SolarTrader: Enabling Distributed Solar Energy Trading in Residential Virtual Power Plants

Yuzhou Feng, Qi Li, Dong Chen, and Raju Rangaswami
School of Computing and Information Sciences
Florida International University

ABSTRACT

Distributed solar energy resources (DSERs) in smart grid systems are rapidly increasing due to the steep decline in solar module prices. This DSER penetration has prompted utilities to balance the real-time supply and demand of electricity proactively. A direct consequence of this is *virtual power plants (VPPs)* that enable solar generated energy trading to mitigate the impact of the intermittent DSERs while also benefiting from distributed generation for more reliable and profitable grid management. However, existing energy trading approaches in residential VPPs do not actually allow DSER users to trade their surplus solar energy independently and concurrently to maximize benefit potential; they typically require a trusted third-party to play the role of an online middleman. Furthermore, due to a lack of fair trading algorithms, these approaches do not necessarily result in "fair" solar energy saving among all the VPP users in the long term.

We propose *SolarTrader*, a new solar energy trading system that enables unsupervised, distributed, and long term fair solar energy trading in residential VPPs. In essence, SolarTrader leverages a new multi-agent deep reinforcement learning approach that enables *peer-to-peer* solar energy trading among different DSERs to ensure that both the DSER users and the VPPs maximize benefit. We implement SolarTrader and evaluate it using both synthetic and real smart meter data from 4 U.S. residential VPP communities that are comprised of ~229 residential DSERs in total. Our results show that SolarTrader can reduce the aggregated VPP energy consumption by 83.8% when compared against a non-trading approach. Furthermore, SolarTrader achieves a ~105% average saving in VPP residents' monthly electricity cost. We also find that SolarTrader-enabled VPPs can achieve a fairness of 0.05, as measured by the *Gini Coefficient*, a level equivalent to that achieved by the fairness-maximizing Round-Robin approach.

CCS CONCEPTS

• **Computing methodologies** → **Multi-agent planning**; **Modeling methodologies**; **Model verification and validation**.

KEYWORDS

Solar Energy Trading, ML, Deep Learning, Reinforcement Learning

BuildSys '20, November 18–20, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-8061-4/20/11…$15.00
https://doi.org/10.1145/3408308.3427611

ACM Reference Format:
Yuzhou Feng, Qi Li, Dong Chen, and Raju Rangaswami. 2020. SolarTrader: Enabling Distributed Solar Energy Trading in Residential Virtual Power Plants. In *The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20), November 18–20, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3408308.3427611

1 INTRODUCTION

The number of solar-powered homes is rapidly increasing due to a steep decline in solar module prices. To illustrate, the cost of solar energy in $/W dropped an estimated 80% from 2010 to 2018 [9], resulting in a 700% increase in distributed solar energy resources (DSERs) in the U.S. over the same period [17]. This solar penetration has prompted utilities to employ DSER management for load reduction, and has given rise to Independent System Operator (ISO) for wholesale market bidding participation. As a result, U.S. utilities have demonstrated an interest in deploying virtual power plants (VPPs) to confront and adapt to these changes. As shown in Figure 1, a VPP is an aggregation of different types of power generation sources, including traditional energy sources and DSERs, which employing the same type of electricity pricing scheme, demand response (DR) approaches, and other net load shifting schemes to ensure the stability and reliability of residential grid management functions. Typical DSERs in residential VPPs can be comprised of small-scale rooftop residential solar photovoltaic (PV) deployments, community shared solar arrays (CSSA) [11, 14], solar farms, and solar energy storage systems (e.g., electric vehicles (EVs), battery storage arrays). VPPs enable the individual DSER owners to have access and visibility across the entire energy market, and benefit from the market to maximize their revenue potential [27]. In addition, grid operation can also benefit from the collaborative and more efficient energy usage of all-VPP available DSER generation and storage capacity.

A key goal in the management of VPPs is enabling a solar generated energy trading platform that achieves the above benefits for both VPPs and the DSER users. Recently, researchers in the areas of micro-grid [18, 19, 38] and community shared solar arrays (CSSA) [11, 14] have proposed several solar energy trading approaches [7, 8, 10, 21, 22, 26, 30, 34, 36, 37] that are candidates for use within a VPPs' solar energy trading system. Unfortunately, these solar energy sharing approaches are not able to extract maximum benefit from the DSERs such as rooftop solar PV arrays, CSSA [11, 14], and other solar energy storage systems. Their primary drawback is an inability to allow DSER users to trade their surplus solar energy independently and concurrently to maximize benefits from solar generated energy, instead requiring a trusted third-party, such as an ISO or the utilities, to play the role of middleman to supervise the trading processes.

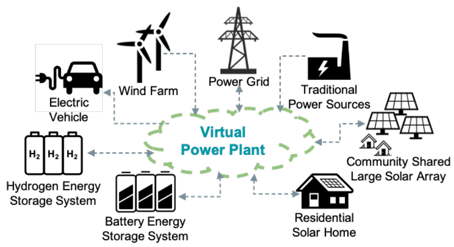


Figure 1: The overview of the virtual power plant (VPP).

In this paper, we propose a new solar energy trading system called *SolarTrader* that can enable unsupervised, distributed, scalable, and long-term fair solar energy trading in residential VPPs. We model a VPP system as a decentralized collaborative environment. Each DSER is abstracted as an agent that is collaboratively evolving in a common environment, the VPP, with other agents. In essence, SolarTrader leverages multi-agent deep reinforcement learning (MADRL) to automatically learn a common and collaborative policy in a shared VPP environment. This policy is then deployed on DSER agents without additional training. The policy is both trained and deployed using a peer-to-peer (P2P) approach manner such that all the DSER agents can trade their solar generated energy concurrently and independently, working as a fully distributed system. Through reward and penalty learning, SolarTrader ensures fairness of solar energy trading outcomes for all the DSER agents in the long term.

To better understand the potential and limitations of SolarTrader, we implemented a SolarTrader prototype and evaluate it using smart meter energy data from ~229 homes of 5 U.S. residential VPPs located in Colorado, Massachusetts, New York, California, and Texas. Our results show that SolarTrader can help DSER users reduce their net grid energy consumption by 109.8% when compared to a non-sharing approach. We also find that SolarTrader enabled VPPs can yield a *Gini coefficient* (a standard measure of an approach's fairness performance) of 0.05, which is the same as that of the most-recent Round-Robin (RR) trading approach that has the best fairness performance in long term trading. Specifically, we evaluate SolarTrader using multiple ways:

- We compare and validate SolarTrader's performance using a wide set of metrics for a synthetic smart meter dataset comprised of 100 residential DSERs from Massachusetts.
- We evaluate SolarTrader's performance using a real smart meter energy dataset which consists of 119 solar-powered homes from 4 different residential VPP communities located in Colorado, New York, California, and Texas using PecanStreet API [5].
- We further validate SolarTrader's trading performance using real-time smart meter data from 10 DSERs from Massachusetts.

Through these empirical studies, we demonstrate that a SolarTrader-enabled VPP is capable of enabling its users to trade solar energy with their neighbors more efficiently to achieve optimal monetary benefits while empowering the VPP with reliable, online, distributed DSER management.

Releasing Datasets and Code. We release all the datasets that are comprised of smart meter data over 229 DSERs and the source code of SolarTrader on our website [33].

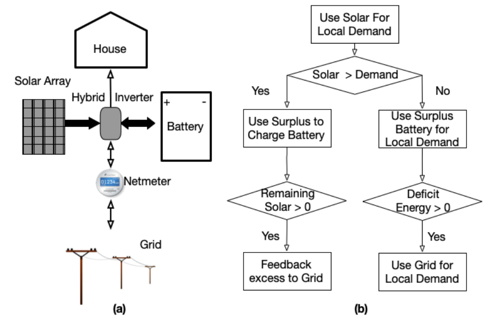


Figure 2: (a) The overview of a residential DSER, and (b) the operation pipeline of a dedicated solar PV array system.

2 BACKGROUND

2.1 Problem Statement

Implementing an efficient residential VPP involves addressing multiple goals. First is a new approach that can automatically gather and abstract all the distributed solar energy resources (DSERs) such that each DSER can be abstracted as an agent in the collaborative VPP environment. Next is a set of new solar energy trading algorithms for the agents to trade their surplus energy in a distributed and P2P manner. In addition, we need to learn trading algorithm performance, including user monetary benefits, long term fairness, and VPP management benefits, when applying each energy trading approach. We outline our specific approach for achieving each of these goals below.

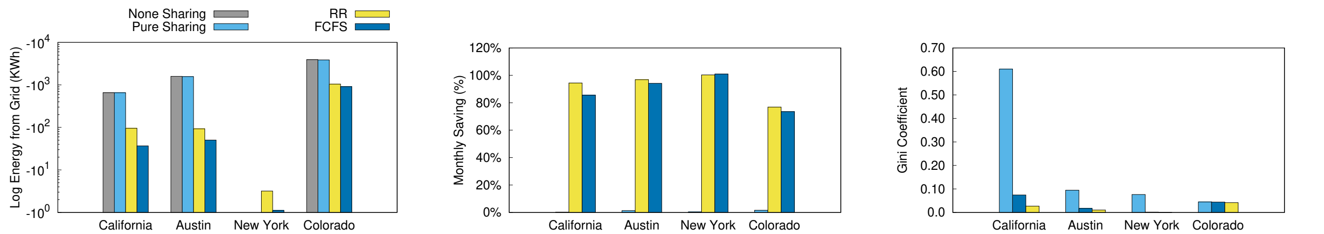
Formally, given a VPP V_i, we abstract each DSER $DSER_i$ in V_i as an agent a_i such that all the DSERs can trade their surplus energy concurrently in the multi-agent VPP system V_i. The objective is to find the minimum energy cost of its operations over time T in the entire VPP environment, which can be defined as follows,

$$\frac{1}{N} \cdot \sum_{i=0}^{T} \sum_{k=0}^{N} [E_{grid_k}(i) \cdot P_{grid_k}(t) - E_{trade_k}(t) \cdot P_{trade_k}(t) \\ -E_{feedback_k}(t) \cdot P_{feedback_k}(t))] \tag{1}$$

where $E_{grid_k}(t)$, $E_{trade_k}(t)$, and $E_{feedback_k}(t)$, represent the electricity that the agent k draws from the grid, trades with its neighbors, and feeds back to the grid, respectively, at moment t. $P_{trade_k}(t)$, $P_{grid_k}(t)$, and $P_{feedback_k}(t)$ denote the electricity cost in \$ per kWh when the agent k draws from the grid, trades with its neighbors, and feeds back to the grid, respectively.

2.2 System Model and Pricing Model

System Model. A general DSER in residential VPPs is a residential rooftop solar PV system that is typically comprised of solar PV arrays, a smart hybrid inverter, and battery arrays. Figure 2 (a) illustrates the system overview of a rooftop solar PV system in daytime when solar generation is available, and Figure 2 (b) shows a dedicated (non-sharing) solar PV system operation pipeline. As we had discussed in Section 1, instead of feeding excess solar generated energy back to grid, sharing surplus solar energy with neighbors can offer greater monetary benefit for DSER owners and also help VPPs better manage grid operation. By enabling solar generated energy trading in a VPP when excess solar energy is available, selling solar



(a) Aggregated Consumption (b) Monthly Energy Saving (c) Fairness in GC

Figure 3: Comparison of different solar energy trading approaches based on (a) the amount of monthly aggregated energy consumption, (b) monthly electricity bill saving, and (c) the Gini Coefficients.

generated energy to homes that have deficit of energy becomes possible. We discuss the details of this approach in Section 4.

Pricing Model. In this paper, we assume that all the DSERs use regular retail electricity pricing schemes. The utilities purchase net-metered electricity at wholesale prices and sell it to VPP residents at retail prices [2, 3]. Instead of purchasing electricity from utilities at retail prices, SolarTrader enables VPP users to purchase electricity from neighbors that have surplus electricity at a price that is higher than the utilities' wholesale price but significantly lower than the utilities' retail price. We also assume VPP users have a billing agreement among all the DSERs to determine the costs of energy borrowed or lent. Finally, the micro-payments necessary for purchasing electricity can be executed using third-party payment systems such as PayPal [4], Venmo [6], etc.

2.3 Comparing Current Approaches

We examine design alternatives for solar generated energy trading approaches, including the dedicated (non-trading) approach, the pure always sharing approach, the first-come-first-service (FCFS) sharing approach, and the round-robin (RR) sharing approach. In doing so, we are able to review a wide range of the most recent sophisticated solar energy sharing or trading approaches [7, 8, 10, 21, 22, 26, 30, 34, 36, 37] in micro-grids and community shared solar PV systems that may be adapted in VPP solar energy trading systems.

Figure 3 quantifies the effectiveness of three approaches by comparing the amount of electricity drawn from the grid, the reduction in their monthly electricity bills, and the Gini Coefficient (a.k.a. Gini Index) of four different VPPs from New York, California, Colorado and Austin, respectively. We report the Gini coefficient (GC) [1], which is a measure of statistical dispersion intended to measure the degree of inequality in a solar energy trading outcome in residential VPPs. GCs range from 0 to 1, where 0 indicates a perfect equality, i.e., each VPP user is receiving an equal electricity saving from solar energy trading, and 1 indicates the worst equality, i.e., only a small group of VPP users are receiving the entire energy saving benefits. GC is discussed further in Section 6. To generate Figure 3, we use smart meter data from 119 residential DSERs which we downloaded from PecanStreet Dataport [5]. Specifically, the dataset contains 6-month data for 25 homes from New York, 5-year data for 23 homes from California, 1-year data for 46 homes from Colorado, and 1-year data for 25 homes from Austin. We pre-process the smart meter data into 1 hour granularity to benchmark the performance of the three different approaches as shown in Figure 3.

Pure Sharing. In this case, a DSER first consumes solar generated electricity locally and then stores surplus solar generated energy into its battery for future use. After the local battery is already fully charged, it starts to sell its excess solar energy to its neighbors who have deficit energy. Thus, the DSER only sends the surplus solar energy after trading with its neighbors back to the grid. We assume that all selling and buying requests are matched perfectly. In doing so, we simulate the maximum monetary benefits that users can receive from solar energy trading.

First-Come-First-Sharing (FCFS). Similar to pure sharing, FCFS consumes solar generated energy in the sequence of meeting load demand, charging battery, trading with neighbors, and feeding back to the grid. Each DSER selling energy to its neighbors maintains a demand queue. All trading requests are enqueued and met according to their arriving order.

Round-Robin (RR) Sharing. The major difference between RR and FCFS is how a DSER maintains its service request queue. RR enables DSERs to trade with their neighbors with deficit energy using a fixed amount of excess solar generated energy divided into equal portions (quanta) assigned to neighbors arranged in round-robin order. This approach handles trading requests without considering their arriving sequence.

Observation: Our results show that the aggregated electricity drawn from the grid per DSER per month in 4 different VPPs has significantly dropped after applying the 3 trading approaches, including pure sharing, RR, and FCFS approaches. Specially, as shown in Figure 3 (a), pure sharing approach reports 21.39%, 6.89%, and 12.28% (with the average of 13.52%) net demand reduction in the VPP communities of New York, California, and Austin, respectively. In addition, by leveraging more efficient solar surplus energy trading approaches of RR and FCFS, DSERs can reduce the aggregated grid demand per DSER per month by the average of 94.66% and 97.28% in the 4 regions, respectively. Figure 3 (b) also shows that users can receive 3%, 88.53%, and 92.16% average savings in their monthly electricity bills per household when applying pure sharing approach, RR, and FCFS in the four VPPs. In addition, we also observe that RR and FCFS help VPP users to achieve more fair monetary benefits. Compared to pure sharing, RR and FCFS decrease their average GC by 71.04% and 66.92% over 12 months in the 4 VPPs as shown in Figure 3 (c). Thus, these native trading approaches can help mitigate bias in solar energy trading systems.

Insights from the Above Approaches: The above results and observations shed the following important insights. First, the most recent trading approaches in the related areas of micro-grid and

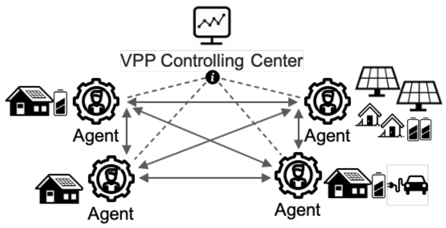


Figure 4: The system overview of the distributed trading approach in SolarTrader.

community-shared solar may have significant potential for net energy savings. Enabling the trading of surplus solar generated energy with DSER neighbors in VPPs has the potential to provide users additional monetary benefits. Second, enabling solar generated energy trading with neighbors can also mitigate unfairness in long-term energy saving outcomes observed in traditional non-sharing and pure sharing approaches. RR is the most stable approach to benchmark a trading algorithm's fairness. These valuable insights will guide the development of our proposed technique, *SolarTrader*.

3 SOLAR ENERGY TRADING CHALLENGES

In this section, we highlight the major challenges in designing and implementing SolarTrader.

Solar Energy Measurement. To enable solar energy trading in a VPP, the first challenge is to accurately describe the available amount of solar energy that may be traded. Prior solar energy trading approaches [7, 8, 10, 21, 22, 26, 30, 34, 36, 37] do not consider the practical limitations, such as battery charging/discharging loss, inverter loss, and weather conditions (primarily cloud/sky cover data relevant in the day-ahead energy market), when designing their solar energy trading techniques. The consideration of these issues are central to SolarTrader's design.

Unequal User Electricity Saving Benefits. Many existing solar energy sharing systems [8, 21, 22, 26, 30] in the areas of microgrid and community shared solar arrays, are built on top of either centralized or priority-based sharing (bound to either hardware or software) approaches. From a long-term (e.g., 1 month or 1 year) perspective, these approaches do not necessarily result in"fair" solar energy sharing among the VPP users. As noted in the released guide [14] of National Renewable Energy Laboratory (NREL), new energy program developers need to ensure all the participants have equal opportunity to join and receive benefits. SolarTrader allows VPP users to achieve long term fair energy saving by enabling concurrent and distributed trading with neighbors.

Inaccurate Net Load Forecasting. Prior approaches [7, 10, 21, 22, 26, 30, 34, 36] leverage multiple strategies to maximize a VPP user's individual monetary benefit without collaboration among the distributed DSERs in a VPP. However, these approaches can result in frequent "unexpected" fluctuations in net, aggregated electricity demand for the VPP. Furthermore, these approaches can also prevent VPP users from receiving the benefit obtained from optimal management of all the available DSERs in a VPP, and can decrease accuracy when forecasting aggregated load demand. SolarTrader leverages deep reinforcement learning-based concurrent and collaborative trading among all the DSERs to mitigate this issue.

Inaccurate Solar Generation Forecasting. Several approaches in the literature have proposed leveraging the day-ahead solar energy trading market [7, 8, 10, 21, 26, 34, 36, 37]. The solar energy generation forecasting in day-ahead market depends on many factors, such as weather conditions (primarily, sky cover), shading due to nearby objects, atmospheric conditions, generation inefficiency, and physical properties of the DSERs. The prior approaches have either not considered these factors or have not modeled the effect of those factors. To ensure the efficiency of its online solar energy trading platform, SolarTrader integrates a more accurate solar generation forecasting model.

4 SOLARTRADER

To address the challenges involved in solar energy trading, we build *SolarTrader* a new VPP solar energy trading system that enables users to maximize the benefits offered by solar energy equally while also allowing the residential VPPs to manage the grid more efficiently.

4.1 System Design

As discussed in Section 1, residential VPPs are mainly comprised of small-scale rooftop solar PV systems, community shared solar PV systems, EVs, etc. We model a residential VPP environment as a multi-agent system (MAS). As shown in Figure 4, we first gather and abstract all the DSERs into MAS agents such that each DSER can be abstracted as an agent in the collaborative VPP environment. This MAS collaborative environment enables the concurrent and distributed energy trading transactions among all DSER agents in a VPP. We then present a set of new energy trading algorithms that leverages a deep reinforcement learning approach for DSER agents to trade their surplus energy in a distributed and peer-to-peer (P2P) manner. In essence, as shown in Figure 5, we first build a discrete state space that can represent all the possible states for DSER agents in the VPPs. Note that each DSER agent must stay in a single state at any time. Then, we build a set of discrete actions for each DSER agent. Each DSER agent may choose one action from this action space based on its current and prior state. We then utilize multiple training and optimization algorithms to learn a stable and accurate reward and penalty model that can be used to evaluate the actions of each agent. Eventually, we employ a scalable hyper-parameter tuning approach to further improve SolarTrader's performance.

4.2 Building State Space

We design a discrete representation for each agent state, including local agent states and global VPP states. The local agent states are designed to track the dynamics of each agent's internal states, such as load demand, energy consumption, solar generation, and battery storage capacity. The VPP environment states aim at providing each agent with a real-time perspective of the entire VPP environment, which may include aggregated VPP energy consumption and Gini Coefficient among all DSERs. Specifically, the state S_i^k of agent k at the time i can be defined as a 6-tuple of tensors,

- **Energy Consumption**. A tensor representing each agent's net load demand which is the total electricity demand in the VPPs minus the solar-generated electricity. This demand is met by a VPP using its traditional electricity generation sources, such as

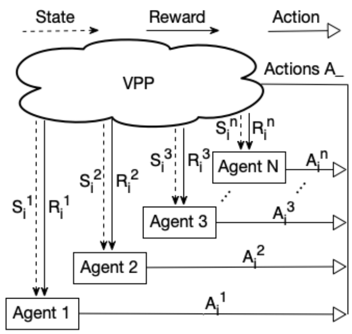


Figure 5: An overview of the multi-agent deep reinforcement learning process in SolarTrader.

natural gas, hydropower, and imported electricity from outside SolarTrader.

- **Net Load Demand**. A tensor representing each agent's actual energy consumption—the total electric demand in the system. This energy consumption is provided by VPPs using a mix of traditional and renewables energy sources.
- **Solar Generation**. A tensor representing each agent's actual solar generation. This energy is generated by the installed solar PV arrays and determined by many local physical characteristics, such as tilt, size, orientation, shade from nearby trees or tall buildings, cloud/sky cover, inverter inefficiency etc.
- **Battery Capacity Level**. A tensor representing each agent's local battery energy storage level. The battery size may vary at different agents. For the community shared battery array users, this tensor represents only their subscribed portions.
- **Aggregated Grid Energy Consumption**. A tensor representing the global VPP environment stability using the grid net energy consumption. This tensor may include the variance or standard deviation of the net energy consumptions. The goal of VPP maintenance is to reduce peaks or spikes in this net consumption curve by making more efficient use of all the available DSERs.
- **DSERs' Fairness**. A tensor representing the DSERs' fairness when participating in solar energy trading processes of SolarTrader. This tensor is used to improve and ensure the long term fairness of trading outcomes.

We observed that by separating the status information into multiple channels rather than mixing them in a single tensor, our DSER agents learn faster and achieve a more stable policy. Note that recording the above states do not require any specialized hardware or software to be installed by VPP users. These states can be monitored and reported at a fine-grained level by smart energy meters (a.k.a net meters) that are already deployed by the utilities.

4.3 Building Action Set

We next build the discrete action set for each DSER agent. Specifically, in the VPP solar energy trading environment, each agent a_k has the following multiple choices to manage and optimize its solar energy trading perspectives at moment t,

- **Action 1: Consuming Solar Generated Energy**. A tensor representing the DSER agent a_k consumes solar generated energy to meet its local electricity demand.
- **Action 2: Charging Battery**. A tensor representing the DSER agent, a_k has surplus solar generated electricity after meeting its local demand. The local battery is not fully charged, and agent a_k is going to charge its battery until full for future usage.
- **Action 3: Discharging Battery to Meet Local Demand**. A tensor representing the DSER agent a_k has decided to discharge its battery stored energy to meet its local electricity demand.
- **Action 4: Discharging Battery to Trade with Neighbors**. A tensor representing the DSER agent a_k has decided to discharge its battery stored energy to lend to its neighbors. Note that at nighttime when solar energy is unavailable, the DSER agent a_k should not perform this action since storing solar energy for future usage is always providing the DSER agent with the best monetary benefit which is equivalent to retail price value.
- **Action 5: Trading Solar Energy by Lending to a Neighbor**. A tensor representing the DSER agent a_k is having excess solar generated energy and decide to lend this surplus energy to its neighbors at moment t. In this case, the solar generated electricity and the battery stored energy has already meet the demand of the DSER agent a_k. Thus, the DSER agent a_k is offering the surplus solar generated energy to its neighbor DSER agent a_l which has deficit energy. Note that the DSER agent a_k may make offers to multiple deficient neighbors simultaneously.
- **Action 6: Trading Solar Energy by Borrowing from a Neighbor**. A tensor representing the DSER agent a_k is having positive deficit energy at moment t. In this case, the solar generated electricity and the battery stored energy can not meet the demand of the DSER agent a_k. Thus, the DSER agent a_k is requesting energy from another neighbor agent a_l which has surplus solar energy to trade. Note that the DSER agent a_k may send requests to multiple neighbors simultaneously.
- **Action 7: Feeding Excess Solar Generated Energy to the Grid**. A tensor representing the DSER agent a_k is feeding surplus energy back to the grid. In this case, the DSER agent a_k's battery is fully charged, and none of its neighbors are requesting energy trading. The DSER agent a_k will receive a green credit in its next monthly electricity bill which is paid based on the wholesale electricity price value.
- **Action 8: Drawing Electricity from the Grid**. A tensor representing the DSER agent a_k is to draw electricity from the grid at moment t. In this case, after using solar generated energy and the traded solar energy from its neighbors, the DSER agent a_k is still having deficit energy situation. Drawing electricity from the grid will be charged according to the retail electricity price.

4.4 Learning Reward and Penalty

We next build and learn an optimal reward and penalty model. The objective function of SolarTrader at each DSER agent is to minimize the energy costs of its operations over a period of time T, and can be formulated as min-R_{cost} at agent k as follows:

$$\begin{aligned} R_{cost_k} = \frac{1}{T} \cdot \sum_{t=0}^{T} (E_{grid_k}(t) \cdot P_{grid_k}(t) - E_{trade_k}(t) \cdot \\ P_{trade_k}(t) - E_{feedback_k}(t) \cdot P_{feedback_k}(t)) \end{aligned} \tag{2}$$

where $E_{grid_k}(t)$, $E_{trade_k}(t)$, and $E_{feedback_k}(t)$ describe the electricity amount of the agent k at moment t draws from the grid, trades with its neighbors, and feeds back to the grid, respectively. $P_{grid_k}(t)$, $P_{trade_k}(t)$, and $P_{feedback_k}(t)$ denote the paid electricity price in \$ per kWh when the agent k is drawing from the grid, trading with its neighbors, and feeding back to the grid, respectively. Note that $P_{trade_k}(t)$ would be a negative value when the agent k is requesting energy from its neighbors. Similarly, $P_{trade_k}(t)$ should be a positive value when the agent k is lending its surplus energy to its neighbors.

Algorithm 1: The Trading Algorithm of SolarTrader

Input: State S, Action A, Reward R
Output: Optimal Trading Policy P
Data: State S, Action A, Reward R
`/* Load Action Space, State Space */`
load tensor vector : $Action$, $State$ ← SolarTrader
while *true* **do**
  $State_k$ ← the VPP environment
  $Action_k$ ← getLegalAction($State_k$)
  if $Random_{number} \geq \epsilon$ **then**
    $action$ ← RandomPick($Action_k$)
  else
    $action$ ← BestActionFromPolicy($Action_k$, $R_(Action_k)$)
  $State_{next}$ ← takeAction($action$)
  `/* Using PPO, PG, APPO, DPPG to train & optimize reward */`
  $Reward_{optimal}$ ← SelectOptimalReward(PPO, $action$)
  `/* Fine-turning parameters using grid search */`
  tuningGridSearch($State_k$, $Action_k$, $Reward_{optimal}$)
  p ← AgentLearning($State_k$, $Action_k$, $Reward_{optimal}$)

Because DSER agents in SolarTrader are designed to be cooperative, we develop a global shared reward function. This global shared objective finds the minimum energy cost of operations across the entire VPP environment, and can be defined as follows:

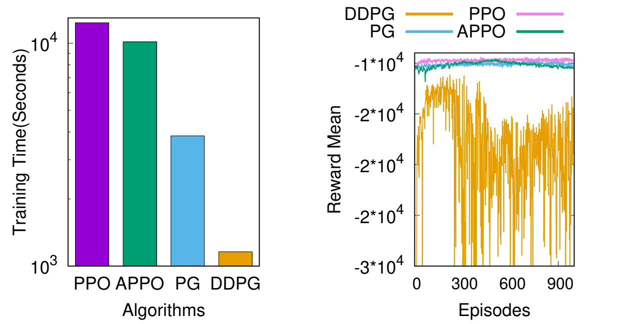
$$R_{all} = \alpha \cdot \frac{1}{N} \cdot \sum_{k=0}^{N} R_{cost_k} + \beta \cdot Var + \gamma \cdot GC \tag{3}$$

where N denotes the amount of agents, Var and GC denote the variance in the aggregated net grid energy consumption and the fairness metric represented by the Gini Coefficient (GC) of monthly cost saving, respectively, across all the DSER users. α, β, and γ are discount factors.

In the reward learning process of SolarTrader, we observe multiple sub-optimal reward outcomes. For instance, a DSER agent is requesting electricity from both the grid and its neighbors when its neighbors are still having surplus solar generated energy. This occurs due to the nature of SolarTrader's distributed and concurrent trading actions. To address such issues, we develop the following penalty function for each reward learning episode:

$$P_{cost_k} = \begin{cases} \lambda \cdot \left|R_{cost_k}\right|, & R_{cost_k} < 0 \\ \mu \cdot R_{cost_k}, & R_{cost_k} > 0 \\ R_{cost_k}, & R_{cost_k} = 0 \end{cases} \tag{4}$$

Here, λ and μ denote the weights of encouraging and penalizing the actions for the DSER agent k. We leverage the policy gradient methods and the grid search approaches discussed in Section 4.5 and Section 4.6 to learn the optimal values for the discount factors.



(a) Training Time (b) Reward Mean

Figure 6: The comparison of 4 different training algorithms in multi-agent reinforcement learning of SolarTrader.

4.5 Training and Optimizing Policy

We then train all the DSER agents using multiple widely used policy gradient methods, including Policy Gradients (PG) [35], Proximal Policy Optimization (PPO) [31], Asynchronous Proximal Policy Optimization (APPO) [25], and Deep Deterministic Policy Gradients (DDPG) [20], to compute and optimize the optimal policy of the solar generated energy trading approach in SolarTrader. PG [35] describes its policy explicitly using function approximators and is also independent of specific value functions. The policy parameters are updated based on the observation of expected reward gradients. PPO [31] is a policy gradient method that optimizes a "surrogate" objective function using stochastic gradient ascent (SGD). In addition, PPO's clipped objective supports multiple SGD passes over the same batch of experiences. APPO [25] is adapting both policy gradient and Q—value learning algorithms to learn using many parallel simulator instances asynchronously. DDPG [20] is a model-free deterministic policy gradient method that is mainly applied in continuous action space problems. And it has combined the benefits from both the actor-critic approaches and Deep Q Network (DQN). Figure 6 (a) and (b) show the comparison results of applying different policy gradient methods for SolarTrader. Interestingly, the policy gradient methods—DDPG and PG achieved the shortest training time, while, PG and PPO yielded the highest average reward mean after training over ~100 episodes. Thus, PG is the best overall performing policy gradient method regrading training time and average reward mean, and is eventually selected to help SolarTrader train and optimize its reward policy.

4.6 Tuning Hyper-parameters

The performance of DSER agent policies are sensitive to the hyperparameter values chosen. Unfortunately, there is no simple approach that allows DSER agents to understand whether a specific value for a given parameter would improve total reward. To address this issue and further increase SolarTrader's performance, we leverage a tuning approach to optimize the SolarTrader's hyperparameters, such as such as discount factors associated with rewards and penalties, and the learning rates. In particular, we employ grid search which allows us to specify the range of values to be considered for each hyper-parameter. The grid search process constructs and evaluates our model using every combination of the hyper-parameters. Finally, we employ cross-validation to evaluate each learned model.

5 IMPLEMENTATION

We implement both a SolarTrader simulator and prototype in `python` using widely available open-source frameworks, including osBrain [23], Pandas [24], Scikit-learn [32] and PyCUDA [15, 28]. The simulator takes smart meter energy traces as input and applies solar energy trading techniques outlined in Section 2. In addition to the simulator, we also build a prototype SolarTrader using the scalable reinforcement learning framework, RLlib [29]. This prototype is able to trade solar energy in real-time using online SolarTrader. We employ a Multi-agent Deep Reinforcement Learning setup, which enable SolarTrader to train, learn and optimization a distributed trading model. The prototype uses actual smart meter data of 119 homes from 4 different U.S. VPP communities. In addition, SolarTrader queries the real-time eGauge [16] smart meter data for an entire VPP in Massachusetts every minute. We implement SolarTrader's algorithms and its optimizations. We use the scalable deep reinforcement learning `python` library RLlib [29] to develop our energy trading solutions in SolarTrader. RLlib supports TensorFlow, TensorFlow Eager, and PyTorch. RLlib provides multi-ways for us to customize the training process of the target environment modeling, neural network modeling, action set building and distribution, and optimal policy learning. We leverage training algorithms, including Deep Deterministic Policy Gradients (DDPG), Asynchronous Proximal Policy Optimization (APPO), Proximal Policy Optimization (PPO), and Policy Gradients (PG), to train and fine-tune a stable global reward policy. Eventually, we schedule the batch jobs on our server to compare the performance of all the 9 different solar energy trading approaches. The evaluation testbed machine configuration includes 2x Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz, an nVidia TITAN X (Pascal) (x8) GPU, and 128GB of DRAM, and runs CentOS 7 Linux.

6 EXPERIMENTAL EVALUATION

Below we describe our datasets, experimental setup, metrics used to evaluate our approaches, and evaluation results.

6.1 Datasets

Dataset 1. We downloaded 1 minute level circuit-level electricity usage and solar generation from 119 rooftop solar-powered houses using Street Dataport [5]. For each solar-powered home listed in this dataset, we can access its metadata data including street address, floor plan, solar generation capacity, and installation details. Specially, this dataset contains 6 months of energy data for 25 homes in New York, 5 years energy data for 23 homes in California, 1 year data for 46 homes in Colorado, and 1 year data for 25 homes in Texas. We preprocess the minute-level energy data to generate hour-level data for analysis and evaluation.

Dataset 2. We also download minute-level netmeter recorded energy data of 100 residential DSERs for 1 year from the dataset which is released by the most recent solar forecasting work [12, 13]. This dataset has four dimensions including timestamps, net demand, ground-truth solar generation, predicted solar generation using the models in work [12, 13]. In addition, this dataset contains detailed metadata of the physical characteristics for each solar site, including size, orientation, tilt, and inverter information.

Dataset 3. We also use live smart meter data from 10 solar-powered residential DSERs from eGuage [16] to evaluate SolarTrader and other trading approaches. This online VPP dataset has detailed metadata of the physical characteristics for each DSER, including solar array size, orientation, tilt, inverter inefficiency, and etc.

6.2 Experimental Setup

We implement and compare 9 different solar energy trading approaches, including non-sharing, pure sharing, Round Robin (RR), First Come First Sharing (FCFS), Shortest Remaining First (SRF), Longest Remaining First (LRF), Longest Sharing First (LSF), Shortest Sharing First (SSF), and the proposed SolarTrader approaches.

6.3 Evaluating Metrics

Gini Coefficient (GC, a.k.a. Gini Index). To access the fairness of our implemented approaches, we employ GC as the measurement of statistical dispersion intended to describe the inequality of VPP user energy trading benefits in their solar energy trading. Note that the GC value is between 0 (indicating a perfect equality) and 1 (indicating the worst equality). It can be formally defined as follows:

$$GC = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| X_i - X_j \right|}{2 \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} X_j} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left| X_i - X_j \right|}{2N \cdot \sum_{i=1}^{N} X_i} \tag{5}$$

where N denotes the amount of agents in the VPP and X_i represents the the energy saving at the agent i.

Mean Absolute Percentage Error (MAPE). To quantify the accuracy of net load forecasting after applying our different energy trading approaches, we compute and use the MAPE between the ground truth net load consumption and that of the different approaches, over all time intervals t. A lower MAPE indicates higher accuracy with a 0% MAPE being perfectly accurate net load demand profiling and forecasting. In doing so, we are examining the stability of the net load forecasting of the VPPs when employing different solar surplus energy trading approaches. MAPE is formally computed as follows:

$$MAPE = \frac{100}{n} \sum_{t=0}^{n} \left| \frac{S_t - P_t}{S_t} \right| \tag{6}$$

where n describes the duration of net load demand S_t, and P_t indicates the predicted net load demand at moment t.

Dicke—Fuller (DF) Test. Time series net demand in smart grid typically follows a combination of deterministic trend and stochastic trend. The Dickey—Fuller (DF) test aims at testing the null hypothesis that assumes a unit root is present in a given time-series dataset. This unit root is the measure that can be leveraged to test whether a times series data is stationary, trend-stationary, or predictable. We employ the DF test to quantify a solar energy trading approach's performance as it informs VPP's peak power reduction and net demand forecasting. The net grid demand can be formally described using the following model:

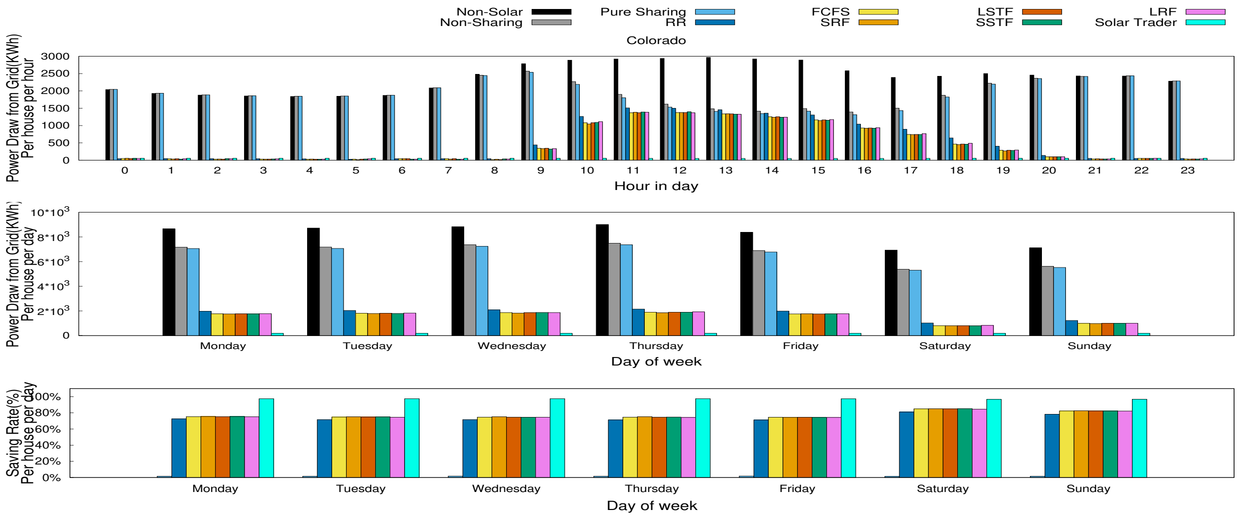


Figure 7: Energy drawn from the net grid over hours in a day (top) , energy drawn from the grid over days of a week (middle), energy saving from the grid over hours in a day (bottom).

$$N_t = \alpha + \beta \cdot t + \phi \cdot N_{t-1} + e_t \quad (7)$$

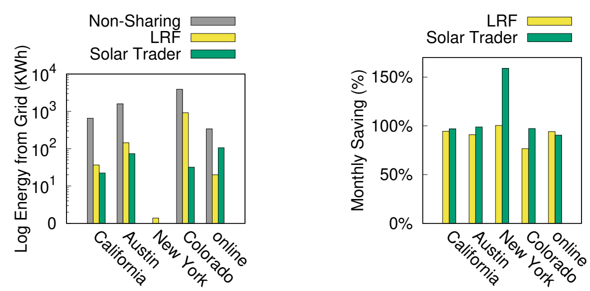
where N_t denotes the net demand of the VPP at moment t, α, β, and ϕ represent the coefficients of net load demand modeling, and e_t is the error term at moment t. Thus, we can have the DF test defined as follows:

$$\Delta N_t = N_t - N_{t-1} = \alpha + \beta \cdot t + \gamma \cdot N_{t-1} + e_t \quad (8)$$

where Δ is the first difference operator, N_t and N_{t-1} denote net demand at moments t and $t-1$, α, β, and γ are the coefficients of net demand modeling, and e_t is the error term at moment t.

6.4 Experimental Results

6.4.1 Quantifying SolarTrader's performance on the reduction of aggregated VPP energy consumption drawn from the grid. As we explained in Section 6.2, we compare SolarTrader with other 9 different solar energy trading approaches. In this case, we are evaluating SolarTrader's performance on reducing the aggregated VPP energy consumption. Unsurprisingly, as shown in Figure 7 (top), all the trading approaches are reporting reduction in their aggregated net energy drawn from the grid over the hours of a day in Colorado VPP. In particular, SolarTrader stands out as the the best performing approach that reduces aggregated VPP energy consumption per-hour across the entire year. Figure 7 (middle) shows the breakdown of the aggregated electricity consumption over the days of a week in the same VPP. We observe the similar result that SolarTrader is always the best performing approach that reduces aggregated VPP energy consumption everyday in a week in the whole year. Note that for this VPP in Colorado, solar energy trading happens between the local sunrise and sunset time in a day, roughly from 8 am to 8:30 pm in 2019 when solar generation is available. In addition, due to the "nature" and daily routines that the residents are consuming more electricity in daytime from 7 am to 10 pm than other times when users are typically sleeping in a day, we also observe that aggregated energy consumption after applying solar



(a) Aggregated energy consumption (b) Monthly saving

Figure 8: The comparison results of LRF and SolarTrader in 4 different U.S. residential VPPs and online.

energy trading approaches is proportional to the net consumption when assuming none of the VPP users have any solar deployment (shown as the longest dark bar in Figure 7 (top)). Eventually, in addition to Colorado VPP, we also examine SolarTrader's performance on the reduction of aggregated VPP energy consumption drawn from the grid in 3 other residential VPPs from California, New York, and Texas. Figure 8 (a) shows the comparison results of these 4 different VPPs and online VPP. SolarTrader is consistently the best performing solar energy trading approach across all the 4 different regions. In addition, SolarTrader resulted in ~538 kWh energy feeding back to the grid, rather than requiring electricity from the grid in New York.

Results: *Comparing with non-sharing approach and the other 8 current solar energy trading approaches, SolarTrader is the best performing approach everyday in the year, and it has resulted in significant average reduction of 83.8% in the aggregated net energy drawn from the grid across 4 different U.S. residential VPPs and online VPP.*

6.4.2 Quantifying SolarTrader's performance on the saving of residents' monthly electricity bills. As we explained in Section 6.2, we compare SolarTrader with 8 other trading approaches with respect to the VPP users' monthly electricity savings. In this case,

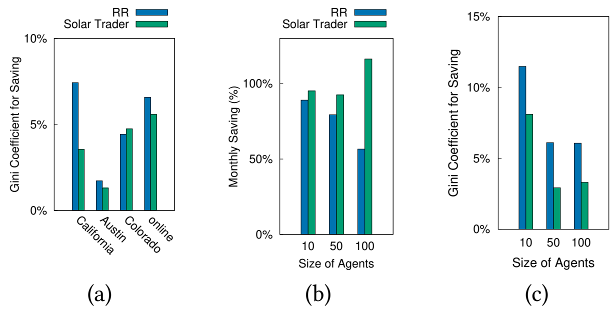


Figure 9: The comparison results of RR and SolarTrader.

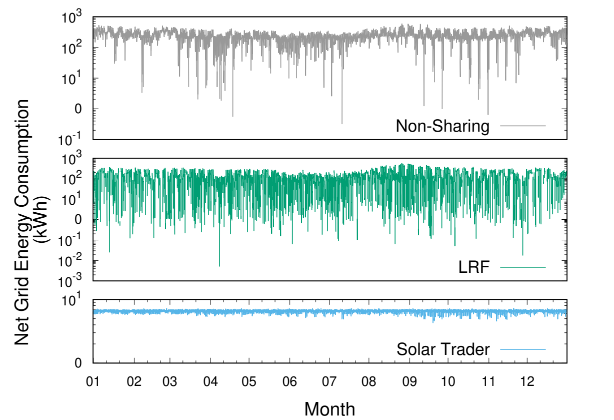


Figure 10: A typical 12-month net grid energy consumption before applying any solar sharing approaches (Top) and after applying LRF (Middle) and SolarTrader (Bottom).

we are evaluating how a DSER user can benefit from trading its solar energy with neighbors. Figure 7 (bottom) shows the statistical breakdown of monthly electricity savings across days of a week. The results show that SolarTrader enables DSER users in the VPPs to achieve the most significant savings in their monthly electricity savings on every day of a week. We next examine the performance of various approaches with respect to the VPP users' monthly energy savings using 4 residential VPPs from 4 different states. As shown in Figure 8 (b), SolarTrader achieves average savings of 109.8% across the 4 different states, and thus is the best performing trading approach in these residential VPPs. Interestingly, in addition to saving energy for the residents in the New York VPP, SolarTrader has resulted in an additional 50% of the monthly energy consumption of non-sharing approach to be sent back to the grid. Thus, SolarTrader allows the VPP to feed more surplus solar energy back to the grid and achieves more benefits for the VPP users.

Results: *Comparing with non-sharing approach and 8 other existing solar energy trading approaches, SolarTrader is the best performing approach everyday in the year, achieving the significant average saving—109.8% in monthly electricity consumption across 4 different residential U.S. VPPs and online VPP.*

6.4.3 Quantifying SolarTrader's performance on the trading fairness for long term. We next examine SolarTrader's performance on the fairness in long term trading process. Figure 9 (a) reports the Gini Coefficient (GC) of RR and SolarTrader in 3 different residential VPPs using the same dataset as Figure 8. Note that based on the insight we learn from Section 2, RR is the best performing approach regarding to the fairness of a solar energy trading approach. As we can see from Figure 9 (a), SolarTrader yields the similar GC to that of RR approach across 3 VPPs of California, Texas, and Colorado.

Results: *SolarTrader yields a GC similar to that of RR approach across 3 different U.S. VPPs. In addition to reducing aggregated VPP energy consumption and increasing monthly savings simultaneously, SolarTrader enforces the best fairness of solar energy trading.*

6.4.4 Quantifying SolarTrader's scalable performance. Next, we examine the trading performance effect on SolarTrader and other trading approaches when using the VPPs that have different amount of DSERs. In doing so, we are evaluating the scalability of SolarTrader. As shown in Figure 9, the larger amount of the DSERs in a VPP has resulted in a better overall trading performance. Specially, Figure 9 (b) shows that SolarTrader can achieve higher efficient saving in VPP user monthly electricity energy savings when the number of the DSERs increase from 10 to 100, and that it is always significantly better than RR approach. As shown in Figure 9 (c), SolarTrader also reduces GC by ~75% when the amount of the DSERs increasing from 10 to 100 over 12 months. Therefore, SolarTrader is able to achieve significantly higher monthly electricity cost savings and lower GCs as the the VPP size increases.

Results: *Larger VPP size has resulted in higher monthly electricity cost savings and lower GCs of SolarTrader. With the growing of VPP size, SolarTrader achieves better overall trading performance regarding to electricity cost saving and long term trading fairness.*

6.4.5 Quantifying SolarTrader's improvement to grid operation. Next, we examine the trading performance effect on demand forecasting of SolarTrader and LRF approach. In doing so, we benchmark the ability of SolarTrader to assist the grid in managing its operations. Figure 10 shows the 1 year comparison results of aggregated net grid consumption after applying SolarTrader (Figure 10 (bottom)) and LRF approach (Figure 10 middle) which is the best performing approach regarding to reducing aggregated energy consumption. We benchmark each approach using Dicke—Fuller (DF) Test. We find that non-sharing, LRF, and SolarTrader yield their test statistic values as -8.16, -9.76, and -11.56, respectively. And these test statistic values are much smaller than the significant critical value (5%) as -2.862. Thus, we can reject the null hypothesis that the aggregated entire VPP energy consumption is not stationary with respect to its forecasting. The more negative test statistic value, the stronger for us to reject this null hypothesis. Thus, SolarTrader-enabled VPP achieves the best net demand forecasting performance.

Results: *A SolarTrader-enabled VPP maximally improves the grid operation performance with respect to its net demand forecasting.*

7 RELATED WORK

There is no signifiant prior work on solar energy trading in VPPs. However, there is significant work in the related ares of microgrid [18, 19, 38] and community shared solar arrays (CSSA) [11, 14]. Several prior approaches in the literature do not allow VPPs users to achieve their maximum benefits from the DSERs such as rooftop solar PV arrays, CSSA, and other solar energy storage systems [7, 8, 10, 21, 22, 26, 30, 34, 36, 37]. The primary reason for this handicap is that these solar energy sharing approaches do not actually allow DSER users to trade their surplus solar generated energy concurrently, and typically require a trusted third-party to play the role of middleman to supervise the trading processes. In addition, due to a lack of fair trading algorithms, these approaches do not actually result in long term "fair" solar energy trading among all the VPP users. In contrast, SolarTrader enables fully distributed, scalable, and long term fair solar energy trading in residential VPPs. SolarTrader leverages multi-agent deep reinforcement learning approach to automatically learn a collaborative policy in a shared VPP environment, which can be deployed on DSER agents without extra training. In addition, the policy is trained on P2P system architecture and executed in a distributed manner such that all the DSER agents can trade their solar generated energy concurrently. Through reward and penalty learning, SolarTrader also ensures solar energy trading fairness across all DSER agents in long term.

8 CONCLUSION

SolarTrader is a new solar energy trading system that enables unsupervised, distributed, and long-term fair solar energy trading in residential VPPs. In essence, it leverages a new multi-agent deep reinforcement learning approach that enables P2P and distributed solar energy trading among different DSERs to ensure both VPP users and the grid to maximize benefits simultaneously. We implemented SolarTrader and evaluated it using both real netmeter data from 4 U.S. residential communities that are comprised of ~219 residential DSERs and online smart meter data from 10 residential DSERs. Our results show that SolarTrader enables DSER users to achieve the significant average saving — as much as 109.8% — in their monthly electricity consumption, while also enforcing the most fair trading process simultaneously. In addition, a SolarTrader-enabled VPP experiences an average reduction of 83.8% in the aggregated net energy drawn from the grid. We plan to deploy a prototype of SolarTrader in two community shared solar PV systems to further understand the benefits of our proposed approach. In addition, we will design new policies and expand the current DSER agent definition to abstract and integrate more DSERs into SolarTrader, such as EVs, utility battery banks, and wind generated energy. Finally, we plan to integrate SolarTrader with the most recent solar generation forecasting model to enable days-ahead trading.

Acknowledgements. This research is supported by Cyber Florida Collaborative Seed Program.

REFERENCES

[1] 2018. Gini Coefficient. https://en.wikipedia.org/wiki/Gini_coefficient.
[2] 2020. Dominion Energy Green Power program. https://www.dominionenergy.com/home-and-small-business/renewable-energy-programs/solar-purchase-program/solar-purchase-program-faqs.
[3] 2020. How We Compensate You for Excess Power Generation. https://www.sdge.com/residential/savings-center/solar-power-renewable-energy/net-energy-metering/billing-information/excess-generation.
[4] 2020. PayPal. https://www.paypal.com/us/.
[5] 2020. Pecan Street Data: Circuit-level Electricity Use and Generation. https://www.pecanstreet.org/dataport/.
[6] 2020. Venmo. https://venmo.com/send-receive/start/.
[7] Nurzhan Zhumabekuly Aitzhan and Davor Svetinovic. 2016. Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing* 15, 5 (2016), 840–852.
[8] Muhammad Raisul Alam, Marc St-Hilaire, and Thomas Kunz. 2019. Peer-to-peer energy trading among smart homes. *Applied energy* 238 (2019), 1434–1443.
[9] G. Barbose, S. Weaver, and N. Darghouth. 2014. *Tracking the Sun VII: An Historical Summary of the Installed Price of Photovoltaics in the United States from 1998 to 2013.* Technical Report. Lawrence Berkeley National Laboratory.
[10] Riccardo Bonetto, Ilya Sychev, and Frank HP Fitzek. 2018. Power to the Future: Use Cases and Challenges for Mobile, Self Configuring, and Distributed Power Grids. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm).* IEEE, 1–6.
[11] Gabriel Chan, Isaac Evans, Matthew Grimley, Ben Ihde, and Poulomi Mazumder. 2017. Design choices and equity implications of community shared solar. *The Electricity Journal* 30, 9 (2017), 37–41.
[12] Dong Chen, Joseph Breda, and David Irwin. 2018. Staring at the sun: a physical black-box solar performance model. In *Proceedings of the 5th Conference on Systems for Built Environments.* ACM, 53–62.
[13] Dong Chen and David Irwin. 2017. Black-box solar performance modeling: Comparing physical, machine learning, and hybrid approaches. *ACM SIGMETRICS Performance Evaluation Review* 45, 2 (2017), 79–84.
[14] community 2012. A Guide to Community Shared Solar. https://www.nrel.gov/docs/fy12osti/54570.pdf.
[15] Cuda 2020. An Even Easier Introduction to CUDA. https://devblogs.nvidia.com/even-easier-introduction-cuda/.
[16] Egauge 2020. eGauge Energy Monitoring Solutions. http://www.egauge.net/.
[17] electricity 2016. U.S. Energy Information Administration, Electricity Monthly Update. https://www.seia.org/solar-industry-research-data.
[18] Jinwei He and Yun Wei Li. 2012. An enhanced microgrid load demand sharing strategy. *IEEE Transactions on Power Electronics* 27, 9 (2012), 3984–3995.
[19] Robert H Lasseter and Paolo Paigi. 2004. Microgrid: A conceptual solution. In *2004 IEEE 35th Annual Power Electronics Specialists Conference (IEEE Cat. No. 04CH37551)*, Vol. 6. IEEE, 4285–4290.
[20] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
[21] Chao Long, Jianzhong Wu, Yue Zhou, and Nick Jenkins. 2018. Peer-to-peer energy sharing through a two-stage aggregated battery control in a community Microgrid. *Applied energy* 226 (2018), 261–276.
[22] Alexandra Lüth, Jan Martin Zepter, Pedro Crespo del Granado, and Ruud Egging. 2018. Local electricity market designs for peer-to-peer trading: The role of battery flexibility. *Applied energy* 229 (2018), 1233–1243.
[23] osbrain 2020. osBrain. https://osbrain.readthedocs.io/en/stable/.
[24] Pandas 2020. Pandas. https://pandas.pydata.org/.
[25] Aleksei Petrenko, Zhehui Huang, Tushar Kumar, Gaurav Sukhatme, and Vladlen Koltun. 2020. Sample Factory: Egocentric 3D Control from Pixels at 100000 FPS with Asynchronous Reinforcement Learning. *preprint arXiv:2006.11751* (2020).
[26] Manisa Pipattanasomporn, Murat Kuzlu, and Saifur Rahman. 2018. A Blockchain-based Platform for Exchange of Solar Energy: Laboratory-scale Implementation. In *2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE).* IEEE, 1–9.
[27] Danny Pudjianto, Charlotte Ramsay, and Goran Strbac. 2007. Virtual power plant and system integration of distributed energy resources. *IET Renewable power generation* 1, 1 (2007), 10–16.
[28] Pycuda 2020. PyCUDA. https://mathema.tician.de/software/pycuda/.
[29] Rlib 2020. RLib: Scalable Reinforcement Learning. https://docs.ray.io/en/master/rllib.html.
[30] Eleonora Riva Sanseverino, Maria Luisa Di Silvestre, Pierluigi Gallo, Gaetano Zizzo, and Mariano Ippolito. 2017. The blockchain in microgrids for transacting energy and attributing losses. In *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData).* IEEE, 925–930.
[31] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2020. Proximal policy optimization algorithms. *preprint arXiv:1707.06347* (2020).
[32] scikit 2020. Scikit-learn Machine Learning in Python. https://scikit-learn.org/stable/.
[33] SolarTrader 2020. Solar, Trader. https://github.com/cyber-physical-systems/SolarTrader.
[34] Wencong Su, Zhiyong Yuan, and Mo-Yuen Chow. 2010. Microgrid planning and operation: Solar energy and wind energy. In *IEEE PES General Meeting.* IEEE.
[35] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems.* 1057–1063.
[36] Luke D Watson and Jonathan W Kimball. 2011. Frequency regulation of a microgrid using solar power. In *2011 Twenty-Sixth Annual IEEE Applied Power Electronics Conference and Exposition (APEC).* IEEE, 321–326.
[37] Ting Zhu, Zhichuan Huang, Ankur Sharma, Jikui Su, David Irwin, Aditya Mishra, Daniel Menasche, and Prashant Shenoy. 2013. Sharing renewable energy in smart microgrids. In *2013 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS).* IEEE, 219–228.
[38] Ting Zhu, Sheng Xiao, Yi Ping, Don Towsley, and Weibo Gong. 2011. A secure energy routing mechanism for sharing renewable energy in smart microgrid. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm).* IEEE, 143–148.