Applied Energy 297 (2021) 117129

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Peer-to-peer energy sharing with battery storage: Energy pawn in the smart grid

Li He, Yuanzhi Liu, Jie Zhang *

The University of Texas at Dallas, Richardson, TX 75080, USA

ARTICLE INFO

Keywords:
Dynamic pricing
Distributed solar
Energy storage
P2P sharing
Q-learning

ABSTRACT

This paper proposes a peer-to-peer (P2P) energy trading framework, allowing distributed photovoltaic (PV) prosumers and consumers to participate in a community sharing market established by a stakeholder, i.e., an energy pawn (EP). The EP is responsible for installing, connecting, managing, and maintaining the specific P2P sharing network, and possesses a publicly accessible battery energy storage (ES) system that can be used to facilitate the energy sharing within the community. A hierarchical P2P sharing market infrastructure is considered, where the interactions among the EP, prosumers, and consumers are modeled by a leader–follower framework. The EP is responsible for i) optimizing the capacity scheduling of the ES system based on forecasting-based rolling-horizon decision-marking, and ii) determining the selling and buying prices within the market. Meanwhile, prosumers and consumers will adjust their energy consumption as response to different sharing prices for maximizing consumption satisfactions based on their utility functions. With the framework, both PV prosumers and consumers can trade with the EP to balance their excess solar generation or insufficient demand to reduce electricity bills. A dynamic pricing algorithm is proposed for EP to determine the internal buying and selling prices simultaneously, and Q-learning is employed to solve the proposed hierarchical decision-making problem. An energy sharing case with 10 agents is studied to validate the effectiveness in terms of the economic benefits and PV sharing enhancement, as well as the reduction of the negawatt fed back into the grid. This study serves to provide a promising win-win-win solution for the utility grid, EP, and P2P market agents.

1. Introduction

As the penetration of distributed energy resources (DERs) such as rooftop photovoltaic (PV) increases, power supplied by distributed generators is anticipated to exceed local consumption demands in certain hours of a day. This fast-growing trend creates a potential to feed negawatt power [1] back into the grid, raising unexpected challenges to power system reliability. One innovative way to address this issue is to establish a peer-to-peer (P2P) energy market [2], in which the prosumers and consumers are able to share their excess resources, such as rooftop PV generation, flexible demand, unused capacity in energy storage (ES) systems, etc. This novel trading mode serves to benefit the whole community by reducing the community's reliance on the main grid, increasing the community resilience, and enhancing the ability to participate in demand response (DR) for achieving better generation-load balance [3].

Extensive explorations have been conducted in the literature on designing and evaluating P2P energy sharing, and game theory has been widely adopted to address the interaction between different agents within the P2P market in recent years. For example, a Nash equilibrium-based game-theoretic approach was applied in energy sharing between storage units [4], multiple prosumers and consumers [5,6], aggregators and DER owners [7], electric vehicles [8], etc. The coalition game has also been widely used in a P2P market, with a target of promoting energy sharing, either in the absence of ES devices [9,10], individual owned [11] or jointly invested ES systems [12]. Moreover, a leader–follower Stackelberg game was utilized to address the interactions between microgrids and a utility grid [13], a combined heat and power community [14], PV and ES sharing between apartments [15], residential units and shared facility controllers [16]. Besides, in a complicated P2P market consisting of multiple agents and stakeholders, hierarchical game-theoretic models [17] could be formulated to solve the conflict between different parties.

It is observed from the aforementioned studies that an appropirate pricing model plays a crucial role in P2P sharing, which directly imposes impacts on P2P participants' incentives of energy sharing and price-based DR. Auction is widely used for P2P market clearing [18],

* Corresponding author.
E-mail address: jiezhang@utdallas.edu (J. Zhang).

https://doi.org/10.1016/j.apenergy.2021.117129
Received 5 March 2021; Received in revised form 7 May 2021; Accepted 21 May 2021

such as one-side auction [19,20], double auction [4,6,16], and reserve auction [21]. However, the auction platform may be complicated in practical implementation, since most of the auction markets rely on a trusted authority or notary architecture, which causes an extra cost of platform construction. And other problems arise when the third party influences the auction proceedings, since it obtains more asymmetric information compared with other auction participants, resulting in an unfair auction process. As an alternative system, a hierarchical community energy market, consisting of a third-party stakeholder (market operator), prosumers, and consumers, has been extensively evaluated in many existing studies to bridge this gap [14,22,23]. To further advance the P2P sharing concept, in these studies, an extra P2P sharing connection is proposed to be installed, managed, and maintained by an operator who possesses its own generators such as DER resources or ES systems; the operator trades energy with the P2P participants at a customized rate that is more cost-competitive compared with the utility price.

With the ever-increasing influence of DERs on the grid, ES has been proven to be an effective solution to the integration of distributed PV in a community, due to its prominent flexibility and fast-response feature [24]. However, installing a large number of distributed ES systems in individual households complicates the demand side response and increases the overall capital cost. The concept of storage sharing [25,26], which requires a joint investment from the prosumers, can effectively reduce the energy costs, avoid unnecessary investments of prosumers, and lead to a higher ES utilization rate. However, not all prosumers would be interested in owning or investing a joint-invested ES system if the benefit of installing does not outweigh the cost, and this can happen due to limited excess solar generation or insufficient monetary stimulus. Thus, the investment of ES should be determined through a long-term budget plan, e.g., based on the break-even cost and annual revenue [27]. Nevertheless, one possible solution to address this challenge is that these prosumers and consumers can sell/purchase their energy to/from a third-party stakeholder, i.e., an energy-sharing provider [28] or EP named in this work, at a better rate compared with the utility grid. The investor-owned battery ES can gain revenue by providing stacked ancillary services and arbitraging energy with the utility grid and P2P agents, which is a newly emerging independent entity to the power grid [29].

Based on the discussion above, there still exist several research gaps in the literature on how to improve the economic efficiency and reliability of a P2P sharing market.

- A majority of existing P2P sharing works evaluated the benefit of ES sharing and different auction models. However, it is important to consider the applicability of installing a large number of ES systems and introducing various market clearing rules in P2P markets. As a newly emerging stakeholder in smart grid, the investor-owned ES system, which possesses a publicly accessible ES, is more practical in future P2P sharing markets to attract those P2P agents without enough incentives of installing household-level ES or participating in market bidding.
- Some previous studies have worked on the different pricing mechanisms of P2P sharing. However, they did not consider the dynamics in the market, in which the agents were able to observe the market evolution. In a practical P2P market, the information of individual PV generation, load, and price-based response of each household is challenging to obtain in advance due to privacy concerns. The research question is how we learn this type of information through the dynamic interaction among market agents. Thus a model-free, adaptive, and concise decision-making model is desired in a P2P market.
- The uncertainty in P2P market operations has not been considered or well addressed in existing research. Since the ES capacity scheduling is a time-coupled problem, it is important to decide the scheduling by look-ahead methods rather than during the sharing process. The uncertainty in PV and load may lead to bad decisions, therefore an appropriate forecasting method is required to optimize the capacity scheduling.

In this paper, we propose a possible stakeholder, i.e., the EP, in future P2P sharing to explore a promising way for addressing the challenges with the ever-increasing distributed renewable capacities and negawatt power in residential communities. The main contributions of this paper are threefold:

- An investor-owned ES sharing-based energy trading framework is proposed in a community P2P network, which benefits the utility grid, EP, prosumers, and consumers.
- A customized dynamic pricing mechanism is developed to incentivize both prosumers and consumers to participate in the P2P sharing platform established by the EP, and Q-learning is adopted to obtain the optimal pricing policy.
- A rolling horizon decision-making process is developed to address the uncertainty in the market, which maximizes both the excess solar consumption and the EP's benefit.

The remainder of this paper is organized as follows. Section 2 describes the architecture of the community energy market. The problem formulation is discussed in Section 3. Section 4 shows a case study with 10 agents to evaluate the market performance. Section 5 concludes the paper and discusses the future work.

2. Market architecture

2.1. Community sharing market

Fig. 1 presents a typical architecture of a community sharing system. The community, consisting of $\mathcal{N}$ agents, are allowed to share energy through a centralized sharing platform executed by an EP. The grid has a continuous supply of energy for the community without interruptions. There is no limit on the feed-in energy from households with excess renewable supply under the Feed-in Tariff (FiT), which refers to the price rate at which the prosumers can sell their excess solar generation to the utility grid [30]. Traditional grid distribution lines are represented as solid lines, and dashed lines stand for the P2P community market connections. Within the framework, there are three major assumptions: (i) The EP is responsible for installing, managing, and maintaining the P2P sharing platform within the community, as well as balancing the supply and demand between the community and the utility grid. (ii) Each residential user is assumed to have a certain portion of the flexible load to respond to different price signals and optimize its energy consumption. (iii) Prosumers are assumed to first consume their own PV generation, and then if their netload (i.e., load minus PV generation in this paper) is negative, these agents have surplus supply to share, and are regarded as sellers ($\mathcal{N}_s$); otherwise, these agents are buyers ($\mathcal{N}_b$); and we have $\mathcal{N} = \mathcal{N}_s \cup \mathcal{N}_b$.

Compared with distributed ES systems, the centralized EP with a shared ES system available to every agent is more beneficial and practical. As a coupling point between the utility grid and P2P agents, the only information that the EP can obtain is the real-time aggregated supply and demand, while the behind-the-meter and actual load data are not available for the EP to fully protect the market participants' privacy.

In a wholesale market, the selling prices of the utility grid (π_s) are divided into two categories: on-peak hours with a higher price and off-peak hours with a lower price, denoted by π_h and π_l, respectively. The FiT (π_f) is also adopted to denote the rate at which prosumers sell their excess generation to the utility grid. In the P2P community market, the EP can design customized buying and selling prices for sellers and buyers to incentivize all P2P community market agents to actively participate in the energy sharing. The buying prices (λ_b) refer to the trade-in rate for sellers, while the selling prices (λ_s) refer to the

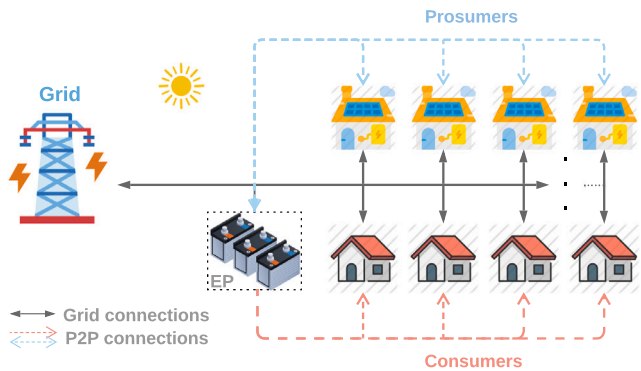

Fig. 1. A community P2P sharing market with EP, prosumers, and consumers. Tradition grid connections are represented as solid lines, and dashed lines stand for the P2P connections. The arrow directions indicate the electricity flow.

charge rate for buyers. The selling price needs to be no less than the buying price to ensure the EP's profit. The buying (λ_b) and selling (λ_s) prices are constrained by the utility price [3] (please note the time superscript h is omitted in this Section).

$$(\lambda_b, \lambda_s) \in [\pi_f, \pi_s] \tag{1}$$

2.2. EP model

Based on the discussion above, the profit function of EP is given by:

$$P = \begin{cases} \lambda_s E_b - \lambda_b E_s - \pi_s \Delta E - c \cdot |x|, & \Delta E \geq 0 \\ \lambda_s E_b - \lambda_b E_s - \pi_f \Delta E - c \cdot |x|, & \Delta E < 0 \end{cases} \tag{2}$$

where E_b and E_s denote the total demand from buyers $E_b = \sum_{i=1}^{\mathcal{N}_b}(l_i - pv_i)$ and supply from sellers $E_s = \sum_{i=1}^{\mathcal{N}_s}(pv_i - l_i)$ inside the community, respectively. The parameter ΔE denotes the imbalance between supply and demand, and $\Delta E = E_b - E_s + x$, which needs to be mitigated by the utility grid. A positive ΔE denotes that the retailer has to purchase power, and a negative value denotes feeding power back to the grid. The parameter x denotes the battery capacity scheduling; a positive value denotes charging and a negative value denotes discharging. The term $c \cdot |x|$ is the equivalent cost of acquiring x units of storage capacity.

There are different sharing models inside the community, e.g., direct sharing and buffered sharing [28]. The direct sharing model refers to the direct energy sharing between buyers and sellers inside the community via the P2P platform, and ES is inactive in this sharing model. The buffered sharing refers to the energy sharing with the assistance of the ES; the EP stores energy when solar generation is excessed or at off-peak hours, and discharge energy when PV is insufficient or at on-peak hours. These two sharing models work simultaneously in the market, and the scenarios can be divided into three main categories:

- EP arbitrages from the utility grid with ES. The utility time-of-use (ToU) price is set to mitigate the gap between load peak and valley, and the EP can charge its battery during off-peak periods and discharge during on-peak periods.
- EP arbitrages from the prosumers and consumers by direct sharing. The EP acts as a middleman to earn profits by buying excess solar from sellers and then directly selling to buyers.
- EP arbitrages from the prosumers and consumers by buffed sharing. The EP acts as a buffer to gain benefits by storing energy from sellers when PV generation is sufficient and then selling the stored energy to buyers.

2.3. Agents model

The prosumers and consumers aim to find the best scheduling of their load across a predefined optimization window (i.e., 24 h in this paper) to minimize the daily cost while achieving the maximum satisfaction level. Instead of using a conventional cost function to describe the agents' behaviors, in this paper the utility function from Ref. [22] is adopted, which describes the agent's consumption behaviors from two aspects: the utility from consuming energy and the cost of trading energy. Different from above-mentioned works in which there only exist consumers, in this paper the agent i's utility function is modified to reflect the fact that agents can act as either a seller [28] or a buyer in different time slots.

$$U_i = \begin{cases} k_i \ln(1 + l_i) - \lambda_s(l_i - pv_i), & l_i \geq pv_i \\ k_i \ln(1 + l_i) - \lambda_b(l_i - pv_i), & l_i < pv_i \end{cases} \tag{3}$$

In Eq. (3), the term $k_i \ln(1 + l_i)$ is the utility achieved by the agent i through consuming energy. The logarithm $\ln(\cdot)$ function has been widely used in economics for modeling the preference of users, since it is closely related to proportionally fair DR. And $(1 + x)$ is a typical modified form to avoid the undesired utility of $-\infty$. Note that k_i is a combination of the utility weight coefficient and consumption preference parameter. It is derived from Eq. (3) that a greater value of k indicates that the agent is willing to consume more energy to earn higher satisfaction. Besides, a higher selling price λ_s will decrease the willingness of a buyer ($l_i \geq pv_i$) to consume more energy, while a higher buying price λ_b will encourage a seller ($l_i < pv_i$) to sell more energy by reducing its consumption. In the dynamic pricing strategy, for any given price λ_b or λ_s, the agent i adapts its consumption to the best response as l_i^* for maximizing its utility U_i. Note that the optimal l_i^*, at which the consumer i achieves its maximum utility in response to a price set (λ_b, λ_s), can be found from Eq. (3),

$$l_i^* = \begin{cases} k_i/\lambda_s - 1, & l_i \geq pv_i \\ k_i/\lambda_b - 1, & l_i < pv_i \end{cases} \tag{4}$$

This optimal solution holds when the obtained value locates within the consumption constraints $[l_{min}, l_{max}]$ that are predefined based on individual agent's preference. Otherwise, the optimal solution l_i^* will always lie on the boundary due to its strict concavity.

3. Problem formulation

3.1. Capacity scheduling and decision-making

The primary goal of EP is to maximize its benefit. According to the constraints defined in Eq. (1), the prices set by the EP always dominate the utility prices. To maximize the renewable energy consumption in a P2P market, the cost can also be evaluated based on the amount of electricity that is not purchased from the utility grid. In an ideal scenario all excess energy is shared among the participants rather than feeding back into the grid. Thus, minimizing the community social cost can be transformed to the minimization of trading with the utility grid, since the energy sharing within the community, i.e., from sellers to EP then to buyers, does not impact the aggregated netload. And the cost function of the community can be transformed as:

$$C^{h\sim H} = \sum_{t=h}^{H} \left[\pi_s^t \cdot \max\left(NL^t + x^t, 0\right) + \pi_f^t \cdot \min\left(NL^t + x^t, 0\right) + c \cdot |x^t| \right] \tag{5}$$

$$-\Lambda/C_{rate} \leq x^t \leq \Lambda/C_{rate} \tag{6}$$

$$SoC_{min} \leq SoC^t \leq SoC_{max} \tag{7}$$

$$SoC^t = \begin{cases} SoC^{t-1} + x^t \cdot \eta, & x_i^t > 0 \\ SoC^{t-1} + x^t/\eta, & x_i^t < 0 \end{cases} \tag{8}$$

where H is the whole optimization horizon (i.e., 24 h), and h is the current time slot. $C^{t\sim H}$ is the look-ahead electricity cost of EP from the current time h to future H. The parameter NL^t denotes the aggregated netload of the community, and $NL^t = \sum_{i=1}^{\mathcal{N}}(l_i - pv_i)$. The parameter x^t represents the battery charging/discharging schedule, and η is the (dis-)charging efficiency. The parameter Λ is an integer number, which

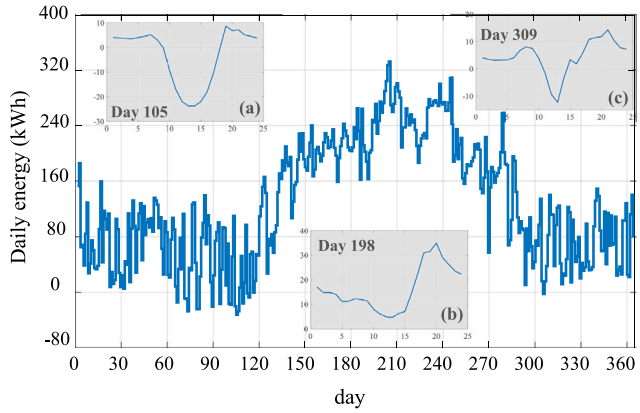

Fig. 2. The community daily netenergy (aggregated amount of ten households over 24 h) in 2018, and hourly netload curves of three typical days in different seasons: (a) April 14, (b) July 17, and (c) November 6.

denotes the nominal capacity of the ES. The terms of $-\Lambda/C_{rate}$ and Λ/C_{rate} are the lower and upper bounds of the (dis-)charging energy in each time slot, respectively, and C_{rate} is the maximum (dis-)charge rate of the ES. The parameter SoC^t is the SoC of the ES at the end of time slot t; SoC_{min} and SoC_{max} are the lower and upper limits of the ES, respectively. In this paper, we consider a maximum C_{rate} of 2, $SoC \in [0.05, 0.95]$ with an initial minimal value 0.05, and $\eta = 0.95$. A nominal capacity optimization [27] of the EP is conducted by minimizing the electricity cost for the whole community, which is expressed as:

$$\Lambda^* = \arg\min \quad C(NL, \Lambda) \tag{9}$$

The determination of the ES capacity is complicated, since the aggregated netload varies all the time. To perform long-term planing of EP, the annual operation should be considered to determine the optimal capacity, which will be further explained in Section 4.

3.2. Netload forecasting

A forecasting-assisted pricing model is developed to help the EP predict the aggregated netload in the market and thus design better internal energy sharing prices. Long short-term memory (LSTM) has shown to be effective in netload forecasting, which is adopted in this study. LSTM is an artificial recurrent neural network architecture with feedback connections, which is capable of processing single data points as well as entire data sequences. A common LSTM unit is composed of a memory cell **c**, an input gate **i**, an output gate **o**, and a forget gate **f**. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information in and out of the cell. As a sequence-based model, LSTM is able to abstract the data pattern, maintain the memory of states, and establish the temporal correlations between information and current circumstances, which means the decision made at time step $t-1$ will affect the decision in the following time slot t. Such characteristic is ideal for load forecasting since the community supply and demand have been proven to follow certain routines due to strong spatial and temporal correlations [6]. To train the LSTM net, the input features including (i) historical netload (ii) calendar information (hour, day, week, and holidays), are used in this study. The data processing and parameters selection are executed following [31], and the simulation is run in the Matlab2018b Deep Learning Toolbox.

3.3. Rolling horizon decision-making

To manage the netload forecasting error and improve the market efficiency, a rolling horizon method [32] is applied to solve a multi-timescale capacity scheduling problem. The rolling horizon approach considers a forecasting period, in which the future netload is treated as a deterministic forecasting value, and the optimal decision-making is obtained.

(i) At the beginning of a given step h, set the forecasting horizon $h \sim H$ and obtain the forecasted netload $\hat{NL}^{h\sim H}$ using LSTM.
(ii) Establish and solve the optimization problem in Eq. (5), and obtain the ES capacity scheduling $x^{h\sim H}$.
(iii) During step h, apply x^h obtained from (ii) into Eq. (2), get the optimal pricing strategy $\lambda_b^{h*}\&\lambda_s^{h*}$ (introduced in the next subsection), and observe the updated NL^h.
(iv) Enter the next step $h+1$, update the future netload forecast $\hat{NL}^{h+1\sim H}$ based on the observation obtained from the previous step, and repeat (i)–(iii) until reaching the end period H.

The proposed rolling horizon approach allows EP to modify forecasting results based on the updated/recent actual netload. Although the solution obtained from the rolling horizon approach might be sub-optimal in practice due to the lack of accurate future information, an appropriate length of look-ahead horizon and a future discount parameter could be leveraged to address this issue.

3.4. Q-learning based dynamic price design

Reinforcement learning (RL) has been applied to a variety of grid problems in recent years, such as demand response [23], battery scheduling [29], forecasting [33], cyber-security [34], etc. As a model-free RL technique, Q-learning is able to learn the optimal policy of finding the best action at every time step, even when some accurate future data is not available during the training process. In the Q-learning algorithm, the agent interacts with the environment through executing sequential actions at a series of states based on the interaction between states and actions, until reaching an ultimate goal. The actions are evaluated by a reward feedback from the environment, which is used to update the Q-value. During the learning process, the Q-values are stored and converge to a maximum value after updating over a sufficient number of iterations. When the optimal Q-value (Q^*) is obtained, the optimal policy can be defined as $\upsilon^* = \text{argmax}\ Q^*$.

In this study, the dynamic pricing problem is modeled as a discrete finite horizon Markov decision process. The EP acts as the learning agent and interacts with the P2P agents (environment) through a set of actions (λ_b, λ_s). The immediate reward function in a specified time t is defined as Eq. (2), while the total reward for the future is defined as the additive inverse of Eq. (5). With state, action, and reward defined, the price design problem is realized by training the Q-learning agent to take action based on the optimal policy. Algorithm 1 presents the Q-learning pseudo-code that is employed to obtain the maximum Q-value and the optimal dynamic prices.

Algorithm 1: Q-Learning Based Dynamic Sharing Price Design

Data: Number of steps H, learning rate θ, discount factor γ, number of iterations N_e.
Result: Obtain the maximum Q-value.
Initialize Q-value;
while $i \le N_e$ **do**
  while $h \le H$ **do**
    Select and execute an action a_h at state s_h by ϵ-greedy policy [23];
    Observe reward $r_h(s_h, a_h)$ by solving Eq. (2) and new state s_{h+1} by solving Eq. (5);
    Update Q-value:
    $Q_h(s_h, a_h) \leftarrow (1-\theta)Q_e(s_h, a_h) + \theta[r_h(s_h, a_h) + \gamma \max \cdot Q_h(s_{h+1}, a)]$;
  end
end

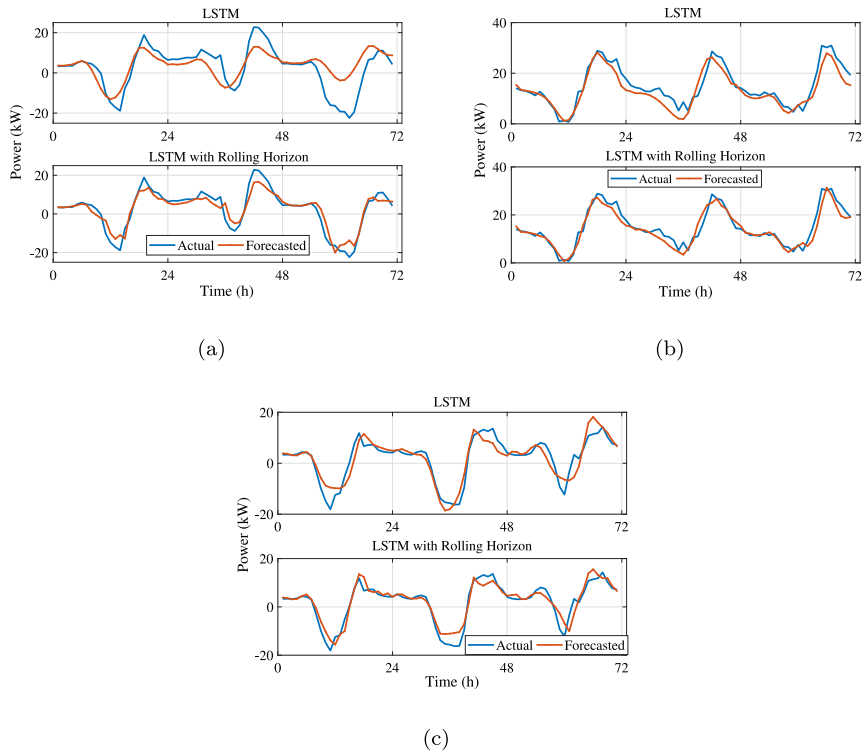

Fig. 3. Aggregated netload forecasts of (a) April 14, (b) July 17, and (c) November 6. Sub-figures on the top show the forecasting results of traditional LSTM, and sub-figures on the bottom show the results of LSTM with rolling horizon update (LSTM-RH).

4. Case study

The developed EP-based energy sharing market is evaluated with a case study containing 10 participants (7 prosumers with PV panels and 3 pure consumers) in Austin, Texas. The detailed household ID, netload, and solar generation can be found in Dataport.[1] The missing or abnormal data has been filled or interpolated based on neighboring observations.

The equivalent cost of ES is defined as the threshold at which using a battery as an energy buffer is profitable. In this work we adopt the prices from Ref. [12], i.e., using a two-period ToU tariff with an on-peak rate of \$0.55/kWh from 7:00 to 23:00, and an off-peak rate of \$0.20/kWh in other hours. The amortized equivalent cost of ES sharing, including the P2P sharing loss and battery cost, is considered as \$0.15/kWh. This equivalent cost is expected to be continuously decreasing with the advancement of energy storage technologies. It is important to note that all parameters or prices employed in this study could be updated to accommodate to other markets, and this does not distort the analysis of the simulation.

4.1. Battery capacity determination

The daily netenergy (aggregated energy over 24 h) curves of the community are shown in Fig. 2. Although there are some variations in netenergy in a same season, some general patterns could be extracted from Fig. 2. For example, in some spring days, the aggregated netenergy drops to *negative* values, which means the PV generation exceeds the demand since there is less consumption for air conditioners. In most summer days, the netenergy is always *positive* due to the high demand in cooling. While in some other days, the aggregated netenergy is relatively low but above zero, which are referred as *off-peak* days. The winter season is not considered here, since 39% of homes in Texas use natural gas as their primary heating source.[2] Thus, three representative days from each category are selected for analysis, i.e., (a) April 14 as a typical *negative* day in spring, (b) July 17 as a typical *positive* day in summer, and (c) November 6 as a typical *off-peak* day in fall.

Table 1
Optimal battery capacity Λ^* (kWh) to minimize the community's daily trading cost with the utility grid under three different day.

	April 14	July 17	November 6
Λ^* (kWh)	37	314	218
Daily cost (\$)	1.1	127.8	78.9

Table 1 shows the optimal ES capacity and corresponding daily cost of the community trading with the grid of the selected days. In typical spring days (e.g., April 14), the ES mainly performs arbitraging from PV prosumers. However, there are fewer buyers in this category due to the low load demand and high PV generation (the average netload is negative as seen in Fig. 2), thus yielding a minimal ES capacity of 37 kWh. In typical summer days (e.g., July 17), the ES mainly performs arbitraging from the utility grid, since there is little excess energy for sharing due to the high load demand. The arbitraging profit from the grid is \$0.35/kWh, which outweighs the battery equivalent cost, thus a large ES capacity of 314 kWh is needed to fully cover the load. While in typical fall days (e.g., November 6), the EP works in both negative and positive netload patterns, and the optimal ES capacity is 218 kWh for minimizing the daily cost. Due to the high cost of energy storage, we mainly focus on battery capacity scheduling and pricing in this research rather than maximizing the capacity. Based on the results from the three representative days, we choose 40 kWh as the nominal capacity of the ES in this study, being able to support the maximum hourly peak load demand (i.e., 43.82 kW, occurs on July 17) for about 1 h. This selection would be helpful to deal with contingencies and also avoid idle capacity.

[1] https://www.pecanstreet.org/dataport/.
[2] https://www.eia.gov/todayinenergy/detail.php?id=47116.

Table 2
Forecasting accuracy summary of two LSTM models in terms of nRMSE (%).

	April 14	July 17	November 6
LSTM	16.46	10.24	10.03
LSTM-RH	7.98	8.15	8.51

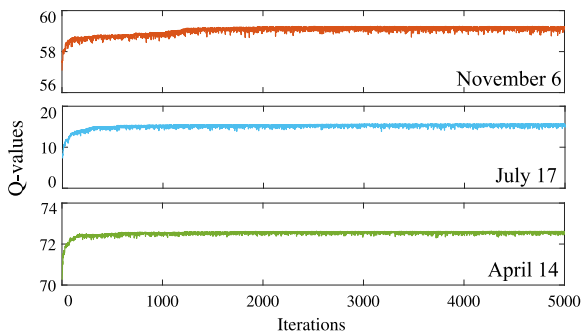

Fig. 4. The convergence of Q-values in three typical days. A higher Q-value indicates that the EP is expected to earn more profit.

4.2. Netload forecasting and capacity scheduling

To train the LSTM net, each category is split into training and testing sets using the latest 30 days netload data prior to the target days: the first 90% for training and the rest 10% for validating. Then the netload forecasts of the three representative days are obtained using the three trained LSTM networks. Fig. 3 compares the result of two different forecasting strategies: benchmark LSTM and LSTM with rolling horizon updating (LSTM-RH, also known as online LSTM), and Table 2 compares the normalized root mean square error (nRMSE) based on the maximum netload. It is observed that the LSTM-RH has a better accuracy in all three days, indicating the effectiveness of rolling horizon update. Overall, the netload forecasting accuracy is satisfactory, since the algorithm only uses the historical netload of the community.

4.3. Optimal community sharing prices and demand response

The simulations were executed on a laptop with a dual Core i7-6600U CPU running at 2.8 GHz and with 16.0 GB RAM. The convergence test of the Q-learning is performed and the learning curves are shown in Fig. 4. The EP learns extremely fast from interacting with the environment at the beginning by trial and error. After about 1200/600/600 iterations, the EP learns slowly and tends to converge, and the computation time of 5000 iterations is 351.3 s, 370.2 s, and 373.5 s on April 14, July 17, and November 6, respectively. After the learning agent converges to the maximum Q-value, it is seen from Fig. 4 that on April 14, the EP obtains the highest Q-value, indicating the EP will gain the highest profit in this typical spring day. While on July 17, the EP gets the lowest Q-value, indicating a lower profit compared with the other two days. The detailed profits of different days are shown in Table 3.

The optimal pricing policy is obtained by Algorithm 1, and the results are shown in Fig. 5. The energy sharing inside the community occurs mainly from 9 a.m. to 5 p.m. (with one exception of 8 a.m. on April 14). The EP designs different buying and selling prices to encourage P2P market agents to participate in the energy sharing, and the community sharing prices are always better than the utility price $\pi_f = \$0.1, \pi_s = \$0.20/\$0.55$ to ensure the economic incentives. The dynamic pricing shows different patterns in those three representative days. For example, for (a) April 14, the community buying price is the lowest among the three cases due to excess solar generation; while for (b) July 17, the community buying price is the highest due to the shortage of solar generation, which also corresponds with the positive netload in Fig. 2. On (c) November 6, the dynamic pricing shows a more volatile pattern, since the aggregated netload is more flexible with the usage of ES, thus the EP can adaptively determine the prices and capacity scheduling. It is also seen that in some hours the EP's buying and selling price are the same in (b) July 17 and (c) November 6, which means the EP encourages prosumers to sell more energy with the aim of balancing the supply and demand even without obtaining an immediate reward from arbitrage.

Once the optimal pricing policy is obtained, the buyers' and sellers' optimal consumption is determined by Eq. (3), and the results are shown in Fig. 6. Under the sharing framework, the negative or valley netload increases while the positive netload decreases in all three cases, which means the excess solar generation is stored for sharing among the community during peak-hour periods and less power is sent back to the utility grid. The capacity scheduling in the three days can also be extracted from Fig. 6. On (a) April 14, the ES is almost inactive in the morning, indicating the EP has no incentive to arbitrage from the utility grid during morning off-peak hours. Then the netload starts to rise after 9 a.m. and decreases after 6 p.m. with the ES, indicating the EP prefers to store excess solar generation in the market and then sell the buffered to market participants in the evening. While on (b) July 17 and the (c) November 6, the EP starts to store energy in the morning off-peak hours by arbitraging from the utility off-peak hours, then discharges the stored energy during morning on-peak hours. However, there is a slight difference between these two cases. On (b) July 17, there is a solar generation shortage in the market due to the high demand in the summer, thus the ES works in a straightforward routine, i.e., simply arbitraging from the utility grid. On (c) November 6, the ES will release some stored energy when on-peak hours start (i.e., 8 a.m.–10 a.m.), then charge again with the excess solar generation in the noon; and after the noon, the ES shows a similar discharging pattern with the other two cases. It is also seen that the negawatt in Fig. 6(c) can be fully offset by the ES.

4.4. Market performance

Table 3 summarizes the EP's profits and 10 agents' utilities on different days (the gray color highlights the pure consumers' indexes, and others are PV prosumers). It is seen that all agents have gained higher utilities with ES sharing and community internal prices, compared with the baseline. On April 14, consumers 5, 8, and 10 have the highest utility growth rates, indicating the negative day benefits buyers more. Similarly, on July 17, prosumers, especially 1, 2, 3 are expected to earn the highest utilities, due to the shortage in PV supply in the community market. While on November 6, almost all prosumers and consumers have shown considerable utility increases. By managing the community-based P2P sharing with ES, the EP has also earned significant daily profit, even with a small community with only 10 agents. By increasing the number of agents and the investment for a larger size battery, the profit of EP is expected to further grow. Overall the investor-owned ES could effectively and significantly enhance renewable energy consumption within the community. The EP earns profits by managing the energy sharing and ES arbitrage, the P2P agents receive higher consumption utilities, and there is less negawatt power fed back into the grid, which leads to a win-win-win situation.

4.5. Scalability and applicability analysis

To analyze the scalability and applicability of the proposed method, additional case studies are conducted by increasing the number of agents. Extra agents' netload data are generated based on the existing dataset by multiplying a random factor from 0.8 to 1.2 due to limited data availability. The computation time of the two different stages, i.e., capacity scheduling and dynamic pricing, is summarized in Table 4. It is seen that both the two stages have a similar computation cost with the increasing number of agents. In the capacity scheduling stage

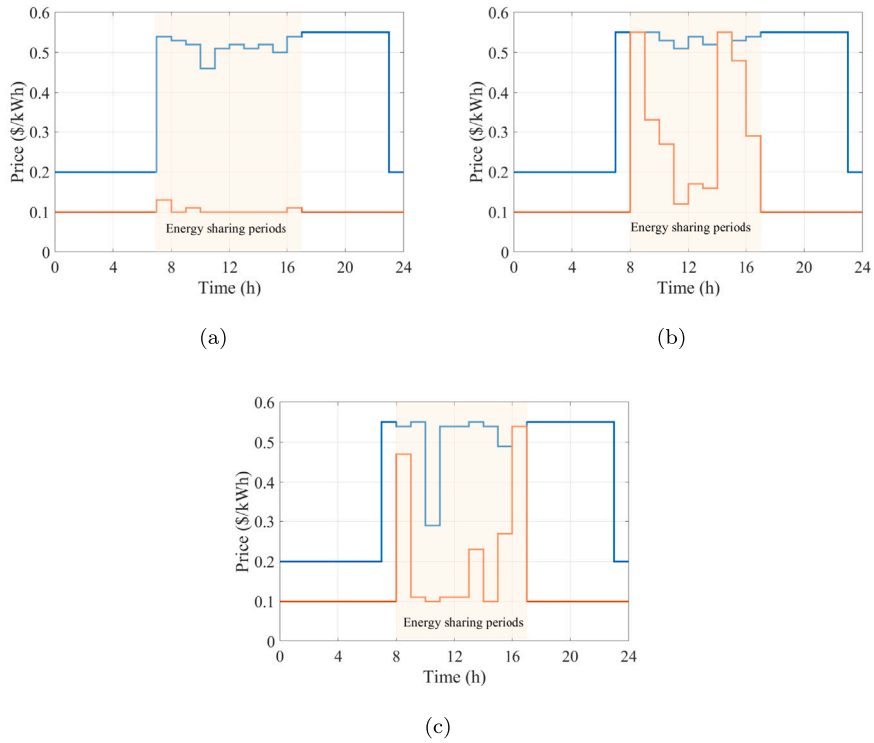

Fig. 5. The internal sharing prices of (a) April 14, (b) July 17, and (c) November 6. Blue lines denote the EP's selling price λ_s, and red lines denote the buying price λ_b. The internal sharing prices are dynamically determined by the EP, which benefits sellers more when the market is in shortage and benefits buyers more when the PV generation is surplus.

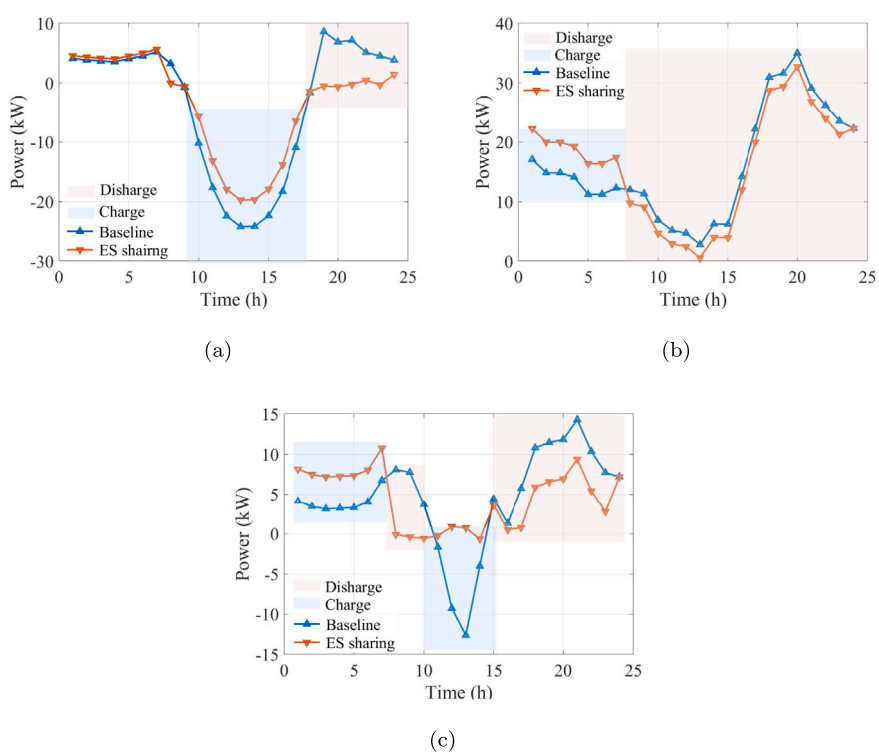

Fig. 6. The aggregated netload response with ES on (a) April 14, (b) July 17, and (c) November 6, illustrating the ES capacity scheduling and how the negawatt is managed through ES in three typical days.

Table 3
Profits of EP and utilities of P2P participants ($). Agents 5, 8, and 10 are pure consumers and others are PV prosumers. This table shows that consumers earn more benefits in a typical spring day and prosumers earn more benefits in a typical summer day. The EP is able to earn significant profit through hybrid ES scheduling and dynamic pricing strategies under different seasons.

	April 14			July 17			November 6		
	Baseline	*Sharing*	Growth	*Baseline*	*Sharing*	Growth	*Baseline*	*Sharing*	Growth
EP	–	19.6761	–	–	17.6349	–	–	18.3344	–
1	4.5513	4.5713	0.44%	9.7669	13.2043	35.20%	3.9008	4.5585	16.86%
2	2.8999	2.9395	1.37%	9.6230	11.1398	15.76%	1.3785	1.9127	38.75%
3	6.5557	6.5825	0.56%	21.0321	22.4565	6.78%	2.2939	2.4133	5.21%
4	3.6979	3.7411	1.17%	87.5938	88.0672	0.54%	3.7182	4.3128	16.00%
5	1.4760	1.6911	14.57%	18.3232	18.6959	2.03%	1.0286	1.1845	15.16%
6	4.8469	4.8904	0.90%	23.7939	24.7626	4.07%	3.3661	3.8033	12.99%
7	4.2917	4.3278	0.84%	23.9202	24.9050	4.12%	6.8507	7.1225	3.97%
8	3.4348	3.7613	9.51%	30.0239	30.4558	1.44%	3.1178	3.3115	6.21%
9	6.8792	6.9327	0.78%	74.7741	74.9448	0.23%	45.5476	46.0663	1.14%
10	1.1245	1.2597	12.02%	5.5181	5.6804	2.94%	1.2287	1.4222	15.75%

Table 4
Computation time with different numbers of agents to show the scalability of the proposed method.

Numbers of agents	10	25	50	100
Computation time for capacity scheduling (s)	0.011	0.036	0.022	0.027
Computation time for 5000 iterations (s)	373.5	365.5	375.5	368.4

(as shown in Eq. (5)), the EP only uses the aggregated netload for calculation, thus the increase of the agents number will not increase the computation complexity. For prosumers and consumers agents, the utility maximization problem (as shown in Eq. (3)) can be fast solved once the internal prices are settled, and all agents run this calculation independently and in parallel. For the Q-learning dynamic pricing process, since the internal prices are identical for buyers and sellers, there is no interaction or competition between buyers and sellers, they only act based on their best responses to different prices. The EP only needs to update the optimal pricing policy based on the aggregated response of all buyers and sellers. Thus, the computational burden in this approach is relatively low, making it scalable and applicable to a larger scale market that may consist of hundreds of agents.

The LSTM method adopted in this study yields a satisfactory accuracy. Note that the forecasting accuracy could be further improved with the availability of more data, such as behind-the-meter PV generation, detailed household load, etc. However, detailed household load and behind-the-meter PV data are challenging to gather due to privacy concerns, since it requires the cooperation from the P2P market agents for data collection. The proposed EP with aggregated netload forecasting can achieve a win-win-win situation without utilizing those detailed data. One limitation of this research is that the cost of ES is still high at the current stage, however, this could be addressed with the advancement of energy storage technologies.

5. Conclusion

This paper proposed an energy pawn (EP) based energy sharing framework in a community market that consists of an investor-owned energy storage system, prosumers and consumers. A rolling-horizon decision-making strategy was developed to maximize the EP's revenue, by solving a forecasting-based capacity scheduling problem and a Q-learning-based dynamic pricing problem. The EP can adaptively determine the community buying and selling prices according to the market participants' response. The proposed sharing framework was applied to a community market with 10 agents, and three representative days with different netload patterns (i.e., typical spring, summer, and fall days) are analyzed to evaluate the effectiveness of storage sharing in peer-to-peer (P2P) energy trading. Results showed that the proposed EP could promote community energy sharing and increase both EP's revenue and market participants' utilities. Overall, the EP can promote renewable energy consumption, reduce the negawatt fed back into the grid, and balance the energy supply and demand in the community, which can be regarded as a win-win-win situation for both the grid, EP, and P2P participants.

Potential future work will further explore (i) a more accurate demand response model when detailed users' data are available; (ii) a hierarchical P2P sharing framework involving different stakeholders, such as retailers, generators, etc.; (iii) methodologies for more accurate netload forecasting to further enhance the market efficiency.

CRediT authorship contribution statement

Li He: Formal analysis, Investigation, Methodology, Validation, Roles/Writing - original draft. **Yuanzhi Liu:** Formal analysis, Investigation, Validation, Roles/Writing - original draft. **Jie Zhang:** Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Tushar W, Saha TK, Yuen C, Smith D, Ashworth P, Poor HV, et al. Challenges and prospects for negawatt trading in light of recent technological developments. Nat Energy 2020;5(11):834–41.

[2] Tushar W, Yuen C, Saha TK, Morstyn T, Chapman AC, Alam MJE, et al. Peer-to-peer energy systems for connected communities: A review of recent advances and emerging challenges. Appl Energy 2021;282:116131.

[3] Liu N, Yu X, Wang C, Li C, Ma L, Lei J. Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers. IEEE Trans Power Syst 2017;32(5):3569–83. http://dx.doi.org/10.1109/TPWRS.2017.2649558.

[4] Wang Y, Saad W, Han Z, Poor HV, Başar T. A game-theoretic approach to energy trading in the smart grid. IEEE Trans Smart Grid 2014;5(3):1439–50.

[5] Chen Y, Mei S, Zhou F, Low SH, Wei W, Liu F. An energy sharing game with generalized demand bidding: Model and properties. IEEE Trans Smart Grid 2020;11(3):2055–66.

[6] He L, Zhang J. A community sharing market with PV and energy storage: An adaptive bidding-based double-side auction mechanism. IEEE Trans Smart Grid 2021;12(3):2450–61. http://dx.doi.org/10.1109/TSG.2020.3042190.

[7] Wang J, Zhong H, Wu C, Du E, Xia Q, Kang C. Incentivizing distributed energy resource aggregation in energy and capacity markets: An energy sharing scheme and mechanism design. Appl Energy 2019;252:113471.

[8] Saad W, Han Z, Poor HV, Başar T. A noncooperative game for double auction-based energy trading between PHEVs and distribution grids. In: Smart grid communications (SmartGridComm), 2011 IEEE international conference on. IEEE; 2011, p. 267–72.

[9] Tushar W, Saha TK, Yuen C, Morstyn T, McCulloch MD, Poor HV, et al. A motivational game-theoretic approach for peer-to-peer energy trading in the smart grid. Appl Energy 2019;243:10–20.

[10] He L, Zhang J. Distributed solar energy sharing within connected communities: A coalition game approach. In: 2019 IEEE power energy society general meeting. 2019, p. 1–5. http://dx.doi.org/10.1109/PESGM40551.2019.8973867.

[11] Tushar W, Saha TK, Yuen C, Azim MI, Morstyn T, Poor HV, et al. A coalition formation game framework for peer-to-peer energy trading. Appl Energy 2020;261:114436.
[12] Chakraborty P, Baeyens E, Poolla K, Khargonekar PP, Varaiya P. Sharing storage in a smart grid: A coalitional game approach. IEEE Trans Smart Grid 2019;10(4):4379–90.
[13] Lee J, Guo J, Choi JK, Zukerman M. Distributed energy trading in microgrids: A game-theoretic model and its equilibrium analysis. IEEE Trans Ind Electron 2015;62(6):3524–33.
[14] Liu N, He L, Yu X, Ma L. Multiparty energy management for grid-connected microgrids with heat- and electricity-coupled demand response. IEEE Trans Ind Inf 2018;14(5):1887–97.
[15] Fleischhacker A, Auer H, Lettner G, Botterud A. Sharing solar PV and energy storage in apartment buildings: Resource allocation and pricing. IEEE Trans Smart Grid 2019;10(4):3963–73.
[16] Tushar W, Chai B, Yuen C, Huang S, Smith DB, Poor HV, et al. Energy storage sharing in smart grid: A modified auction-based approach. IEEE Trans Smart Grid 2016;7(3):1462–75.
[17] Paudel A, Chaudhari K, Long C, Gooi HB. Peer-to-peer energy trading in a prosumer-based community microgrid: A game-theoretic model. IEEE Trans Ind Electron 2018;66(8):6087–97.
[18] Khorasany M, Mishra Y, Ledwich G. Market framework for local energy trading: a review of potential designs and market clearing approaches. IET Gener Transm Distrib 2018;12(22):5899–908.
[19] Zhong W, Xie K, Liu Y, Yang C, Xie S. Auction mechanisms for energy trading in multi-energy systems. IEEE Trans Ind Inf 2018;14(4):1511–21.
[20] Shamsi P, Xie H, Longe A, Joo J-Y. Economic dispatch for an agent-based community microgrid. IEEE Trans Smart Grid 2015;7(5):2317–24.
[21] Cintuglu MH, Martin H, Mohammed OA. Real-time implementation of multiagent-based game theory reverse auction model for microgrid market operation. IEEE Trans Smart Grid 2015;6(2):1064–72.
[22] Ma L, Liu N, Zhang J, Tushar W, Yuen C. Energy management for joint operation of CHP and PV prosumers inside a grid-connected microgrid: A game theoretic approach. IEEE Trans Ind Inf 2016;12(5):1930–42. http://dx.doi.org/10.1109/TII.2016.2578184.
[23] Lu R, Hong SH, Zhang X. A dynamic pricing demand response algorithm for smart grid: reinforcement learning approach. Appl Energy 2018;220:220–30.
[24] Long C, Wu J, Zhou Y, Jenkins N. Peer-to-peer energy sharing through a two-stage aggregated battery control in a community microgrid. Appl Energy 2018;226:261–76.
[25] Kalathil D, Wu C, Poolla K, Varaiya P. The sharing economy for the electricity storage. IEEE Trans Smart Grid 2019;10(1):556–67.
[26] Barbour E, Parra D, Awwad Z, González MC. Community energy storage: A smart choice for the smart grid? Appl Energy 2018;212:489–97.
[27] Sun M, Chang C-L, Zhang J, Mehmani A, Culligan P. Break-even analysis of battery energy storage in buildings considering time-of-use rates. In: 2018 IEEE green technologies conference. IEEE; 2018, p. 95–9.
[28] Liu N, Cheng M, Yu X, Zhong J, Lei J. Energy sharing provider for PV prosumer clusters: A hybrid approach using stochastic programming and stackelberg game. IEEE Trans Ind Electron 2018;65(8):6740–50.
[29] Huang B, Wang J. Deep-reinforcement-learning-based capacity scheduling for PV-battery storage system. IEEE Trans Smart Grid 2021;12(3):2272–83.
[30] Couture TD, Cory K, Kreycik C, Williams E. Policymaker's guide to feed-in tariff policy design. Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States); 2010.
[31] Kong W, Dong ZY, Jia Y, Hill DJ, Xu Y, Zhang Y. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans Smart Grid 2017;10(1):841–51.
[32] Silvente J, Kopanos GM, Pistikopoulos EN, Espuña A. A rolling horizon optimization framework for the simultaneous energy supply and demand planning in microgrids. Appl Energy 2015;155:485–501.
[33] Feng C, Sun M, Zhang J. Reinforced deterministic and probabilistic load forecasting via Q-learning dynamic model selection. IEEE Trans Smart Grid 2019;11(2):1377–86.
[34] Kurt MN, Ogundijo O, Li C, Wang X. Online cyber-attack detection in smart grid: A reinforcement learning approach. IEEE Trans Smart Grid 2018;10(5):5174–85.