A unified decision making framework for supply and demand management in microgrid networks

Raghuram Bharadwaj *, Sai Koti Reddy †, Krishnasuri Narayanam †, Shalabh Bhatnagar *

∗ Department of CSA, Indian Institute of Science, Bangalore, India
† IBM Research, Bangalore, India.

31 October 2018

The concept of 'Smart Grids'

- Solution concept aimed at improving traditional power grid operations.
- Distributed energy network composed of intelligent nodes (or agents) that can either operate autonomously or communicate and share energy.

Objectives that are considered

- Demand side management
  - Minimize the peak Demand : Dynamic Pricing and integrating the customers participation.
  - Minimize the cost of consumption of power.
- Supply side Management
  - Minimize the Demand-Supply power deficit.
  - Minimize the use of fossil fuels.

What are 'Microgrids'

- A microgrid encompasses a networked group consisting of renewable energy generation sources, battery storage with the aim of providing energy to small areas.
- An important alternative to the conventional scheme with large power stations transmitting energy over long distances.
- Can operate in Island Mode or Cooperative mode.

Introduction

- We consider two problems - one on the supply side and another on the demand-side.
- On supply-side : Energy sharing among microgrids with the goal of maximizing profit obtained from selling power while meeting the desired customer demand.
- On Demand-side : Optimally scheduling the time-adjustable demand - i.e., of loads with flexible time windows in which they can be scheduled.

Setup

- N microgrids denoted by $\{1, \ldots, N\}$, which are inter-connected through the central electric grid distribution network.

Setup

- N microgrids denoted by $\{1, \ldots, N\}$, which are inter-connected through the central electric grid distribution network.
- Each microgrid i is equipped with a renewable energy source r_i and battery of capacity B_i units.
- At every time instant t, microgrid i gets a demand d_t^i from its customers.

Setup

- N microgrids denoted by $\{1, \ldots, N\}$, which are inter-connected through the central electric grid distribution network.
- Each microgrid i is equipped with a renewable energy source r_i and battery of capacity B_i units.
- At every time instant t, microgrid i gets a demand d_t^i from its customers.
- The time-adjustable demand is also known as 'Activities of Daily Life (ADL)'.
- Each microgrid i has a set of ADL jobs ($J_t^i = \{\gamma_1^i, \ldots, \gamma_n^i\}$), where the j^{th} ADL job $\gamma_j^i = (a_j^i, f_j^i)$.

Setup

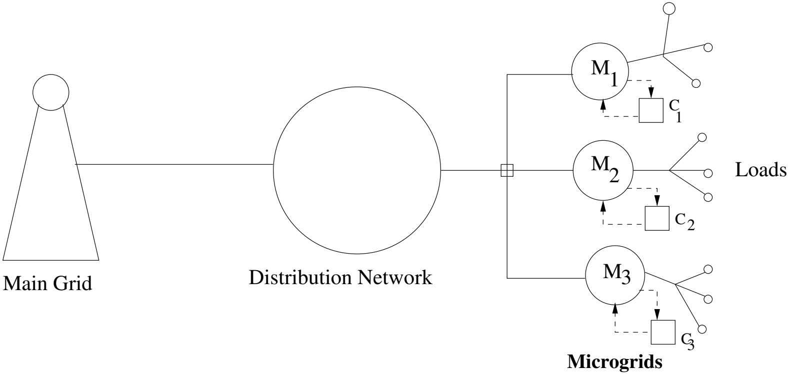


Figure: Cooperative Energy Exchange Model

Setup

Each microgrid i at every time instant t, has to decide on the following:
(a) Amount of energy it needs to buy (sell).
(b) The subset of ADL jobs it needs to schedule.

Model

- We formulate this problem in the framework of Markov Decision Process (MDP) [1].
- MDP is characterized by $< S, A, P, R >$.
- We divide a day into t decision time units.
- The state s_t^i at time instant t for the microgrid i is as follows:

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

where the net demand $nd_t^i = r_t^i + b_t^i - d_t^i$. Here b_t^i is the battery level of microgrid i at time t, and p_t denotes the price per unit energy.

[1] Sutton, Richard S., Andrew G. Barto, and Francis Bach. Reinforcement learning: An introduction. MIT press, 1998.

Model

- We formulate this problem in the framework of Markov Decision Process (MDP) [1].
- MDP is characterized by $< S, A, P, R >$.
- We divide a day into t decision time units.
- The state s_t^i at time instant t for the microgrid i is as follows:

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

  where the net demand $nd_t^i = r_t^i + b_t^i - d_t^i$. Here b_t^i is the battery level of microgrid i at time t, and p_t denotes the price per unit energy.
- If $nd_t^i > 0$, then there is excess of power after meeting the non-ADL demand and if $nd_t^i < 0$, there is a deficit in power even to meet the non-ADL demand.

[1] Sutton, Richard S., Andrew G. Barto, and Francis Bach. Reinforcement learning: An introduction. MIT press, 1998.

Model

- Actions : Each microgrid need to make two decisions : u_t^i and v_t^i.
- Goal: Maximize profit.
- Single-stage reward for microgrid i:

$$
\begin{aligned}
r^i(s_t^i, u_t^i) = p_t * u_t^i + c * min(0, nd_t^i - u_t^i) \\
- c * \sum_{k=1}^{n} I_{\{f_k^i = 0\}} a_k^i,
\end{aligned}
\tag{2}
$$

where c (≥ 0) is penalty per unit of unmet demand.

Model

- We formulate this problem in the Average-cost MDP setting, where the objective is to maximize the expected average profit of the microgrids
- The long-run average reward objective function $J(\pi)$ of the microgrid i for a given policy π is given as follows:

$$J(\pi) := \lim\sup_{n \to \infty} \frac{1}{n} E\left(\sum_{t=0}^{n-1} r^i(s_t, u_t) \middle| \pi \right), \tag{3}$$

where $E(.)$ denotes the expected value. Here we view a policy π as the map $\pi : S \to A$ which assigns for any state s, a feasible action a.

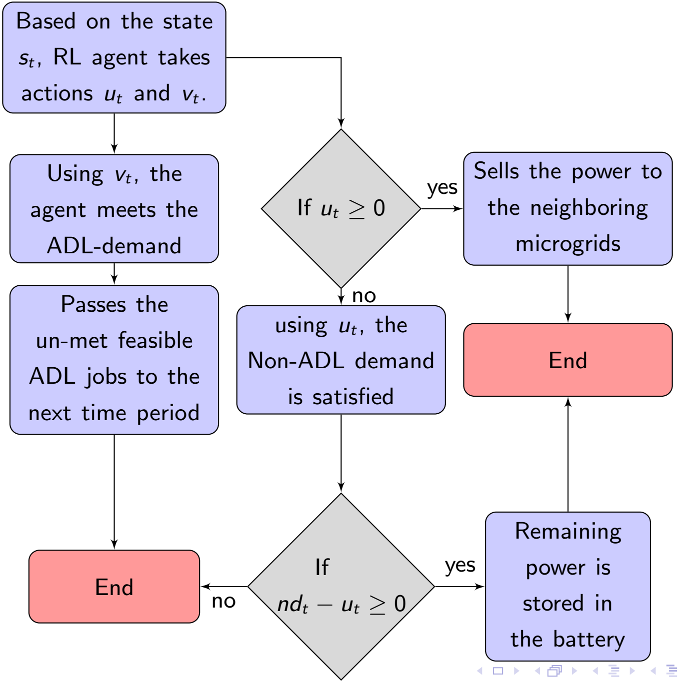
Algorithm

- We do not assume any model of the system (i.e., probability transition model of the demand, prices).
- We apply the Relative Value Iteration (RVI) based Q-Learning algorithm [2]. The Q-values $Q^n(s, u)$

$$Q^{n+1}(s,u) = Q^n(s,u) + \alpha(n)(r(s,u,s^{'})+ max_u Q^n(s^{'},u) - max_u Q^n(s_0,u) - Q^n(s,u)), \quad (4)$$

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

[2] Abounadi, Jinane, D. Bertsekas, and Vivek S. Borkar. "Learning algorithms for Markov decision processes with average cost." SIAM Journal on Control and Optimization 40.3 (2001): 681-698.


Based on the state s_t, RL agent takes actions u_t and v_t.
Using v_t, the agent meets the ADL-demand
Passes the un-met feasible ADL jobs to the next time period
End
If $u_t \geq 0$
yes
no
Sells the power to the neighboring microgrids
End
using u_t, the Non-ADL demand is satisfied
If $nd_t - u_t \geq 0$
yes
no
Remaining power is stored in the battery

Experimental setup

- We used the RAPSim [3] simulator for estimation of the renewable energy.
- We implement our models on a network with three microgrids and five microgrids respectively.

[3] Pochacker, Manfred, Tamer Khatib, and Wilfried Elmenreich. "The microgrid simulation tool RAPSim: Description and case study." In Innovative Smart Grid Technologies-Asia (ISGT Asia), 2014 IEEE, pp. 278-283. IEEE, 2014.

Implementation

For comparison purposes, we also implement the following models.

- **Greedy-ADL model**: Here, microgrids exhibit greedy behavior, share power only after filling their respective batteries completely.
- **Non-ADL model**: Here, ADL demand is treated as normal demand. Penalty is levied immediately if the demand is not met in the current time slot.

Simulation setup

- The number of decision time periods (t) in a day is taken to be 4.
- The maximum size of the battery (B_t^i) and maximum power that a microgrid can obtain from the main grid (M_t^i) are considered to be 8 and 10 energy units respectively.

Simulation setup

- The number of decision time periods (t) in a day is taken to be 4.
- The maximum size of the battery (B_t^i) and maximum power that a microgrid can obtain from the main grid (M_t^i) are considered to be 8 and 10 energy units respectively.
- At each microgrid, we consider 3 ADL jobs, $\{\gamma_1^i = (1,2), \gamma_2^i = (1,3), \gamma_3^i = (2,4)\}$ at the start of the day, where ADL job $\gamma_j^i = (a, b)$ requires a units of energy within b time slots.

- We use the average profit obtained by each microgrid as a performance metric to evaluate the models.

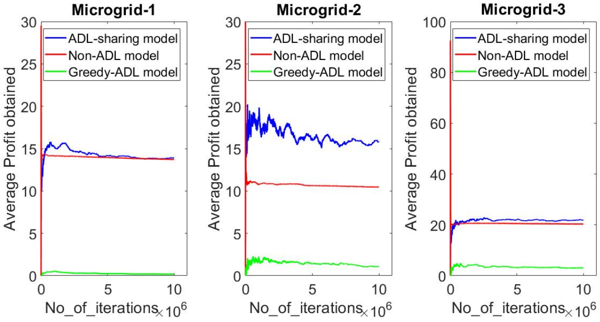


Figure: Convergence of algorithms when $c = 0$ for a three microgrid network.

Results

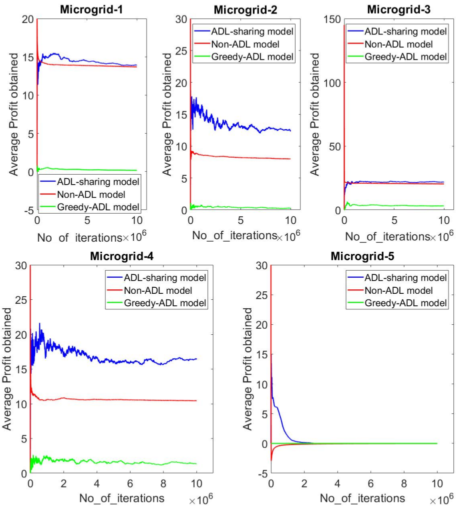


Figure: Convergence of algorithms when $c = 0$ for the five microgrid network.

Results

We run the trained models for 1000 runs to obtain the average reward in each case.

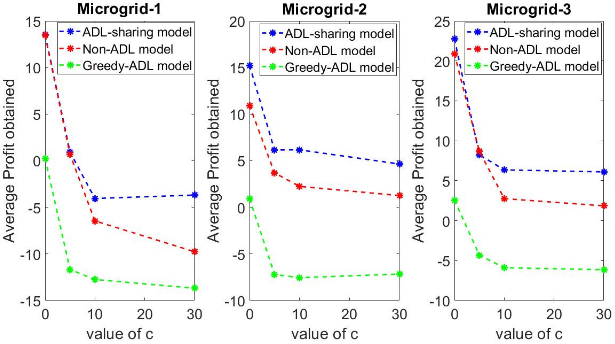


Figure: Comparison of algorithms on the three microgrid network

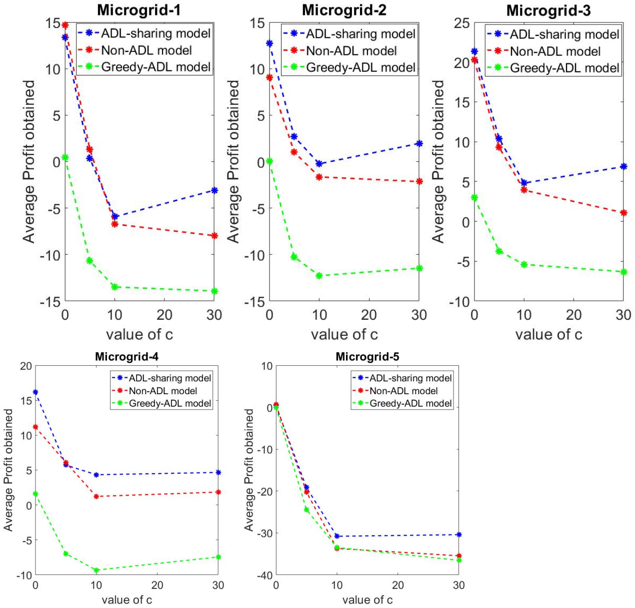


Figure: Comparison of algorithms on the five microgrid network

Discussion

- ADL-Sharing model provides more profits to the microgrids by:
  - Intelligently sharing the power with the neighboring microgrids.
  - Making optimal use of flexible ADL jobs.

Future Work

- We would like to consider the pricing mechanism for microgrids.
- Asynchronous multiagent RL algorithms will play a role here.
- We would like to use efficient RL algorithms with function approximation to scale the proposed algorithms.

Acknowledgement

We would like to acknowledge and thank the donors for the financial support from "Lakshmi and Aravamudan Student Travel fund" award obtained through Office of Development and Alumni Affairs (ODAA) IISc.

Thank You