Agent based decision making system for stochastic supply & demand management in inter connected microgrid networks

D. Sai Koti Reddy,
Krishnasuri Narayanam
IBM Research - India
Email: saikotireddy@in.ibm.com

D. Raghuram Bharadwaj,
Shalabh Bhatnagar
Department of Computer Science and Automation
Indian Institute of Science, Bangalore, India

***Abstract*—One of the key challanges of Smart Grid management is ensuring to meet the demand while having control on the supply cost. It involves performing optimal actions both at the production and consumption sites of electric grid. In this work, we consider the problem of managing multiple microgrids attached to a central smart grid. Each of these microgrids are equipped with batteries to store renewable power. At every instant, each of them receive a demand to meet. Depending on the supply (i.e., currently available battery energy, power drawn either from the central grid or from the peer microgirds), each of the microgrids take a decision on from where to draw the energy to meet the demand. When a microgrids buys energy either from the central grid or from the peer microgrids, it can use that energy either to meet the current demand or to store in the battery storage for future use. Hence, there is a control decision problem at each microgrid on the number of power units to be bought or sold at every time instant. We note that both the forecasted demand and predicted renewable supply impact this decision by each microgrid. Further, we consider some amount of the forecasted demand to be adjustable in terms of the time when it can be met by the microgrid. Such an adjustable demand is attributed due to the activities of daily living pertraining to Smart Homes connected to the microgrid networks. Hence, we formulate this problem in the framework of Markov Decision Process and apply Reinforcement Learning algorithms to solve this problem. Through simulations, we show that the policy we obtain performs an significant improvement over traditional techniques.**

I. Introduction

The smart grid is a distributed energy network composed of intelligent nodes (or agents) that can either operate autonomously or communicate and share energy [1]. The purpose of a smart grid is to efficiently deliver energy to consumers as well as store and convert energy produced, e.g., according to prices, supply and demand.

A microgrid is a networked group of distributed energy sources with the goal of generating, converting and storing energy. While the main power stations are highly connected, micro-grids with local power generation, storage and conversion capabilities, act locally or share power with a few neighboring micro-grid nodes [2]. This scenario is being envisaged as an important alternative to the conventional scheme with large power stations transmitting energy over long distances.

In order to take full advantage of the modularity and flexibility of micro-grid technologies, smart control mechanisms are required to manage and coordinate these distributed energy systems so as to minimize the costs of energy production, conversion and storage, without jeopardizing grid stability.

The implementation of such smart controls is by no means easy for the following reasons: **(i)** Small scale energy production and storage is intrinsically related to intermittency of wind/solar energy and to variability in the load profile. So an important challenge is to increase resilience and reliability under stochastic supply and demand. **(ii)** Micro-grids can operate in two different modes: (a) when they are connected to the main power grid, and (b) in the isolated or island mode. Moreover, they can share energy with other microgrids that require energy. Thus, one needs to make dynamic decisions on (a) when to operate in the connected (to the power grid) or isolated modes, (b) when to share energy with other microgrids and when to store energy for future use, and (c) which form to store energy given that storage management itself involves heterogeneous storage technologies with different operating characteristics.

In this paper, we address two problems. First, energy sharing among the microgrids under stochastic supply and demand (mentioned above as (i) and (ii)) along with the optimal battery scheduling of a microgrid from the supply-side management (SSM) perspective. Second is from the demand-side management (DSM) perspective, which is efficiently scheduling the time adjustable demand from smart appliances in the smart home environment, called as an ADL (activity of daily living) demand along with the normal demand. Our goal here is to reduce the energy demand and supply deficit in the long-run. We address this learning and scheduling problem by modeling them as a Markov decision process (MDP) [3], [4].

A. Supply-side management problem

Supply-side management (SSM)[] deals with developing techniques to generate, transmit and distribute energy efficiently at supply-side. Cooperative energy exchange among microgrids is a popular technique in SSM for efficient energy distribution. Local energy sharing/exchange between micro-

grids has the following advantages: (a) it can significantly reduce power wastage that would otherwise result over long-distance transmission lines, and (b) it helps satisfy demand and reduce reliance on the main grid. Figure 1 shows a cooperative energy exchange model with multiple microgrids (on the distribution side of the network) that can cater to their individual local loads. Each microgrid controls its local sub-network through its controller (labelled C_1, C_2 etc.) that mainly has access to its local state information.

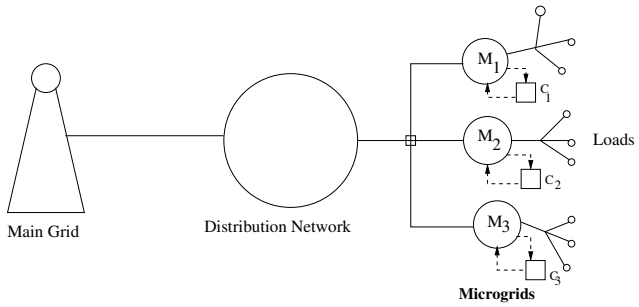

Fig. 1: Cooperative Energy Exchange Model

In classical power grids, system level optimization is done based on a centralized objective function, where as microgrid network has heterogeneous nature right from the manner in which electricity is generated such as from wind turbines, solar farms and diesel generators to energy storage devices such as batteries and capacitors. Because of this heterogeneity and the fact that energy can be shared between microgrids depending on requirements, one needs to consider asynchronous distributed techniques to control and optimize a smart grid system with a microgrid distribution network.

Related work : [5] provides a survey on game theoretic approaches for microgrids where both cooperative energy sharing models as well as non-cooperative game models for distributed control of microgrids are examined when the system model is known. Since models for energy dynamics are very unreliable [6], one has to use model-free algorithms to address these problems. Because of their model-free nature, reinforcement learning [7] approaches that are primarily data-driven control techniques are playing a significant role in these problems.

In [8], distributed reinforcement learning algorithm for coordinated energy sharing and voltage restoration in a islanded DC microgrid is proposed. In [9], reinforcement learning algorithm for optimal battery scheduling under the dynamic load environment and sloar power is proposed with the goal of reducing energy consumption from the main grid. In this paper, we consider the coordinated energy sharing among the grid connected microgrids with optimal battery scheduling problem when stochastic supply and adjustable stochastic demand is available.

B. Demand-side management problem

Load shifting is a popular technique used in demand-side management (DSM) [10]. It involves moving the consumption of load to different times within an hour, or within in a day, or even within a week. It doesn't lead to reduction in net quantity of energy consumed, but simply involves changing the time when the energy is consumed. Advantage due to load shifting for the customer is reduction in the energy consumption cost, and the advantage for the smart grid is in managing the peak load consumption. Hence load shifting is beneficial for both the consumers and the smart grid.

With the increased use of the smart appliances and smart home environments, the concept of load shifting is becoming increasingly handy for the smart grid as the demand from smart appliances is time adjustable in general. One or more of these smart appliances collectively achieve some activity in the smart home environment, called as an ADL (activity of daily living). It's possible to monitor and identify the ADLs in the smart home environments [11], [12]. When an ADL is active, the smart appliances associated with that ADL are switched on to perform the activity defined by the ADL thus adding load on the smart grid. With the help of the smart home technology, it's possible to find the amount of load each ADL puts on the grid, and also the allowed time window during which the ADL would perform the activity (e.g., washing machine running for an hour to clean the cloths anytime between 3PM to 6PM). If the time window for the ADL lets the smart grid have more than one possible way of scheduling the load, it's considered as flexible ADL. On the other hand, if the time window for the ADL lets the smart grid have exactly one possible way of scheduling the load, it's considered as non-flexible ADL (e.g., washing machine running for an hour to clean the cloths anytime between 3PM to 4PM, is not flexible since there is only one option of switching on the washing machine at 3PM). Thus the demand from the flexible ADLs need not be met at a fixed time period, instead could be met at any time period within a flexible time window. With the help of the advanced metering infrastructure (AMI) [13] that provides a two-way communication between the utility and customers, it's possible to take the decision of when to schedule the ADL demand at the smart grid and convey the same to the customer's smart meter.

There is other regular demand that needs to be met at fixed time periods, apart from the zero or more ADL related demand associated with any customer. This regular demand along with the zero or more non-flexible ADL demand of a smart home is considered to be non-ADL demand for the rest of the paper. Similarly, the demand due to zero or more flexible ADLs of the smart home is considered to be ADL demand.

There is prior art around scheduling the ADL-demand using the load shifting technique for handling the peak load scenarios [14]. However, they precisely know the supply profile while doing such a scheduling of the ADL-demand. In this paper, we propose scheduling of ADL-demand using the load shifting technique with uncertainty in the supply profile generated (e.g., renewable energy sources like solar or wind being the primary sources of power generation).

Our contributions : We summarize our contributions as follows :

(i) To the best of out knowledge, we are the first one to

integrate both the Demand-side and Supply-side management problems in a single Markov decision process framework. We used reinforcement learning algorithms which do not require knowledge of the underlying model to address these problems. Our algorithms are easy to implement and also scalable.

(ii) The Optimal scheduling of ADL demand at microgrid level, where both the demand and power generation is stochastic is first time introduced through this work.

The rest of the paper is organized as follows. The next section describes the important problems associated with the microgrids and solution techniques to solve them. Section III presents the results of experiments of our algorithms. Section IV provides the concluding remarks and Section V discusses the future research directions.

II. Problem formulation and MDP model

We consider N microgrids denoted by $\{1, \ldots, N\}$, which are inter-connected through distribution network. In this paper, we consider the case when microgrids are connected to the main grid (i.e, they are operated in grid connected mode). Each microgrid comprise of the distributed small scale renewable power generating sources and also equipped with energy storage devices. Let B_i be the maximum energy storage capacity of microgrid i. At every time step t of a day, ith microgrid controller C_i is having the following information:

(a) Total generated energy from all its distributed renewable energy sources denoted by r_t^i.
(b) Accumulated non-ADL demand denoted by d_t^i, from each load in the ith microgrid.
(c) Set of all ADL jobs at microgrid i denoted J_t^i. J_t^i is of the form $\{\gamma_1^i, \ldots, \gamma_n^i\}$, where jth ADL job γ_j^i, is a tuple consists of number of units of energy required to finish that job denoted by a_j^i and an integer f_j^i, which denotes number of future time slots remaining by which one can schedule that job without incurring the penalty. Let it be represented as follows $\gamma_j^i = (a_j^i, f_j^i)$. Let the total ADL demand be denoted by $A_t^i = \sum_{j=1}^{n} a_j^i$.
(d) Total energy available in the storage device of microgrid i, denoted by b_t^i.

In this paper, we are considering the cooperative energy exchange model under which microgrids can share energy among themselves. From the above available information, microgrid controller C_i at every time step t has to decide on the following choices: (a) Amount of energy it needs to buy/sell from the main grid, (b) Amount of energy it needs to buy/sell from the neighboring microgrids, (c) Amount of the energy it needs to store/take from the storage device, and (d) Sub-set of ADL jobs it needs to schedule. Both the demand and energy generated at microgrid i is uncertain/random due to random nature of loads (d_t^i and A_t^i) and renewable energy generation (r_t^i).

Markov decision process (MDP) is a general framework for modeling problems of dynamic optimal decision making under uncertainty. *****Need to write about MDP*******. In the next sub-section we provide the details of our MDP model.

A. MDP framework

1) State space: The state s_t^i at time instant t for microgrid i is as follows:

$$s_t^i = (t, nd_t^i, p_t, J_t^i), \tag{1}$$

where the net demand $nd_t^i = r_t^i + b_t^i - d_t^i$, which denotes whether there is an excess power or deficit. And $p_t and J_t^i$ denotes the price per unit energy and set of all ADL jobs at time instant t respectively. If net demand nd_t^i is positive then it implies that there is excess of power after meeting the non-ADL demand and if negative implies that there is a deficit in power even to meet the non-ADL demand. Current time slot is included in the state since optimal action can depend on time. For example, optimal action for solar microgrid is to sell the power in the morning as it can generate enough power in the afternoon. But it is not optimal to sell in the evening as it can not generate power in the night. Optimal action in the morning is to sell whereas in the evening is store.

2) Action space: At each time instant t microgrid controller needs to take two decision u_t^i and v_t^i. The first action u_t^i, if positive, denotes the number of units that the microgrid is willing to sell and if negative, represents the number of units that the microgrid is willing to buy. The second action V_t^i pertains to the scheduling decision of ADL jobs taken by microgrid i.

Let P_t^i be the power set of J_t^i, which consists of all possible combinations of the ADL jobs that can be scheduled at time instant t at microgrid i. Let A_t^i consists of the total aggregated demand for every subset in P_t^i. For example, let the jth element of $A_t^i(j) = \sum_{k=1, \gamma_k^i \in P_t^i(j)}^{n} a_k^i$, where $P_t^i(j)$ is jth element of P_t^i. The feasible region for action u_t^i is bounded as follows:

$$\begin{aligned} -min(M_t^i, B_t^i - nd_t^i + \max_{1 \le j \le 2^n} A(j)) \le u_t^i \\ \le max(0, nd_t^i - \min_{1 \le j \le 2^n} A(j)), \end{aligned} \tag{2}$$

where M_t^i denotes maximum amount of power main grid can give it to microgrid i. This constraint is to maintain stability of the main grid. Above bounds represents the following, microgrids can sell the power only after meeting its non-ADL demand and can buy to fill its battery after meeting its both ADL and non-ADL demands.

After controller picks action u_t^i, we construct the feasible set F_t^i, which is a subset of P_t^i. It consists of all possible subsets of ADL jobs that can be scheduled with u_t^i. More formally, each element j of F_t^i has to satisfy the following condition : $A_t^i(j) \le u_t^i$, where $A_t^i(j)$ is total energy required to finish all the ADL jobs in it. Now, controller picks action v_t^i which is an element in F_t^i, which results in scheduling all the ADL jobs in that subset. The remaining power is used to meet the non-ADL demand and for storing in the battery for future use.

Let $\widehat{J}_{t+1}^i$ be the new set of ADL jobs received by controller at time instant $t+1$. Depends on action v_t^i, some of the ADL

jobs will not get scheduled. We pass them to time step $t+1$, if they can be scheduled without incurring the penalty. The set of all ADL jobs at time instant $t+1$ is union of the new ADL jobs and old ADL jobs which are not scheduled after reducing there f_j^i by one (number of future time slots remaining by which one can schedule that job without incurring the penalty) . $J_{t+1}^i = \widehat{J}_{t+1}^i \cup \widetilde{J}_t^i$, where $\widetilde{J}_t^i = \{(a_1^i, f_1^i - 1), \ldots, (a_n^i, f_n^i - 1)\}$, and $(a_j^i, y_j^i) \in \overline{J}_t^i$. And $\overline{J}_t^i = J_t^i - v_t^i$.

The storage device battery information is updated as follows:

$$b_{t+1}^i = max(0, nd_t^i - u_t^i), \tag{3}$$

which denotes the power available after meeting the non-ADL demand and ADL demand is stored in the battery for future use.

3) Single stage reward function: In this paper, we want to maximize the profit of each microgrid obtained by selling power while reducing the demand and supply deficit. Our singe stage reward function has both the reward obtained by selling power and penalty for unmet demand. The single stage reward function for our MDP is as follows :

$$g_t^i(s_t^i, u_t^i) = p_t * u_t^i + c * (min(0, nd_t^i - u_t^i)) + c * \sum_{k=1}^{n} I_{f_k^i=0} a_k^i, \tag{4}$$

The first term represents the cost/gain incurred for buying/selling the power, and the second and third terms represents the penalty incurred for not meeting the non ADL demand and ADL demand respectively. Here, c is penalty per unit of unmet demand and $I_{f_k^i}$ is indicator random variable which is equal to one if $f_k^i = 0$ and zero otherwise. ******Need to write about transition probability kernel **********.

B. Average cost setting

Finally, the objective of the microgrid i is to maximize the following [15]:

$$limsup_{n\to\infty} 1/n \sum_{k=0}^{n} E(g^i(s_k, u_k)), \tag{5}$$

where $E(.)$ is the expectation.

We also consider the long run discounted cost formulation. The objective here is to maximize the following:

$$limsup_{n\to\infty} \sum_{k=0}^{n} \gamma^k * E(g^i(s_k, u_k)), \tag{6}$$

where γ is the discount factor.

III. Algorithm

We first note that the renewable generation is uncertain in nature. That is, we do not know in the current time period, the renewable generation in the future time periods. Also, we do not know the probability transition model of the demand. Hence we employ the RL algorithms that provides optimal solution under model-free environments.

To solve the above average cost formulation, we apply a popular RL algorithm, Q-Learning for the average cost. Our objective is to obtain a stationary optimal policy $\pi : S \to A$, which is a mapping from state space to action space.

We apply the Relative Value Iteration (RVI) Q-Learning described in [15]. In this algorithm, we update the Q-values in each iteration according to the following rule :

$$Q^{n+1}(s,a) = Q^n(s,a) + \alpha(n)(g(s,a,s^{'}) + \tag{7}$$

$$max_u Q^n(s^{'}, u) - max_u Q^n(s_0, u) - Q^n(s,a), \tag{8}$$

where α is the learning rate and s_0 is any prescribed state.

The $Q(s,a)$ in each iteration represents the average reward obtained in state s by taking an action a. In [15], they show that under appropriate learning rate, the algorithm converges to the optimal policy.

Each microgrid will run the algorithm independently under convergence. Then the optimal policy is obtained as follows:

$$\pi^*(s) = max_u Q(s,u) \tag{9}$$

That is, the optimal policy in state s is selected by taking the maximum of all actions of Q-value of corresponding state.

IV. Simulation Experiments

We implemented our models on network with 3 microgrids. Two microgrids are operating on solar renewable generation and the other on wind renewable energy. To simulate the renewable generation, we use RAPsim software [16]. RAPsim is an open source simulator for analyzing the power flow in microgrids. It has a provision for simulating the renewable generation, which is the main feature that we use in our experiments. We construct our microgrid model as shown in the Fig 2. We can see that there are three microgrids, two of them operating on the solar energy and the other on the wind energy. The solar microgrid in the right has more capacity than that of the one in the left. These microgrids also have electrical connections from the main grid. Each microgrid provides power to the respective houses on their power line.

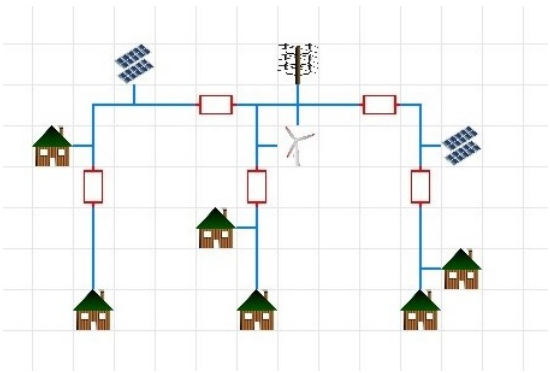

Fig. 2: Experimental Setup

A. Implementation

We implement the model we described in the section 2. We call this model as $ADL - sharing$ model. For comparison purposes, we implement following models.

- **Greedy-ADL model**: In this model, the microgrids will exhibit the greedy behavior. They share the power only if there is excess power after filling up the battery. That is, the action in each instant is bounded by

$$-min(M, B - nd) \leq max(0, nd - B) \qquad (10)$$

  That is, if the net demand is negative, decision is taken on amount to power to buy to meet the demand and fill the battery. On the other hand, if the net demand is positive, it is first used to fill the battery and only if there is any excess power left, it will be sold to other microgrids.
- **Non-ADL model**: This model is similar to the $ADL - sharing$ model, but without the integration of ADL demand. In this model, the ADL demand is added to the main demand. Unlike the $ADL - sharing$ model, there is no flexibility of intelligently scheduling the ADL demand.

B. Setup

We simulate the above setup for the month of September 2017 and collect the wind and solar renewable power generated each day every hour. The parameters for our experiments are described below. The number of decision time periods is taken to be 4. We consider 3 demand values for all the microgrids - 2, 4 and 6 units. The probability transition matrices for all the 3 microgrids are given below :

$$P_1 = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.1 & 0.2 & 0.7 \\ 0.8 & 0.1 & 0.1 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0.2 & 0.2 & 0.6 \\ 0.8 & 0.1 & 0.1 \\ 0.2 & 0.7 & 0.1 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{bmatrix}$$

The price values is considered to be 5, 10 and 15. The Probability transition matrix for the price vector is given below:

$$Q = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.1 & 0.5 & 0.4 \\ 0.5 & 0.4 & 0.1 \end{bmatrix}$$

Maximum size of battery and renewable power generated is taken to be 8 units. The maximum power that a microgrid can obtain from the main grid is set to 10 units.

We consider 3 ADLs in our experiment. We assume that all these 3 ADL's are known to microgrids in the first time period. First ADL requires 1 unit of power that needs to be satisfied within the second time period. Second ADL requires 1 unit of power within the third time period. Third ADL requires 2 units of power within the fourth time period.

With this setup, we compare our proposed models. The algorithms are trained for 10^6 iterations. For comparison purposes, we plot value of threshold c on X-axis and Average reward on Y-axis. Average reward is computed as follows. We run the trained models for 1000 runs and average the reward obtained by each microgrid.

C. Observations

D. Discussion

In Figure 1, we compare our $ADL - sharing$ model with the $Greedy - ADL$. As discussed earlier, in the latter model the sharing of power is done only when there is excess after filling the battery. We can see that the first model outperformed the second model. Even though there will be less buying of power in $Greedy - ADL$, there will be no selling of power as well. Therefore the overall profit obtained will not be higher than that, when the intelligent decisions are made. Hence we can conclude that the intelligent sharing of power among microgrids yields more profit than that of the non-sharing case.

In Figure 2, we compare our $ADL - sharing$ model with the $Non - ADL$ model. We can observe from the plot that, our model 1 outperforms the model 3. The reason for that is discussed below. Consider $c \geq 15$, which is the maximum price. In $ADL - sharing$, there is a flexibility to intelligently schedule the ADL activities according to the non-ADL demand and price. But this is not the case with the $Non - ADL$ model. In this case, the penalty will be immediately levied if the demand (including the ADL demand) is not met. This results in the poor performance of $Non - ADL$. Hence we can conclude that intelligently scheduling the ADL demand results in the better performance.

[update other cases after results]

From the above discussion we can conclude that our proposed algorithm along with the flexible ADL demand integration, is the best algorithm that provides more profits to the microgrids.

V. Experimental Results

The following experimentals results are desired to be observed:

- With different ADLs being scheduled along with the non-ADL demand, few of the ADLs are expected to be scheduled at the beginning of the allowed execution time window of the ADL, few other ADLs are expected to be scheduled at the end of the allowed execution time window of the ADL, while some other ADLs get scheduled at the mid of the allowed execution time window. This ensures that the MDP learning agents exploit the fact that the ADL demand is flexible to meet in a given range of time window. On the other-hand, it is not desired that the learning agent schedules all the ADLs either at the beginning or at the end of the allowed time window of execution.

- With surplus energy available at a microgrid at any moment, it is desired not to sell this surplus to other microgirds if there is more demand than supply in the near feature. For example, if the renewable energy source for a microgrid is solar energy, then if there is surplus energy(i.e., excess energy available after meeting the demand at some moment) at the microgrid during the midday, the microgrid could sell that surplus energy to the other microgrids (because it is expected to generate more supply as the day progresses); on the other-hand, if there is surplus energy at the microgrid during the end of the day, the microgrid might not want to sell that surplus energy to the other microgrids (because there may not be much supply possible for the rest of the day).
- How are we ensuring this? If there is 5 units of surplus at time t. If the demand at time (t+1) is 5 units, it's possible to meet that demand by storing 5 units at time t. Other possibility is, sell the 5 units in time t, and buy 5 units in time $t + 1$ from some other microgrid. However the first option is most desired. How are we ensuring this in our experiments? One possible way to implement this is by ensuring the buying cost to be more than the selling cost for one unit of energy.

VI. Conclusion

VII. Future Work

Acknowledgment

The authors would like to thank Robert Bosch Centre for Cyber-Physical Systems, IISc, Bangalore, India for supporting part of this work.

References

[1] G. Weiss, *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press, 1999.

[2] H. Farhangi, "The path of the smart grid," *IEEE power and energy magazine*, vol. 8, no. 1, 2010.

[3] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[4] D. Bertsekas, "Dynamic Programming and Optimal Controlvolume 2," 1999.

[5] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, 2012.

[6] R. Zamora and A. K. Srivastava, "Controls for microgrids with storage: Review, challenges, and research needs," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 7, pp. 2009–2018, 2010.

[7] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.

[8] L. Zifa, L. Ya, Z. Ranqun, and J. Xianlin, "Distributed reinforcement learning to coordinate current sharing and voltage restoration for islanded dc microgrid," *Journal of Modern Power Systems and Clean Energy*, pp. 1–11.

[9] R. Leo, R. Milton, and S. Sibi, "Reinforcement learning for optimal energy management of a solar microgrid," in *Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS), 2014 IEEE*. IEEE, 2014, pp. 183–188.

[10] B. Davito, H. Tai, and R. Uhlaner, "The smart grid and the promise of demand-side management," McKinsey, 2010.

[11] I. Bae, "An ontology-based approach to adl recognition in smart homes," *Future Generation Computer Systems*, vol. 33, pp. 32–41, 2014.

[12] G. Baryannis, P. Woznowski, and G. Antoniou, "Rule-based real-time ADL recognition in a smart home environment," in *Rule Technologies. Research, Tools, and Applications - 10th International Symposium, RuleML 2016, Stony Brook, NY, USA, July 6-9, 2016. Proceedings*, ser. Lecture Notes in Computer Science, vol. 9718. Springer, 2016, pp. 325–340.

[13] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 473–484, 2014.

[14] C. O.Adika and L. Wang, "Smart charging and appliance scheduling approaches to demand side management," *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 232–240, 2014.

[15] J. Abounadi, D. Bertsekas, and V. S. Borkar, "Learning algorithms for markov decision processes with average cost," *SIAM Journal on Control and Optimization*, vol. 40, no. 3, pp. 681–698, 2001.

[16] M. Pochacker, T. Khatib, and W. Elmenreich, "The microgrid simulation tool rapsim: description and case study," in *Innovative Smart Grid Technologies-Asia (ISGT Asia), 2014 IEEE*. IEEE, 2014, pp. 278–283.