# A Reinforcement Learning-based Secure Demand Response Scheme for Smart Grid System

Aparna Kumari, Student Member, *IEEE*, Sudeep Tanwar, Senior Member, *IEEE*

***Abstract*—Smart Grid (SG) systems necessitate secure Demand Response Management (DRM) schemes for real-time decisions making to increase the effectiveness and stability of SG systems along with data security. Motivated from the aforementioned discussion, in this paper, we propose Q-SDRM, a secure DRM scheme for Home Energy Management (HEM) using Reinforcement Learning (RL) and Ethereum Blockchain (EBC) to facilitate energy consumption reduction and decrease energy costs. In cooperation with RL, Q-learning is adopted to make optimal price decisions using Markov Decision Process (MDP) to reduce energy consumption, which benefits both consumers and utility providers. Then, Q-SDRM uses Ethereum Smart-Contract (ESC) to deal with data security issues and incorporate with off-chain storage InterPlanetary File System (IPFS) that handles data storage costs issue. Experimental results reveal the effectiveness of the proposed Q-SDRM scheme, which significantly reduces energy consumption and energy cost. The proposed scheme also provides secure access to energy data in real-time compared with state-of-the-art approaches regarding different evaluation metrics such as scalability, overall energy cost, and data storage cost.**

***Index Terms*—Artificial Intelligence, Demand Response Management, Reinforcement Learning, Blockchain, Home Energy Management, Q-Learning.**

## I. Introduction

With the advancement in SG system, the concept of DRM in residential houses has gained widespread popularity [1] [2]. It facilitates end-consumers to make well-versed decisions about energy consumption and helps utility providers to reduce the peak demand for energy. It is one of the most reliable and cost-effective approach to manage energy demand during peak hours [1] [3]. According to the Department of Energy, United States, DRM is a program to reduce energy usage by making changes in the energy prices overtime or provide an incentive to a consumer to encourage lower energy usage [4]. It is categorized into two classes: (i) price-based (changes energy consumption pattern using time-dependent energy prices) and (ii) incentive-based (provides incentives to consumers based on energy load reduction during peak hour) [5]. The price-based DRM motivates consumers to change their energy usage patterns based on the time-varying energy prices; another side incentive-based DRM provides incentives to consumers based on energy load reduction during peak hours [5] [6].

Several research works in price-based DRM programs have been done globally [7][8]. To design a HEM system, energy consumption by household appliances is a major factor in price-based DRM programs; for example, Mixed-integer linear programming (MILP)-based system has been designed It enables to determine the optimal appliance scheduling thereby improve energy-efficiency and reducing consumer costs [9] [10] [11]. Likewise, Yu *et al.* [12] developed a price-based DRM approach for optimal management of home appliances. In this case, the Stackelberg game has been used for virtual energy trading wherein appliances can buy energy from the HEM system that acts as a retailer and offers retail prices, virtually [12]. Then, Shafie-Khah *et al.* [13] developed a price-based DRM approach to maximize consumer comfort and minimize the energy cost, and Park *et al.* [14] developed an approach for the small-scale generation of renewable energy considering the uncertain availability of it. Currently, most methodologies rely on conventional deterministic rules or abstract models, but these rules or models have certain drawbacks: (i) while handling volatile energy systems, deterministic rules fail to ensure optimality and may result in financial losses, (ii) the use of abstract models may be unrealistic while collating to actual systems and heavily relies on the SG operators skill, and (iii) in large-scale SG systems, game-theoretic/MILP optimization suffers from scalability issue due to a high number of binary values.

To handle the aforementioned issues, Artificial Intelligence (AI) is a prominent solution. The advancement in AI has shifted its focus towards optimal decision-making using RL. AI has been used to address realistic issues with limited breakthroughs: *AlphaGO Zero* and *AlphaGO* (introduced by Google), which reveal an excellent decision-making ability by using RL without prior knowledge of the environment. RL is *model free* that consists of an agent and environment interact by discrete-time steps. Fig. 1 illustrates the working of RL agents who interact with the stochastic environment for cumulative reward maximization. At each step, an action is selected by the agent and sent to the environment; then the environment moves to a new state, and the agent gets


Fig. 1: Reinforcement learning approach

A. Kumari and S. Tanwar (Corresponding Author) are with the Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India 382481, e-mail: (17ftphde22@nirmauni.ac.in, sudeep.tanwar@nirmauni.ac.in).