteli *et al.* [8] released CityLearn, an OpenAI Gym environment for demand response in groups of buildings, enabling fair comparison of RL agents. Table I summarizes these works.

TABLE I Summary of Literature Survey

III. PRELIMINARIES

1. Markov Decision Processes

Energy management is modelled as a Markov decision process: the state includes indoor temperature, appliance status, battery level, solar output and price; actions switch or schedule devices; and the reward combines cost, comfort and energy use. Fig. 1 summarizes controllable resources and RL algorithm families.

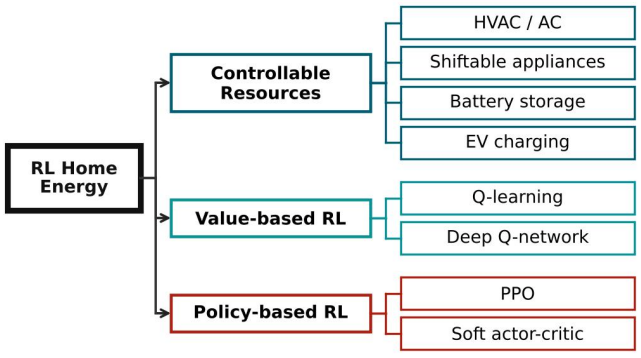


Fig. 1. Controllable resources and RL algorithm families in smart homes.

2. Algorithm Families

Value-based methods such as Q-learning and deep Q-networks [9] learn the value of actions; policy-gradient methods such as PPO [10] optimize policies directly; and actor-critic methods such as soft actor-critic [11] handle continuous actions such as thermostat set-points.

IV. PROBLEM STATEMENT AND OBJECTIVES

1. Identified Research Gaps

G1: RL agents need many interactions to learn, which is impractical in a real home.

G2: Safety and comfort constraints are not always guaranteed during learning.

G3: Policies trained for one home rarely transfer to homes with different devices and occupants.

G4: Occupant preferences and behaviour are hard to model.

G5: Indian tariff structures and appliance mixes are rarely studied.

2. Objectives Of The Review

To formulate smart-home energy management as a reinforcement learning problem.

To review RL algorithms and their applications to household devices.

To propose an IoT-integrated RL energy management architecture.

To compare RL algorithms with rule-based and optimization approaches.

To identify challenges and future research directions.

3. Proposed Framework

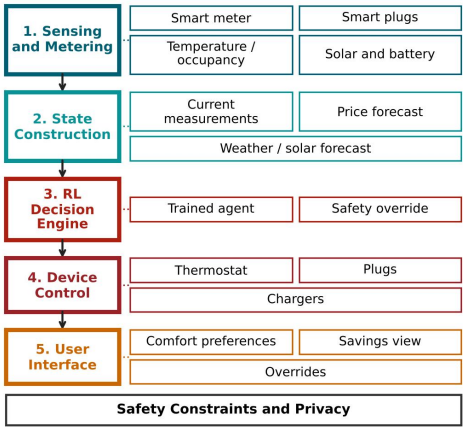


Fig. 2 shows the proposed architecture.

Fig. 2. Proposed IoT-integrated RL energy management architecture.

4. Sensing And Metering Layer

Smart meters, smart plugs, temperature and occupancy sensors, solar inverters and battery systems report data through the home gateway.

5. State Construction Layer

Measurements are combined with forecasts of price, weather and solar generation to form the agent's state.

6. Rl Decision Engine

A trained agent selects actions every few minutes, such as adjusting the AC set-point, delaying the dishwasher, charging the battery or EV. A safety layer overrides actions that would violate comfort or device limits.

7. Device Control Layer

Commands are sent to smart plugs, thermostats and chargers via standard protocols.

8. User Interface Layer

Residents set comfort preferences and priorities, view savings and override decisions; overrides are used as feedback.

V. WORKING METHODOLOGY

Fig. 3 outlines the control loop. The agent learns to maximize expected discounted reward; for Q-learning, the update is

$$Q(s,a) \leftarrow Q(s,a) + \alpha\,[\, r + \gamma \max_{a'} Q(s',a') - Q(s,a)\,] \quad (1)$$

and the reward balances cost and comfort,

$$r = -(\, p_t \cdot E_t + \beta \cdot D_comfort\,) \quad (2)$$

where p_t is the price, E_t the energy used and D_comfort the deviation from preferred conditions. Agents are pre-trained in simulation and fine-tuned in the home.

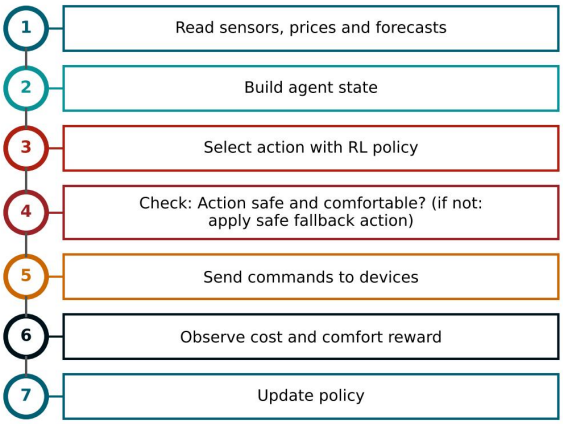


Fig. 3. RL-based home energy control loop.

VI. DISCUSSION

1. Comparison Of Approaches

Table II compares control approaches. Rule-based control is simple, optimization requires accurate models, tabular RL suits small problems and deep RL scales to many devices and continuous actions.

Approach	Adaptivity	Model Needed	Scalability
Rule-based schedules	Low	No	High
Optimization (MPC / game)	Medium	Yes	Medium
Tabular Q-learning	High	No	Low
Deep RL (DQN / PPO / SAC)	Very high	No	High

TABLE II Comparison of Control Approaches

2. Open Challenges And Future Work

Table III summarizes challenges. Future directions include multi-agent RL for neighbourhoods, transfer and meta-learning across homes, safe RL with formal constraints and integration with grid demand-response programmes.

Challenge	Mitigation
Sample inefficiency	Simulation pre-training
Safety and comfort	Constraint layer; safe RL
Transfer between homes	Transfer / meta-learning
Occupant behaviour	User feedback; preference learning
Data privacy	Local processing at home gateway

TABLE III Challenges and Mitigation

VII. CONCLUSION AND FUTURE WORK

Reinforcement learning allows smart homes to learn how to use energy efficiently under changing prices, weather and habits. The reviewed literature shows promising cost and energy reductions for HVAC, appliances, batteries and EV charging. With simulation-based pre-training, safety layers and user feedback, RL-based energy management can reduce bills and support a cleaner, more stable grid.

1. Acknowledgment

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science, Jaihind College of Engineering, Pune, for providing the facilities, guidance and encouragement needed to carry out this work. The authors also thank the faculty members and colleagues whose valuable suggestions helped improve the quality of this review.

References

1. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
2. J. R. Vázquez-Canteli and Z. Nagy, "Reinforcement learning for demand response: A review of algorithms and modeling techniques," Appl. Energy, vol. 235, pp. 1072–1089, 2019.
3. A.-H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," IEEE Trans. Smart Grid, vol. 1, no. 3, pp. 320–331, 2010.
4. T. Wei, Y. Wang, and Q. Zhu, "Deep reinforcement learning for building HVAC control," in Proc. 54th Annu. Design Autom. Conf. (DAC), 2017.
5. E. Mocanu et al., "On-line building energy optimization using deep reinforcement learning," IEEE Trans. Smart Grid, vol. 10, no. 4, pp. 3698–3708, 2019.
6. Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time EV charging scheduling based on deep reinforcement learning," IEEE Trans. Smart Grid, vol. 10, no. 5, pp. 5246–5257, 2019.
7. Z. Zhang, D. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," CSEE J. Power Energy Syst., vol. 6, no. 1, pp. 213–225, 2020.
8. J. R. Vázquez-Canteli, J. Kämpf, G. Henze, and Z. Nagy, "CityLearn v1.0: An OpenAI Gym environment for demand response with deep reinforcement learning," in Proc. 6th ACM Int. Conf. Syst. Energy-Efficient Buildings (BuildSys), 2019, pp. 356–357.

9. V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, pp. 529–533, 2015.
10. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv:1707.06347, 2017.
11. T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in Proc. ICML, 2018, pp. 1861–1870.