Data-Driven Stochastic Game With Social Attributes for Peer-to-Peer Energy Sharing

Liudong Chen, *Graduate Student Member, IEEE*, Nian Liu, *Member, IEEE*, Liangying Liu, *Graduate Student Member, IEEE*, Xinghuo Yu, *Fellow, IEEE*, and Yusheng Xue, *Life Member, IEEE*

***Abstract*—Distributed energy resources create a prosumers era. Peer-to-peer (P2P) energy sharing is an effective way to conduct prosumers-based energy management. Integrating prosumers' social attributes in energy engineering substantially affects their decisions, which brings challenges to the energy scheduling under the cyber-physical-social environment. In this paper, a data-driven stochastic game model with prosumers' social attributes is proposed for P2P energy sharing. According to social psychology, the prosumers' social attributes are expressed as subjective probabilities, which are studied by the spatial-temporal graph convolutional networks. In the network, a double-layer feature graph is built to learn the social attributes based on the social survey data and load metering data. The P2P energy sharing incurred randomness comes from social attributes of interactive prosumers, which is formulated as a stochastic game model. In this game, a subjective utility model is proposed for prosumers, and the energy scheduling is conducted with the designed dynamic interval adjustment method. Numerical analysis reveals the results of the learning network and generalized Nash equilibrium. Through comparing with rational scenarios, the influence of prosumers' social attributes on P2P energy sharing is concluded.**

***Index Terms*—Cyber-physical-social system, spatial-temporal graph convolutional network, social attributes, P2P energy sharing, stochastic games.**

NOMENCLATURE

Parameter and Variable in Energy Sharing

N, i	Number and index of prosumers
H, h	Number and index of time slot
K, k	Number and index of load interval
$iter^h$	Number of iterations during game process
N_s^h, N_b^h	Number of sellers and buyers
E_i^h	PV production
$L_{i,k}^h$	Load consumption, which includes initial reference load $L0_i^h$ and optimal load L_i^{h*}
$fL_{i,k}^h$	Flexible load
$p_{i,k}^h$	Subjective probability
$U_{i,k}^h$	Subjective utility
δ_i^h	Constant preference parameter of the prosumers
$fl_{i,min}^h$, $fl_{i,max}^h$	Lower and upper bounds of flexible load
$TL_{i,k}$	Flexible load in the entire time slots
C^h	Feeder capacity
$pri_{s.k}^h$, $pri_{b.k}^h$	Dynamic selling and buying prices
TSL_k^h, TBL_k^h	Total selling and buying energy of the system
TNL_k^h	Total netload of the system
α	Profit parameter of the prosumers
λ_s, λ_b	Utility grid selling and buying prices
SDR_k^h	Supply and demand ratio
Δl_i^h, ΔL_i^H	Load different between initial and optimal load, and cumulative load difference in previous H time slots
$\Delta l_i^{h'}$, $\Delta l_i^{h''}$	Load different between initial and optimal load after the first and second adjustment
$IL_{i,k}^h$	Length of load interval
m	Number of adjusted intervals
b, d	Coefficient of interval adjustment constraints, which includes b_1, b_2, b_3, d_1, d_2, and d_3.

Parameter and Variable in STGCN

j, r	Index of feature combination and STGCN layer
sN	Number of nodes in undirected complete graph, which is also the number of social attributes' types and feature combinations
Hs,Hs_{r-1}, hs	Length of time slice respective in dataset and layer r, and the index of the time slice
sM	Length of longest feature combination
$\boldsymbol{\chi}$	Dynamic node feature matrix
$\boldsymbol{\chi}_{\mathrm{S}}$, $\boldsymbol{X}_{\mathrm{S},hs}^{j}$, $x_{\mathrm{S},hs}^{c,j}$	Social attributes matrix, feature combination j's element in the matrix, and social survey question c in the feature combination j

Manuscript received December 5, 2020; revised April 10, 2021; accepted June 21, 2021. Date of publication June 30, 2021; date of current version October 21, 2021. This work was supported by the National Natural Science Foundation of China under Grant 51877076, and in part by Australian Research Council Discovery Program under Grant DP200101197. Paper no. TSG-01807-2020. *(Corresponding author: Nian Liu.)*

Liudong Chen, Nian Liu, and Liangying Liu are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China (e-mail: liudong@ncepu.edu.cn; nianliu@ncepu.edu.cn; 120202201434@ncepu.edu.cn).

Xinghuo Yu is with the School of Engineering, RMIT University, Melbourne, VIC 3001, Australia (e-mail: x.yu@rmit.edu.au).

Yusheng Xue is with State Grid Electric Power Research Institute, Nanjing 210003, China (e-mail: xueyusheng@sgepri.sgcc.com.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2020.3093587.

Digital Object Identifier 10.1109/TSG.2021.3093587

$\boldsymbol{\chi}_{\mathrm{Ls}}$, $\boldsymbol{X}^{k}_{\mathrm{Ls},hs}$ Load distribution matrix, and interval k's element in the matrix

$\boldsymbol{S}_{\mathrm{amls}}$, $\boldsymbol{S}_{\mathrm{ams}}$ Dynamic interactive influence between load intervals, and between social attributes

$\mathfrak{R}$, $\mathfrak{R}'$, $\mathfrak{R}'^{j_1,j_2}$ Dynamic spatial attention matrix, normalization type, and element in normalization matrix

$\mathfrak{I}$, $\mathfrak{I}'$, $\mathfrak{I}'^{hs_1,hs_2}$ Dynamic temporal attention matrix, normalization type, and element in normalization matrix.

I. INTRODUCTION

WITH the increasing number of distributed energy resources (DERs), such as photovoltaic (PV), the distribution system has entered the era of prosumers who has the capability of both power consumption and production and can serve as sellers or buyers to participate in the system. Although the prosumers can help promote the consumption of the DERs, they also complicate the control, power flow, and information flow of the system due to distributed allocation and fluctuation of PV production [1]. Peer-to-peer (P2P) energy sharing between prosumers is a feasible way to conduct prosumers-based energy management [2].

P2P energy sharing is an emerging framework that offers different sources of value to prosumers by allowing energy balance to a great extent and reducing energy supply from the utility grid [3]. During the P2P energy sharing, each prosumer is the self-organized entity, aimed at achieving their goal of maximum profit by setting their energy strategies. The researches can be grouped into two parts: (1) P2P energy sharing in the deterministic scenario, (2) P2P energy sharing with uncertainties. In the deterministic scenario, because the essence of P2P energy sharing is a multi-agent profit-related problem, game theory has been widely used with advantages in dealing with agents' competitive interests. The process can be generally formulated by the non-cooperative game, cooperative game, and Stackelberg game [4]. For the non-cooperative game, prosumers are modeled as a selfish entity with profit competing, and the Nash equilibrium can be achieved accompanied by optimal electricity prices [5]–[6]. Prosumers can also be cooperative as a coalition to participate in the P2P sharing, building a cooperative game model for the P2P coalition and finding an equilibrium solution to keep each coalition stable [7]. Stackelberg game is used to model the interaction between the prosumers and a third party, which manages the energy sharing system and sets pricing strategies for the energy sharing [8]. This framework includes a cooperative type that prosumers can form a coalition to interact with the third party [9], and a non-cooperative type that prosumers are competing during the interaction with the third party [10]–[11]. However, the stochastic factors exist in P2P energy sharing, which affects the optimization results of deterministic game models (i.e., static game). Researches related to stochastic factors in energy sharing mainly study the stochastic factors using static process, and cannot reflect prosumers interactions in decision-making processes through the game process, e.g., for the uncertainties of each DER, the method includes stochastic programming [12], two-stage optimization strategies [13], and reserved regulating capacity mechanisms [14].

The social attributes of prosumers have gradually become a special type of stochastic factor in P2P energy sharing [15]. According to the strong relation between decision theory and psychology, the decision-making process of prosumers is affected by multiple subjective factors [16], which is described by psychology theory [17]–[18]. Besides, prosumers are active participators in energy sharing and can make profits [19]. Thus, with the integration of social attributes, the P2P energy sharing system can be regarded as a cyber-physical-social system (CPSS) in energy [20]. Studies that focus on the prosumers' social attributes can be mainly separated into two parts: model-based method and data-based method. The model-based method relies on the motivational psychology framework and prospect theory (PT). The motivational psychology framework consists of many motivational models, which can combine with the energy sharing model to study the participation willingness of each prosumer in P2P energy sharing [18], [21]. PT describes specific human behavior expressions (i.e., radicalness when losing and conservative when gaining) under risk conditions [22], which has been used in storage batteries scheduling of prospect prosumers [19], one-hour time scale energy optimization with uncertain human behavior [23], and P2P energy sharing with subjective prosumers and time-coupled load constraints [24]. However, specific psychology theory can only reflect limited social attributes of commercial behavior under the risk condition, rather than the diverse social attributes (e.g., family member, income, and power usage habits) of prosumers which could have impacts on the decision of energy sharing.

The data science with the social survey data and other types of data can comprehensively measure the prosumers' social attributes. Only a limited number of researches consider the social preferences of persons through direct data analysis. For example, the cluster of typical consumer types is analyzed based on load metering data and customer consumption features extraction method [25], the EV users' travel willingness is discussed through designing questionnaires to get data to replace decision-making of human participants by multi-agent [26], and the individual driving pattern of EV is studied by statistical methods combined with the national travel survey [27]. However, the existing research only considers single data types and uses statistical summaries or encoding system methods. The integration of social survey data and engineering data is still a problem amidst emerging social-technical complexities.

From the perspective of CPSS, the problems of how to incorporate the social attributes into P2P energy sharing still exist. As social science and energy engineering are two totally different areas, there are two major challenges to be addressed. (1) How to merge the social attributes data and load metering data to quantify the prosumers' social attributes for their subjective decision-making process? (2) The dynamic game model for P2P energy sharing under the incorporation of interactive prosumers with social attributes, in which the prosumers' social attributes reflect as the stochastic factors, so that the game environment will be dynamic, then the static framework is no longer applicable.

To address the aforementioned problems, the formulation is based on a richer class of games, namely stochastic games [28], to capture the dynamic energy sharing environment. A data-driven stochastic game model with social attributes is proposed for P2P energy sharing. We start with the social survey data and load metering data, transform the prosumers' social attributes into subjective probability (SP) through a Spatial-temporal Graph Convolutional Networks (STGCN). The SP can be dynamically modeled among interactive prosumers by the stochastic game with dynamic states and state transition probabilities. This is a new framework to connect the social survey and the P2P energy sharing, which has not been reported, as far as we know. The detailed contributions include two aspects.

(1) A subjective utility model is proposed that considers the prosumers' social attributes in P2P energy sharing. The SP is formulated in the subjective utility model through the STGCN. In the network, a double-layer feature graph is built to learn the prosumers' social attributes based on the social survey data and load metering data, and the relation of prosumers' social attributes and sequential load distributions is studied by the spatial and temporal attention mechanisms.

(2) A stochastic game model is built for P2P energy sharing among prosumers with SPs, where the interactive influence of prosumers is dynamically reflected in the game process by state transitions. The iterative solution algorithm incorporates with STGCN is introduced to get the generalized Nash equilibrium, and the dynamic interval adjustment (DIA) method is designed in the algorithm to solve the game.

II. System Model and Subjective Utility Model

A. Sociality Theoretical Basis

The P2P energy sharing framework includes prosumers and energy sharing provider (ESP). The energy sharing implements among prosumers with the help of the user energy management system (UEMS) [2]. In this framework, each prosumer sets load strategies according to the dynamic prices set by the ESP to maximize their profit. A communication system is existed between ESP and prosumers' UEMS to realize the information exchange, which is realized based on the wireless channels in private 4G/5G networks with Virtual Private Network.

For the development of CPSS, acquiring profit is no longer the only goal of prosumers to participate in energy sharing. According to psychology, other subjective factors affect their decision [29]. Some theories are used to measure a person's subjective action, such as PT [19], theory of reasoned action, and decision theory [17], however, lack of method and interface respective to quantify the social survey data and reflect in the P2P energy sharing. In the rational scenario, each prosumer has equal probabilities for different decisions, which enables prosumers to select decisions totally based on their utility. However, because each prosumer has a different preference for different decisions in the subjective scenario, the probability of choosing each decision is different for prosumers, which reflects as SP. This paper takes the theory in [17] as an example to provide the interface between energy engineering and many kinds of data in social science.

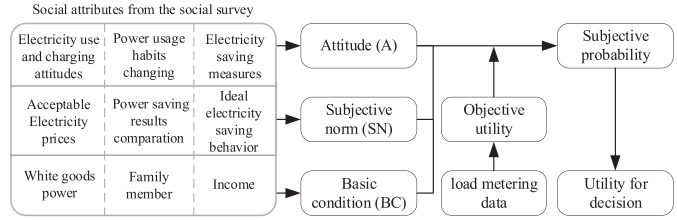

Fig. 1. Subjective decision with SP.

SP can measure an entity's preference for each strategy in the decision, and largely affected by the subjective factors with mapping mechanisms, like attitude (A), basic condition (BC), and subjective norm (SN) [17], which can be obtained through the social survey (Fig. 1). The decision model can be expressed as:

$$p = f(A, BC, SN), \tag{1}$$

$$U = p * u, \tag{2}$$

where p is the SP, and $f(\cdot)$ is the mapping function between subjective factor and probability, U and u are the subjective and objective utility, respectively. Generally, u can be determined by persons' estimation or mathematical economic calculation.

B. Subjective Utility Model of Prosumers

Supposing the prosumers have PV production and load consumption, where the prosumers' load pattern consists of two parts: fixed load and flexible load. There will be multiple possible load strategies of prosumers due to the subjective factors, which are a discrete form and express as the load interval. Each load interval (i.e., prosumers' load strategy) has a unique corresponding SP that presents the prosumers' willingness to choose the strategies.

The prosumers act as a seller (e.g., $E_i^h - L_{i,k}^h > 0$) or a buyer (e.g., $E_i^h - L_{i,k}^h <= 0$) in the system according to the PV production and load consumption. The subjective utility model is formed by three-part series: consume load, share energy, and trade with the utility grid.

$$U_{i,k}^h = \begin{cases} p_{i,k}^h \Big(\delta_i^h ln\Big(1 + L_{i,k}^h\Big) \\ \qquad + pri_{s.k}^h \Big(E_i^h - L_{i,k}^h - \frac{max\left(TNL_k^h, 0\right)}{N_s^h} \Big) \Big) \\ \quad + p_{i,k}^h \frac{\alpha * max\left(TNL_k^h, 0\right)}{N_s^h}, \; E_i^h - L_{i,k}^h > 0 \\ p_{i,k}^h \Big(\delta_i^h ln\Big(1 + L_{i,k}^h\Big) \\ \qquad + pri_{b.k}^h \Big(E_i^h - L_{i,k}^h - \frac{min\left(TNL_k^h, 0\right)}{N_b^h} \Big) \Big), \\ \quad + p_{i,k}^h \frac{\alpha * min\left(TNL_k^h, 0\right)}{N_b^h}, \; E_i^h - L_{i,k}^h <= 0, \end{cases} \tag{3}$$

$$TNL_k^h = TSL_k^h - TBL_k^h, \tag{4}$$

$$TSL_k^h = \sum_{i=1}^{N} max\left(E_i^h - L_{i,k}^h, 0\right), \tag{5}$$

$$TBL_k^h = -\sum_{i=1}^{N} min\left(E_i^h - L_{i,k}^h, 0\right), \tag{6}$$

(1) The first term $\delta_i^h ln(\ 1+L_{i,k}^h)$ is the utility of consuming load, δ_i^h express the sense of getting utility by consuming load. Higher δ_i^h means prosumer wants to consume more load to get more utility. The logarithmic format of the utility function $\delta_i^h ln(\ x)$ has been widely used in economics for modeling the preference of users and the decision-making process [30]. With the development of demand response management in the smart grid, the function that can lead to proportional fair demand response is shown to be suitable for the design of utility for power consumers. However, the users will get a payoff of $-\infty$ when $x = 0$, so the utility function for prosumers has been changed to $\delta_i^h ln(\ 1+x)$ [31], i.e., $\delta_i^h ln(\ 1+L_{i,k}^h)$.

(2) The second term $pri_{s.k}^h\left(E_i^h - L_{i,k}^h - \frac{max(\ TNL_k^h,0)}{N_s^h}\right)$ is the utility of participating in P2P energy sharing, $pri_{s.k}^h$ and $pri_{b.k}^h$ are set by ESP during the decision-making process. The calculation method is based on the Re. [2]:

$$pri_{s,k}^h = \begin{cases} \frac{\lambda_s*\lambda_b}{(\lambda_b-\lambda_s)*SDR_k^h+\lambda_s}, & 0 <= SDR_k^h < 1 \\ \lambda_s, & SDR_k^h <= 1, \end{cases} \tag{7}$$

$$pri_{b,k}^h = \begin{cases} pri_{s,k}^h * SDR_k^h + \lambda_b\left(1-SDR_k^h\right), & 0 = SDR_k^h < 1 \\ \lambda_s, & SDR_k^h <= 1, \end{cases} \tag{8}$$

$$SDR_k^h = \frac{TSL_k^h}{TBL_k^h}, \tag{9}$$

The energy that can use in P2P energy sharing is coming from the netload of each prosumer, but the P2P sharing energy of each prosumer should minus one part of the system's total netload, i.e., $\frac{max(\ TNL_k^h,0)}{N_s^h}$, because of the energy balance requirements of the whole system. This part of the load cannot be satisfied by other prosumers through P2P energy sharing with dynamic prices and require selling to the utility grid with utility grid prices.

(3) The third term $\frac{\alpha*max(\ TNL_k^h,0)}{N_s^h}$ is the utility by sharing the utility of selling netload to the utility grid among sellers, $\frac{\alpha*min(\ TNL_k^h,0)}{N_b^h}$ is the cost by sharing the cost of buying netload to the utility grid among buyers. For sellers, α equals to the utility grid selling prices, and utility grid buying prices for the buyer.

It is noted that the SP affects all three parts of the utility because all the three parts relate to load consumption, which is the decision of prosumers and is affected by the social attributes. Then the influence is reflected in the utility for each prosumer to participate in the P2P energy sharing. Combined with social sciences, the SP is not a fixed distribution, but a subjective value that varies with everyone and every time slot.

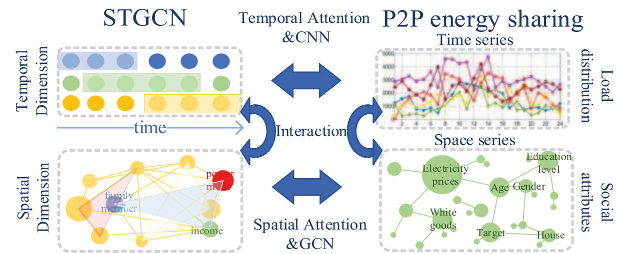

Fig. 2. Applicability of STGCN in the context of energy sharing.

The adjustment of the flexible load is not infinite, there exist lower and upper bound constraints and summation constraints:

$$fl_{i,min}^h <= fl_{i,k}^h <= fl_{i,max}^h, \tag{10}$$

$$\sum_{h=1}^{H} fl_{i,k}^h = TL_{i,k}, \tag{11}$$

To keep the normal operation of the system, the constraints of feeder capacity should also be considered, which limits the sum of all the prosumers netload:

$$\sum_{i=1}^{N}\left(E_i^h - L_{i,k}^h\right) <= C^h. \tag{12}$$

In the subjective environment, the prosumers' willingness for P2P energy sharing will be affected by their social attributes. However, the subjective utility can consider the influence of prosumers' social attributes and the gaining profit during energy sharing. Therefore, if the subjective utility can be increased through participating in the P2P energy sharing, prosumers will be willing to join the energy sharing.

III. Social Preference Based on Learning Method

A. *Applicability of STGCN in the Context of Energy Sharing*

In the context of energy sharing with subjective prosumers, prosumers' decision-making is highly affected by their social attributes, so that the load distribution data and social attributes data should be considered at the same time. These two kinds of data are highly different, load metering data is continuous data, while social survey data is categorical data. General neural networks usually learn features of input data together based on their quantities, which makes it difficult to process data of different nature. Therefore, it is better to process these two kinds of data in different dimensions to distinguish them.

The STGCN can learn the invisible correlation from time horizon and spatial domain, consider the spatial and time dependence at the same time, capture the dynamic spatial-temporal (ST) characteristic, and accurately predict the information of future moments [32]. As Fig. 2 shows, in the energy sharing, the relation of social attributes and load distribution is reflected as the temporal and spatial dimension then modeled by the STGCN. Prosumers' load distribution will present specific characteristics in one day, which is a kind of time series and regard as the temporal dimension. It is noted that the load distributions are affected by the social attributes, which can be regarded as the variables in the spatial dimension. Based on statistical social attributes data and load distribution data, a double-layer feature graph can be built during the learning process of STGCN to study the relations of prosumers' social attributes and sequential load distributions, then getting the SP.

B. The Input Features of STGCN With Data Pre-Processing

An undirected complete graph is defined as $G = (V, E, \boldsymbol{A}_{dj})$, in which V is a finite set with the number of sN nodes, E is the set of edges, $\boldsymbol{A}_{dj}$ is the adjacent matrix of G [33]. The dynamic node feature matrix of graph G includes two parts: social attributes in the spatial domain and load distribution in time horizon, which respectively expressed by the social attributes matrix and load distribution matrix. Graph G is physically extended to a double-layer feature graph based on these two matrices. Data pre-processing of social survey data and load distribution data is required for the establishment of these two feature matrices.

1) Social Attributes Matrix: The original social attributes data from the social survey is disorder, which should be classified and selected. A feature combination is composed of a set of social survey questions and define as one type of prosumers' social attribute. Personal social attributes affect subjective strategies. Each prosumer has sN types of social attributes (i.e., feature combinations) in the STGCN network. sM is the length of the longest feature combination in the matrix, which is calculated by:

$$sM = max\left\{len\left(\left\{\boldsymbol{X}^{j}_{\mathrm{S},hs}|j \in [1, sN]\right\}\right)\right\}, \tag{13}$$

where $len(\cdot)$ calculates the length.

Once the value of sM is confirmed, all the feature combinations are processed by zero paddings. Therefore, the social attributes matrix for each prosumer can be expressed as follows.

$$\boldsymbol{\chi}_{\mathrm{S}} = \begin{bmatrix} \boldsymbol{X}^{1}_{\mathrm{S},1} & \cdots & \boldsymbol{X}^{j}_{\mathrm{S},1} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{S},1} \\ \cdots & & \cdots & & \cdots \\ \boldsymbol{X}^{1}_{\mathrm{S},hs} & \cdots & \boldsymbol{X}^{j}_{\mathrm{S},hs} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{S},hs} \\ \cdots & & \cdots & & \cdots \\ \boldsymbol{X}^{1}_{\mathrm{S},Hs} & \cdots & \boldsymbol{X}^{j}_{\mathrm{S},Hs} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{S},Hs} \end{bmatrix} \in Hs \times sN \times sM,$$

$$\boldsymbol{\chi}_{\mathrm{S},hs} = \left[\boldsymbol{X}^{1}_{\mathrm{S},hs} \cdots \boldsymbol{X}^{j}_{\mathrm{S},hs} \cdots \boldsymbol{X}^{sN}_{\mathrm{S},hs}\right] \in sN \times sM, \tag{14}$$

$$\boldsymbol{X}^{j}_{\mathrm{S},hs} = \left[x^{1,j}_{\mathrm{S},hs}, \ldots, x^{c,j}_{\mathrm{S},hs}, \ldots, x^{sM,j}_{\mathrm{S},hs}\right]^{T} \in 1 \times sM, \tag{15}$$

The $\boldsymbol{\chi}_{\mathrm{S},hs}$ can be dynamically changed in different time slice or keep unchanged in the whole time span, which is determined by the social attributes data.

2) Load Distribution Matrix: The initial load data is aggregated to an hourly expression to support the energy scheduling. To get the $\boldsymbol{\chi}_{\mathrm{Ls}}$, data pre-processing of the hourly load is required on the basis of statistical probabilities. The maximum load value is divided into K load intervals for each time slot. In the time slice hs, interval k contains two types of information: the middle load value $x^{sM+1,k}_{\mathrm{Ls},hs}$ and probability $x^{sM+2,k}_{\mathrm{Ls},hs}$ that reflect the possibility of the load value falls in this interval. It is noted that the interval k is corresponding to feature combination j.

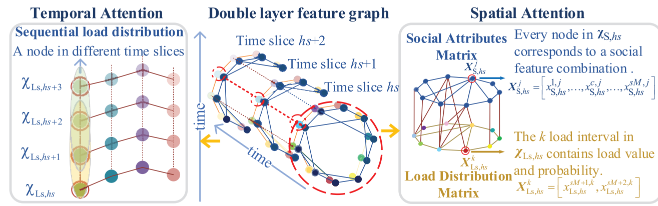

Fig. 3. The special and temporal attention mechanism.

The load distribution matrix consists of the prosumers' social preference for load strategies, which is expressed as:

$$\boldsymbol{\chi}_{\mathrm{Ls}} = \begin{bmatrix} \boldsymbol{X}^{1}_{\mathrm{Ls},1} & \cdots & \boldsymbol{X}^{k}_{\mathrm{Ls},1} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{Ls},1} \\ \cdots & & \cdots & & \cdots \\ \boldsymbol{X}^{1}_{\mathrm{Ls},hs} & \cdots & \boldsymbol{X}^{k}_{\mathrm{Ls},hs} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{Ls},hs} \\ \cdots & & \cdots & & \cdots \\ \boldsymbol{X}^{1}_{\mathrm{Ls},Hs} & \cdots & \boldsymbol{X}^{k}_{\mathrm{Ls},Hs} & \cdots & \boldsymbol{X}^{sN}_{\mathrm{Ls},Hs} \end{bmatrix} \in Hs \times sN \times 2,$$

$$\boldsymbol{\chi}_{\mathrm{Ls},hs} = \left[\boldsymbol{X}^{1}_{\mathrm{Ls},hs} \cdots \boldsymbol{X}^{k}_{\mathrm{Ls},hs} \cdots \boldsymbol{X}^{sN}_{\mathrm{Ls},hs}\right] \in sN \times 2,$$

$$\boldsymbol{X}^{k}_{\mathrm{Ls},hs} = \left[x^{sM+1,k}_{\mathrm{Ls},hs}, x^{sM+2,k}_{\mathrm{Ls},hs}\right]^{T} \in 1 \times \ 2.$$

C. The Structure of ST Blocks in STGCN

Based on graph G, STGCN aims at extracting node features. The structure of STGCN contains several ST blocks and a fully connected (FC) layer. Each ST block includes 4 parts: spatial attention mechanism, temporal attention mechanism, convolution neural network (CNN), and graph convolution network (GCN).

Graph G includes a spatial dimension to explore the dynamic impact of the social attributes on load probability distribution through learning dynamic weighting. To better capture the dynamic ST correlation, two attention mechanisms are adopted in this paper: spatial attention mechanism and temporal attention mechanism, which is shown in Fig. 3.

1) Spatial Attention Mechanism: In the spatial dimension, the nodes project from the temporal dimension, which includes their own social attributes and the preference of a certain load strategy. The spatial attention mechanism (Fig. 3 right) is adopted based on the build double-layer feature graph, which is used to adaptively capture the dynamic correlation between nodes, then promote the accuracy of the prediction. Two types of influence are considered and express as two matrices: the dynamic interactive influence between social attributes, and between load intervals, which comes from the natural language processing area that can find dependency relationships based on training data [34].

$$\boldsymbol{S}_{\mathrm{amls}} = \left(\boldsymbol{\chi}^{r-1}_{\mathrm{Ls}}\boldsymbol{W}_{1L}\right)\boldsymbol{W}_{2L}\left(\boldsymbol{W}_{3L}\boldsymbol{\chi}^{r-1}_{Ls}\right)^{T}, \tag{16}$$

$$\boldsymbol{S}_{\mathrm{ams}} = \left(\boldsymbol{\chi}^{r-1}_{S}\boldsymbol{W}_{1S}\right)\boldsymbol{W}_{2S}\left(\boldsymbol{W}_{3S}\boldsymbol{\chi}^{r-1}_{S}\right)^{T}, \tag{17}$$

where $\boldsymbol{W}_{1L}$, $\boldsymbol{W}_{2L}$, $\boldsymbol{W}_{3L}$, $\boldsymbol{W}_{1S}$, $\boldsymbol{W}_{2S}$, $\boldsymbol{W}_{3S}$ are the parameters of learning model, $\boldsymbol{\chi}^{r-1}_{S} = (\boldsymbol{X}^{1}_{S}, \boldsymbol{X}^{2}_{S}, \ldots, \boldsymbol{X}^{Hs_{r-1}}_{S})$ is the element of $\boldsymbol{\chi}_{\mathrm{S}}$, which is the input data of spatial layer r in the STGCN.

Based on the $\boldsymbol{S}_{\mathrm{amls}}$ and $\boldsymbol{S}_{\mathrm{ams}}$, the dynamic spatial attention matrix $\mathfrak{R}$ is calculated by the Hadamard product [35], to extract the connection between social attributes and load strategy preference. The weight relationships between each node in spatial dimension can be got, which is dynamically adjusted with the spatial attention matrix $\mathfrak{R}$ and adjacent matrix A_{dj} in the process of graph convolution. The matrix $\mathfrak{R}$ should be normalized as the input of the spatial layer, and the value of the element $\mathfrak{R}'^{j_1,j_2}$ semantically represents the strength of the dependencies at the node j_1, j_2.

$$\mathfrak{R} = \boldsymbol{V}_s \cdot \sigma(\boldsymbol{S}_{\text{amls}} \odot \boldsymbol{S}_{\text{ams}} + \boldsymbol{b}_s), \tag{18}$$

$$\mathfrak{R}'^{j_1,j_2} = \frac{exp\left(\mathfrak{R}^{j1,j2}\right)}{\sum_{j2=1}^{J} exp\left(\mathfrak{R}^{j1,j2}\right)}, \tag{19}$$

where $\boldsymbol{V}_s$, $\boldsymbol{b}_s$ are the parameters of the learning model, $\sigma(\cdot)$ is the activation function, and $\odot$ is the Hadamard product.

2) Temporal Attention Mechanism: Fig. 3 (left) shows the temporal attention mechanism in the time horizon, and the node weight is different between each time slot. The dynamic temporal attention matrix $\mathfrak{I}$ should be normalized as the input of the STGCN in the temporal dimension. The value of the element $\mathfrak{I}'^{hs_1,hs_2}$ semantically represents the strength of the dependencies at the time slice hs_1, hs_2:

$$\mathfrak{I} = \boldsymbol{V}_e \cdot \sigma\left(\left(\left(\left(\boldsymbol{\chi}_{Ls}^{r-1}\right)^T \boldsymbol{Ua}_1\right)\boldsymbol{Ua}_2\left(\boldsymbol{Ua}_3\boldsymbol{\chi}_{Ls}^{r-1}\right) + \boldsymbol{b}_e\right), \tag{20}$$

$$\mathfrak{I}'^{hs_1,hs_2} = \frac{exp\left(\mathfrak{I}^{hs1,hs2}\right)}{\sum_{hs_2=1}^{Hs_{r-1}} exp\left(\mathfrak{I}^{hs1,hs2}\right)}, \tag{21}$$

where the $\boldsymbol{V}_e$, $\boldsymbol{b}_e$, $\boldsymbol{Ua}_1$, $\boldsymbol{Ua}_2$ $\boldsymbol{Ua}_3$ are all the parameters of the learning model, $\boldsymbol{\chi}_{Ls}^{r-1} = (\boldsymbol{X}_{Ls}^1, \boldsymbol{X}_{Ls}^2, \ldots, \boldsymbol{X}_{Ls}^{Hs_{r-1}})$ is the element in $\boldsymbol{\chi}_{\text{Ls}}$, which is the input data of temporal dimension in layer r.

3) GCN and CNN: GCN and CNN are the learning components in the ST block for extracting spatial and temporal information [36]–[37].

D. Learning Process in STGCN

The STGCN is shown in Fig. 4. The node feature matrix $\boldsymbol{\chi}$ is the initial input of the entire STGCN. The spatial dimension of $\boldsymbol{\chi}$ is handled by GCN while the temporal part is handled by CNN for extracting more information about social attributes and load distribution. The extracting process of GCN and CNN is respectively assisted by spatial attention mechanism and temporal attention mechanism.

In the spatial dimension, the dynamic correlation between each node is captured by the spatial attention mechanism, which is expressed as the spatial attention matrix after normalization $\mathfrak{R}'$. The matrix can learn the node weights considering social attributes and load distribution, which can help GCN explore the node features of $0 \sim (k-1)$ adjacent nodes, and enhance the efficiency of feature extraction. The node features are transferred to the temporal dimension for the learning process of CNN.

In the temporal horizon, $\widehat{\boldsymbol{\chi}}_{Ls,hs}^{(r-1)}$ is an intermediate variable in the learning network for dynamically adjusting the input of the temporal dimension. The variable is calculated by the dynamic temporal attention matrix $\mathfrak{I}'$ and the input data $\boldsymbol{\chi}_{Ls}^{r-1}$:

$$\begin{aligned}\widehat{\boldsymbol{\chi}}_{Ls,hs}^{(r-1)} &= \left(\widehat{\boldsymbol{X}}_{Ls}^1, \widehat{\boldsymbol{X}}_{Ls}^2, \ldots, \widehat{\boldsymbol{X}}_{Ls}^{Hs_{r-1}}\right) \\ &= \left(\boldsymbol{X}_{Ls}^1, \boldsymbol{X}_{Ls}^2, \ldots, \boldsymbol{X}_{Ls}^{Hs_{r-1}}\right) \cdot \mathfrak{I}'. \end{aligned} \tag{22}$$

The learning process of CNN is based on the spatial node features extracted by GCN, which realize the connection of spatial and temporal horizons. The node features are updated by merging relevant information on the adjacent time slots through CNN. The spatial and temporal attention mechanism enhances the nodes' correlation during the learning process and improves the training accuracy of STGCN. Finally, FC and activation function rectified linear unit (ReLU) are used to normalize the output, which includes the predicted results of the SP and the corresponding load value. These two results are output as matrices with H time slots and K load intervals.

IV. Stochastic Game Formulation and Solution Algorithm

A. Formulation of the Stochastic Game Model

Prosumers' load strategies are continuous in the time horizon, and the social attributes in the previous time slots have an influence on the next time slot for continuous learning of SP, Besides, many possible scenarios will arise from the subjective uncertainties. Therefore, the energy scheduling environments are dynamic in the time horizon, and the stochastic game model is built to formulate the stochastic and dynamic process of P2P energy sharing. Different from the general static game model, it includes five elements: game player, action, state, state transition probability, and utility. The game implements in a dynamic way with multi-stage, in which the stage is corresponding to the time slot in energy scheduling. SP dynamically change in different stages, resulting in the different states (i.e., scenarios) in each stage, and the state will change in the next stage with the state transition probability. Based on the subjective utility model of each prosumer, the game model is expressed as follows:

$$\boldsymbol{\theta} = \left\{N, \boldsymbol{S^h}, \boldsymbol{A^h}, \boldsymbol{P^h}\left(A_{i,k}^h \;\middle|\; S_i^{h-1}\right), \boldsymbol{U^h}, h \in H\right\}, \tag{23}$$

$$S_{i,k}^h = \left\{L_{i,k}^h, p_{i,k}^h\right\}, S_{i,k}^h \in \boldsymbol{S^h}, \tag{24}$$

$$A_{i,k}^h = \left\{L_{i,k}^h | k \in [1, K]\right\}, A_{i,k}^h \in \boldsymbol{A^h}, \tag{25}$$

$$P_{i,k}^h\left(L_{i,k}^h \;\middle|\; S_i^{h-1,*}\right) = \frac{p_{i,k}^h \prod_{h=1}^{H-1} p_{i,k}^h}{\prod_{h=1}^{H-1} p_{i,k}^h} = p_{i,k}^h, \tag{26}$$

$$U_{i,k}^h(\text{see Eq.(3)}), U_{i,k}^h \in \boldsymbol{U^h}, \tag{27}$$

where the superscript h is the stage index of the dynamic game, N is the prosumers set, and they are the players of the game. $\boldsymbol{S^h}$ is the state set of the game, and each state includes load patterns and SP. $\boldsymbol{A^h}$ is the action set, which includes load consumption. $\boldsymbol{P^h}(A_{i,k}^h | S_i^{h-1,*})$ is the state transition probability, which is based on SP that acquires from the STGCN and implicates the probability of a chosen action, where $S_i^{h-1,*}$ is the optimal state of time slot *h-1*. According to the definition of the stochastic game [28], the state transition probability is a conditional probability related to the previous state $S_i^{h-1,*}$. $\boldsymbol{U^h}$ is

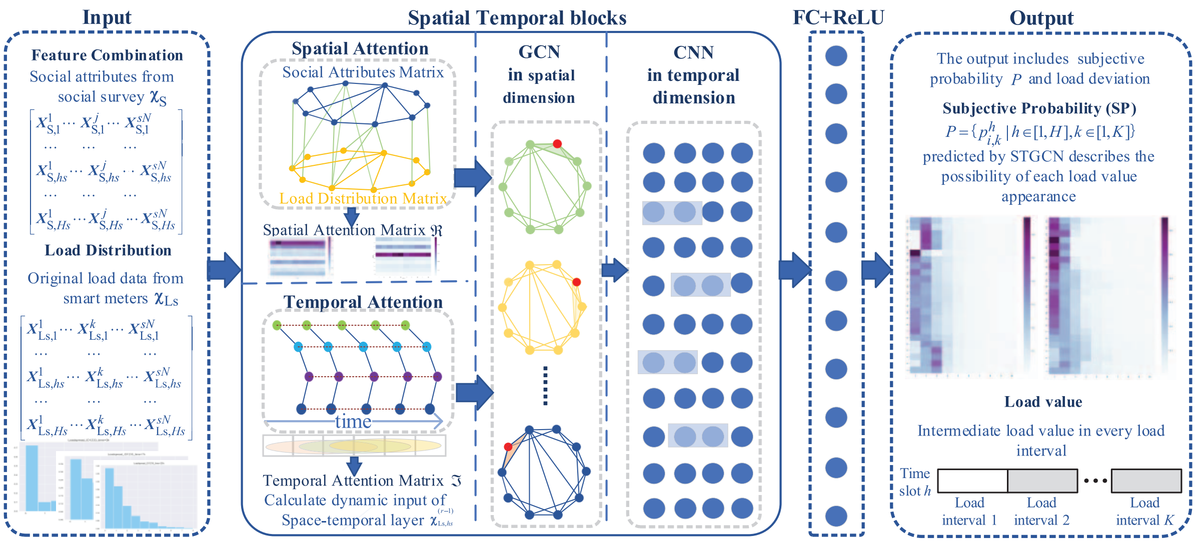

Fig. 4. The STGCN for learning of SP.

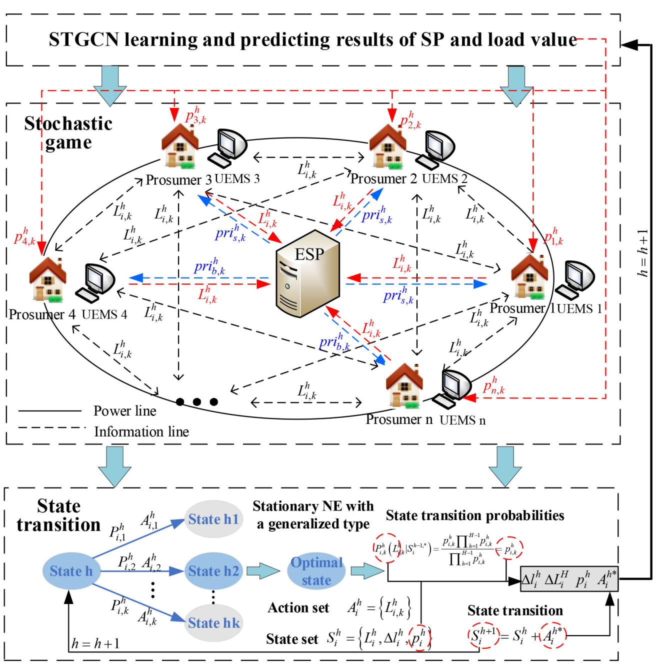

Fig. 5. The stochastic game process.

the utility set of prosumers. For each prosumer, the utility is based on the energy sharing model, which aims at maximizing the subjective utility under the consideration of SP.

Fig. 5 shows the stochastic game process. The game is implemented between prosumers, and an ESP is a middleman (not a game player) to support the game process. During the game, ESP calculates the dynamic electricity prices and transmits them to prosumers through a communication system, while prosumers determine their load actions based on their utility under the influence of SP. It is noted that only the load strategies of each prosumer should transmit to the ESP, which are all general information shared in the game model [38], so that doesn't cause privacy issues related to the prosumer's social attributes.

Prosumers' actions (i.e., load strategies) will cause state transitions in the current time slot following the state transition probability, which determines the states in the next time slot. The load actions in the next time slot will also be affected because of the time-coupled constraints (11). Besides, the SP in the next time slot is predicted based on the last time slot's updated probability in the temporal horizon and historical learnable parameters. By regarding all prosumers as a cluster, then the solving process between different time slots is a single agent process, which satisfies the Markov decision process (MDP). The MDP following the optimality principle of dynamic programming, i.e., the sub-result of the optimal results is always optimal and called as 'without aftereffect' [39]. Therefore, the stochastic game implements as MDP accompanied by state transitions and SP learning, until reaches H time slot.

B. Nash Equilibrium

The optimal action in each time slot is Nash equilibrium (NE). From a general view of NE, the decision of prosumer i includes the load strategies in response to the decisions of other prosumers because of the coupled constraints (12), so that the NE will become a generalized type [6].

Definition 1: A set of optimal action $A_i^{h*} = \left\{L_i^{h*}|h = 1, 2, \ldots H\right\}$ of player i is the NE with a generalized type in time slot h , accompanied with an optimal dynamic selling and buying prices pri_s^h and pri_b^h, if and only if the action of player i is the optimal action of the combination of other prosumers action, that is the following expression is satisfied:

$$U_i^h\left(A_1^{h*}, \ldots, A_i^{h*}, \ldots A_N^{h*}\right)$$
$$<= U_i^h\left(A_1^{h*}, \ldots, A_{i-1}^{h*}, A_i^{h*}, A_{i+1}^{h*}, \ldots A_N^{h*}\right)$$
$$\forall i \in N, h \in H, A_i^{h*} \in \boldsymbol{A^h}. \tag{28}$$

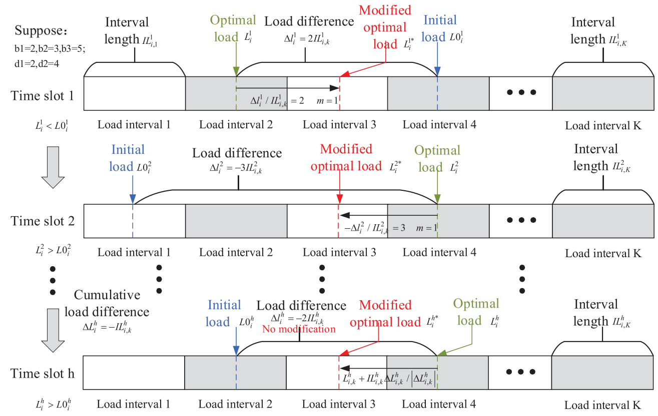

Fig. 6. The DIA method.

Theorem 1: The unique NE with a generalized type always exists in each time slot in the proposed stochastic game $\boldsymbol{\theta}$.

Proof: The proof is based on two aspects. (i) The action set of each prosumer is nonempty, compact, and convex; (ii) The utility function is always concave for each prosumer [6], [19]. For the first aspect, it is obvious that the utility function of prosumers (3) is a continuous function, and the set is nonempty, compact, and convex. Then, the concavity and convexity of prosumers' utility function in each load interval k should be clarified. For both sellers and buyers, the Hessian matrix with respect to $L_{i,k}^h$ is expressed as:

$$HM = -\delta_i^h / \left(1 + L_{i,k}^h\right)^2. \tag{29}$$

The Hessian matrix is negative definite with respect to $L_{i,k}^h$, and thus $U_{i,k}^h$ is a strictly concave function.

$U_{i,k}^h$ should be maximized during the game process, therefore, there exists a unique NE with a generalized type and Theorem 1 is proved. ■

C. Solution Algorithm With DIA

During the game, it is difficult to satisfy the time-coupled load constraints (11), because the game $\boldsymbol{\theta}$ implements hour-by-hour. The DIA method is adopted during the decision-making process of each time slot, which is a kind of heuristic method and is shown in Fig. 6. The method is conducted based on two mechanisms: (i) forward compensation to increase the optimal load L_i^{h*} for acquiring higher load value, and (ii) backward curtailment to decrease the optimal load for limiting the load value. There are two kinds of adjustment:

1) Inside a Time Slot: There is a load difference $\Delta l_i^h = L0_i^h - L_i^h$ between the initial reference load $L0_i^h$ and the optimal load L_i^{h*} in a time slot. The optimal load decreases or increases m intervals based on the load difference to get the adjustment load $\Delta l_i^{h'}$, and the load will be further adjusted as Eq. (31) shows to get $\Delta l_i^{h''}$. The adjustment is in a wider range to reduce the deviation of flexible load, and further satisfy the time coupled constraints. Then the corresponding probability also changes to the interval after adjustment.

$$\begin{aligned}
&\underset{\Delta l_i^h <= 0}{decr} \begin{cases} m = 1, & b_1 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h} <= b_2 \\ m = 2, & b_2 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h} <= b_3 \\ m = 3, & b_3 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h}, \end{cases} \\
&\underset{\Delta l_i^h > 0}{incr} \begin{cases} m = 1, & d_1 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h} <= d_2 \\ m = 2, & d_2 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h} <= d_3 \\ m = 3, & d_3 < \dfrac{-\Delta l_i^h}{IL_{i,k}^h}, \end{cases}
\end{aligned} \tag{30}$$

$$\underset{\Delta l_i^{h'} <= 0}{decr} \begin{cases} m = 1, & k > 0 \\ m = 2, & k > 1, \end{cases} \underset{\Delta l_i^{h'} > 0}{incr} \begin{cases} m = 1, & k < 9 \\ m = 2, & k < 8. \end{cases} \tag{31}$$

2) Inter Time Slots: The cumulative load difference ΔL_i^H also exists. Because the major adjustment conducts inside each time slot and great adjustment may affect the decision in the current time slot, the adjustment range is half or one interval. When $|\Delta L_i^H| < 0.5IL_{i,k}^h$, the forward compensation and backward curtailment are conducted inside an interval, when $|\Delta L_i^H| <= 0.5IL_{i,k}^h$, the adjustment range is one interval, and the adjustment is conducted many times, which is automatically determined by the utility. Then the optimal load can be modified as:

$$\begin{cases} L_{i,k}^{h'} = L_{i,k}^h + \dfrac{0.5IL_{i,k}^h \Delta L_i^H}{|\Delta L_i^H|}, \; k' = k, \; |\Delta L_i^H| < 0.5IL_{i,k}^h \\ L_{i,k}^{h'} = L_{i,k}^h + \dfrac{IL_{i,k}^h \Delta L_i^H}{|\Delta L_i^H|}, \quad k' = k + \dfrac{\Delta L_i^H}{|\Delta L_i^H|}, \; |\Delta L_i^H| \\ \qquad\qquad\qquad\qquad\qquad\qquad <= 0.5IL_{i,k}^h, \end{cases} \tag{32}$$

$$\Delta L_i^H = \sum_{h=1}^{H-1} \Delta l_i^{h''}. \tag{33}$$

There are parameters that affect the adjustment interval to satisfy the time coupled constraints during the game process. Some pre-analysis can be conducted to choose the best parameters for the specific application, then get better performance for the algorithm, like many heuristic methods [40]. In the energy sharing application, the sensitiveness analysis is listed in the Case Study.

In the stochastic game model, the stochastic factors result in the problem of the "curse of dimensionality" when solving the game. Because each prosumer has K possible actions, which will cause K possible states, and N prosumers will have K^N possible states. If simultaneously considering all prosumers' actions, the number of states exponential growth with the actions K and prosumers N in one stage. The problem makes the game cannot be directly solved by the mathematical method.

The solution algorithm solves the problem by the iteration method. Each prosumer separately decides their strategies and updates their strategies through the iteration with ESP. When each prosumer decides their strategies, the state transition probability is also considered, and the DIA is conducted to consider the time coupled constraints. The dynamic prices serve as the intermediary to connect each iteration and each

Algorithm 1 Iterative Solving Method With DIA for Game $\boldsymbol{\theta}$

1. Setting the utility buying and selling prices λ_b, λ_s, initial load distribution $L_{0,i,k}^h$, preference parameter of prosumer δ_i^h, constraints parameter $fl_{i,min}^h$, $fl_{i,max}^h$, $TL_{i,k}$, C^h.

2. Initial the STGCN network, and training it based on $\boldsymbol{\chi}_S$, $\boldsymbol{\chi}_{Ls}$, then getting the weights and learnable parameters.

3. Conduct the stochastic game $\boldsymbol{\theta}$ with the learning model.

For time slot $h = 1$

4. ESP set the utility grid prices as the initial dynamic prices.

For iteration $iter = 1$

(1). Based on the dynamic prices, each prosumer gets multiple utilities $U_{i,k}^{h,iter}$ in each interval k followed the strategy $L_{i,k}^h$ and corresponding SP $p_{i,k}^h$.

(2). Calculate the load difference Δl_i^h and conduct the adjustment based on the DIA.

(3). Choosing an optimal utility $U_i^{h,iter*}$ that satisfies the following condition, and gets corresponding optimal strategy L_i^{h*} and SP p_i^{h*}.

$$U_i^{h,iter*} > \max_k U_{i,k}^{h,iter}$$

(4). Based on the prosumers' load strategy, ESP calculates dynamic selling and buying prices $pri_{s,k}^h$, $pri_{b,k}^h$ according to Eqs. (7-9), then transmit them to prosumers.

If $\left|U_i^{h,iter*} - U_i^{h,iter-1*}\right| <= 1e-4, \forall i \in N$

Reach the generalized Nash equilibrium in time slot h;

Break;

End if

End for iteration

5. Pass the SP $p_i^{h*}, \forall i \in N$ and load strategies $L_i^{h*}, \forall i \in N$ in time slot h to predict the SP in time slot $h+1$ and satisfy the constraints (11).

If $h = H$

stop the game process;

Break;

End if

End for

6. Get the load strategies $\boldsymbol{L}_i^*$ and corresponding SP $\boldsymbol{p}_i^*$ of all prosumers in the whole time slot.

prosumer. Then the K^N solution space (i.e., possible states) can be reduced to $K \times N \times iter^h$ in one time slot. Besides, the prosumers coupled constraint (12) is also considered in the decision-making process through the iteration in each stage, where each prosumer should satisfy the constraint when they separately optimize their load strategies. The solution algorithm is expressed as follows.

It is noted that the solution to the optimization model and heuristic DIA are included in the iterative process. There are negative feedbacks between dynamic selling/buying prices and load strategies during iteration to guarantee the convergence of the algorithm. The dynamic prices are calculated as Eqs. (7)-(9) shows, it is noted that the SDR determines the prices and is affected by the total selling energy TSL_k^h and buying energy TBL_k^h. When the SDR increases, it can be equivalent to an increase in TSL_k^h or a decrease in TBL_k^h, and both the dynamic selling prices and buying prices will decrease. Therefore, the increased TSL_k^h is restrained because selling energy can get less profit, and the decreased TBL_k^h is also restrained because buying energy requires less cost. When the SDR decreases, the TSL_k^h can be regarded as a decrease or TBL_k^h regarded as an increase, and both dynamic selling and buying prices increase. Therefore, the TSL_k^h increase to get more selling profit and TBL_k^h decrease to avoid more cost.

V. Case Study

A. Basic Data

The P2P energy sharing test system consists of an ESP and 11 resident prosumers with PV production connected to a medium-voltage (10 kV) feeder of a distribution network. Each prosumer on the STGCN learning network includes load probability distribution and social attributes features. The load data and prosumers social attributes were taken from the Smart Metering Electricity Customer Behavior Trials (CBTs) by Commission for Energy Regulation of Irish Social Science Data Archive [41], where the social survey is conducted during 2009 and 2010 with over 5,000 Irish homes and businesses participating and more than 600 questions. 10 types of social attributes can be extracted from 90 questions through the data pre-processing method: family member, income, white goods power, small household appliances' power, acceptable electricity prices, power usage habits changing, electricity saving measures, ideal electricity saving behavior, power saving results comparison, and electricity use and charging attitudes. The hourly load data is collected by smart meters, which includes 17 months. Besides, the number of time slots for energy sharing is 24 hours, and each prosumer has 10 possible load strategies in one time slot. The data are uploaded in the ScholarOne Manuscripts system.

B. Results of SP and Comparison With Other Neural Networks

1) SP From STGCN: In the learning network, the number of prosumers' social attributes is 10, each social attribute has 13 dimensions. Therefore, the feature of the learning network can be got, which is also the input of the graph G. The load distribution data is a time series that vary in time, while the social survey is only conducted once, so that the social attributes data is static, which means different rows of the social attributes matrix is the same. The parameter of the learning network is set as: the horizontal dimension of input channels is 15, the terms of Chebyshev polynomial is 4, convolution kernels of GCN and CNN are 16 and 4, the batch size is 2, and the learning rate is 0.0001. Based on graph G, two spatial attention matrices are respectively calculated between social attributes and between load probability distribution.

Then according to Eq. (18), the spatial attention matrix between social attributes and the load probability distribution is got by Hadamard product of the above two matrices, which is shown in Fig. 7 with prosumer ID 5407. The element represents the correlation between load distribution and social attributes, and they both affect the SP. Like node (1,1) is got from Hadamard product by two parts: (i) the correlation

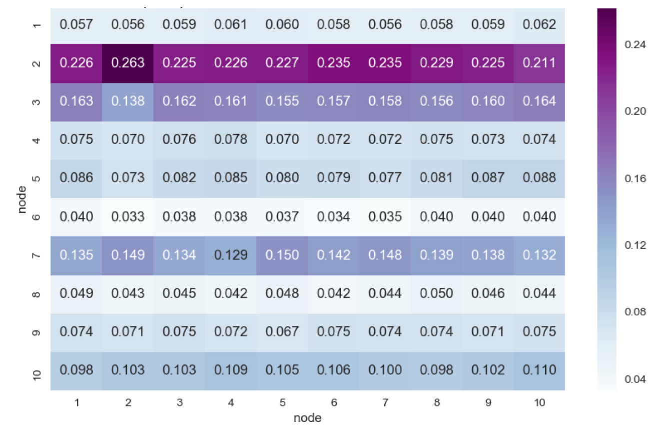

Fig. 7. The spatial attention matrix of prosumer ID5407.

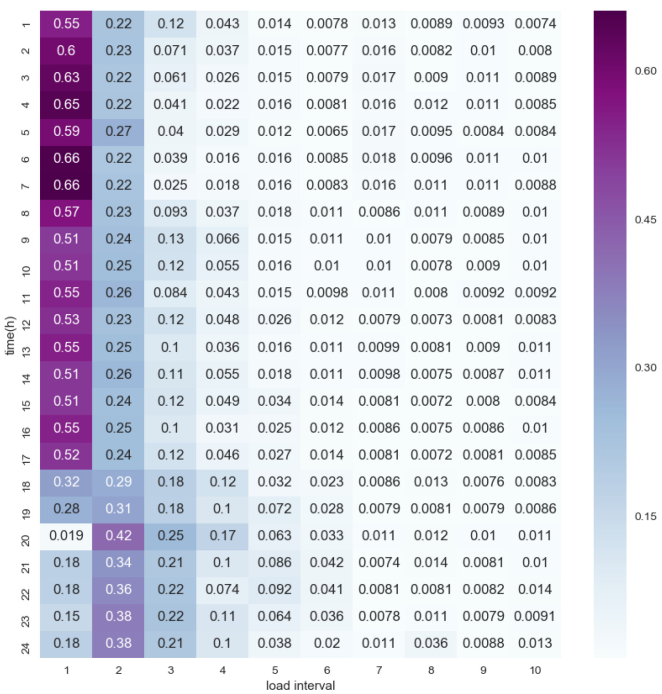

Fig. 8. The SP of prosumer ID 5407.

between load interval 1 and other load intervals in matrix χ_{Ls}, (ii) the correlation between social attribute 1 with other social attributes in matrix χ_S. Therefore, the number 0.057 indicates the comprehensive influence between these two matrices and on the SP, and the influence is less than the node (2,2) 0.263.

After the learning process, the SP can be got with the game occurring. Fig. 8 shows the prosumer ID5407 SP as a thermodynamic diagram. It is clear that there is no regular distribution that can quantify the SP because it is largely affected by the prosumer's social attributes. In time slots 18-24, the probability is larger in second than the first load interval because the initial load peak appears in the evening, indicating the prosumer subjective preference to each load value. Besides, the SP is concentrated on the first three load intervals, due to the rare appearance of large load strategies.

2) Comparison With Other Neural Networks: To show the effectiveness of STGCN, we have conducted the comparison

TABLE I
MAE OF LEARNING RESULTS

ID	STGCN	LSTM		GRU	
		Univariate	Multivariate	Univariate	Multivariate
1233	0.0334	0.0429	0.0408	0.0398	0.0402
1506	0.0289	0.0371	0.0361	0.0344	0.0347
1846	0.0359	0.0497	0.0475	0.0465	0.0471
2086	0.0385	0.0506	0.0484	0.0474	0.0481
2379	0.0381	0.0528	0.0531	0.0493	0.0499
2499	0.0294	0.0381	0.0370	0.0361	0.0366
2850	0.0259	0.0405	0.0382	0.0370	0.0374
4509	0.0269	0.0363	0.0352	0.0338	0.0347
4563	0.0353	0.0516	0.0488	0.0481	0.0485
5363	0.0305	0.0430	0.0409	0.0399	0.0407
5407	0.0253	0.0360	0.0347	0.0337	0.0341

results with the long short-term memory (LSTM) network [42] and gated recurrent unit (GRU) network [43]. Two kinds of inputs are designed: one is to input only load distribution (i.e., univariate with one time series feature), the other is to input both social attributes data and load distribution data in the form of one-dimensional matrix (i.e., multivariate with multiple time series features). Mean absolute error (MAE), which is a classic evaluation index, is used as the evaluation metric, and the learning results are shown in Table I.

Through the comparison, STGCN is superior to the LSTM and GRU, achieving the best results. The reason is that the STGCN model can simultaneously process these two kinds of data in the spatial and temporal dimension, which helps STGCN extracts abundant information. During the learning process, social attributes and load distribution are linked through a double-layer feature graph to learn the SP with the help of attention mechanisms that can capture the dynamic changes of social attributes data and load distribution data.

C. Results of Stochastic Game and Social Preference Analysis

The results of the stochastic game include load strategies of prosumers, dynamic prices set by ESP, and the utility of prosumers. To show the influence of considering prosumers' social preference, the results of subjective scenarios are compared with the rational scenarios, which are conducted by the non-cooperative game for maximizing utilities. The initial load and utility are got from the median of the 17 months data.

1) Results of Load Strategies: Fig. 9 shows the load strategies of prosumers with ID5407. The major fluctuations of the curve arise from discrete load strategies. It is clear that the optimal load is almost increased in time slots 8-15, which has a higher PV production. Prosumers can increase their utility by avoiding buying energy in other time slots where PV production is insufficient. The load peak in time slots 18-23 is also reduced by the optimization, because these time slots have no PV production, and the buying prices are higher. However, the load in time slots 16-17 is a particular case, this phenomenon is influenced by SP and DIA. From Fig. 8, the SP in the second load interval of time slots 16-17 is smaller than time slots after 18, which is the peak time slots. Besides, the positive load deviation is accumulated in the time slots 8-15 according to the DIA. Therefore, the prosumer abandons the larger load

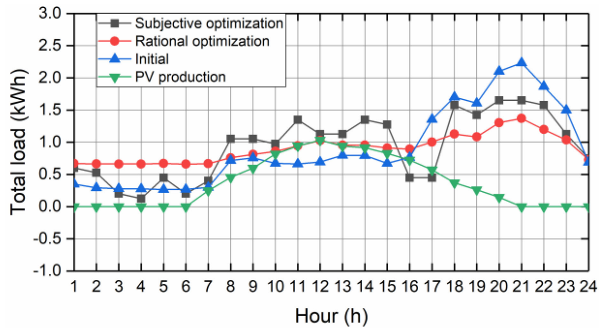

Fig. 9. The load strategies of prosumer ID5407.

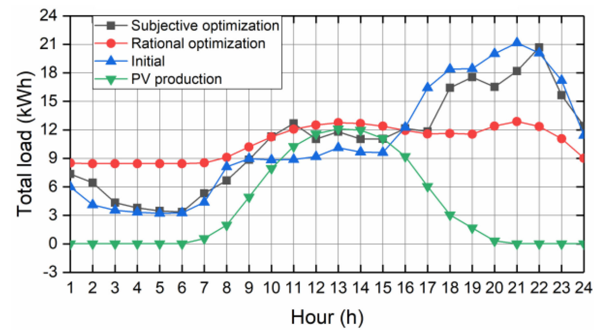

Fig. 10. The load strategies of all prosumers.

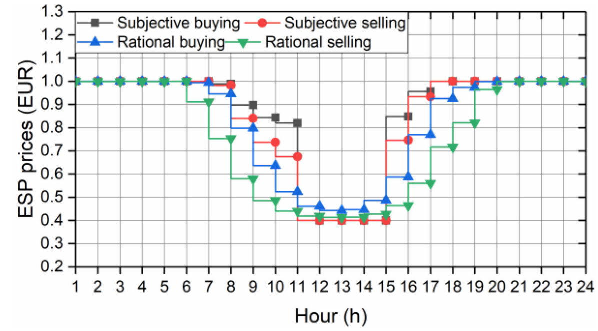

Fig. 11. The dynamic electricity prices set by ESP.

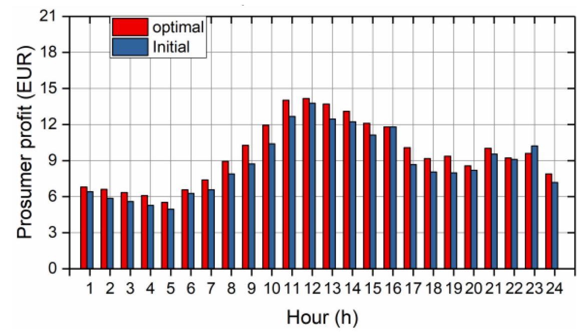

Fig. 12. The subjective utility of all prosumers.

strategies with high utility and chooses the first small interval for his social preferences and load constraints.

Fig. 10 shows the load strategies of all prosumers. The optimal load distribution is also increased in the time slots 9-16 for the same reason that was discussed before. The original load peak in time slots 19-23 is reduced, resulting in the flat load distribution and decreasing the influence of uncertainty PV production on the utility grid. Besides, the load also shows a DIA method can satisfy the load constraints in the energy scheduling process. From the perspective of social attributes, the optimal load strategies in a subjective scenario almost follow the previous load pattern than rational prosumers, indicating the influence of social preferences on their load strategies.

2) Results of Dynamic Prices: Fig. 11 shows the hourly dynamic prices set by ESP. The selling and buying prices are lower in time slots 8-17 under the higher PV production. It is because the prosumers are encouraged to buying energy (i.e., consuming energy) in these time slots to increase the local consumption of PV energy. The selling prices are always lower than or equal to the buying prices to keep the gaining of the utility grid. In time slots 12-15, the selling and buying prices are equal to the minimal value. The reason is that the higher PV production indicates the selling energy in these time slots far beyond buying energy.

Comparing with the rational scenario, the fluctuation of the curve in the subjective scenario is more volatile than in the rational scenario. The reason comes from two aspects: one is the discrete load strategies because of the social attributes, the other is the larger SP in small load intervals, resulting in lower load strategies and large deviation with original prices.

3) Results of Prosumers' Utility: Fig. 12 shows the utility of prosumers, and the utility of 24 hours increases from 211.01EUR to 229.34EUR through optimization. The utility of prosumers comes from saving buying energy from the utility grid and replaced by their own PV generation so that the profit is higher in time slots 8-17. The dynamic prices can help prosumers get more profit than utility gird prices, so that the increment of optimal profit is mainly in the time slots 8-17 because of the higher PV generation in these time slots. Besides, the optimal load strategies in some time slots (like 16, 22) are the same as that of the initial value. The reason is that the influence of social preference, resulting in the optimal utility is the same as the initial utility.

D. Analysis of the Parameters in the Solution Algorithm

To show the sensitivity of parameters in the solution algorithm, multiple comparison experiments with different parameters combinations are set, and the results list in Table II. The deviation refers to the optimal load deviation with the initial load in terms of the time-coupled constraint (11) and expresses it in a percentage form. Max deviation is the maximum value that comes from one prosumer, while mean deviation is the mean value of all prosumers. It is clear that the DIA can reduce the flexible load deviation of the time coupled constraints, and these combinations have similar results, among them, parameters [1, 2, 3], [1, 2, 3] have the best performance with the comprehensive consideration of deviation, profit, and iteration numbers. Therefore, the parameter combination is used in this application.

TABLE II
SENSITIVE ANALYZE OF PARAMETERS IN SOLUTION ALGORITHM

	b_1,b_2, b_3	d_1,d_2, d_3	profit	max deviation	mean deviation	Iterations
Without DIA	\	\	234.2	19.57%	7.51%	12
With DIA	[1,3,5]	[1,3,5]	229.0	7.49%	2.28%	18
	[2,3,5]	[2,3,5]	227.4	9.41%	3.05%	19
	[2,3,5]	[2,4]	227.4	9.41%	3.05%	19
	[2,4,6]	[2,4,6]	227.4	9.41%	3.05%	15
	[1,2,3]	[1,2,3]	229.3	5.87%	1.60%	15
	[2,3,4]	[2,3,4]	227.4	9.41%	3.05%	19
	[3,4,5]	[3,4,5]	227.1	9.41%	2.72%	15
	[4,5,6]	[4,5,6]	227.1	9.41%	2.72%	17
	[5,6,7]	[5,6,7]	227.1	9.41%	2.72%	16

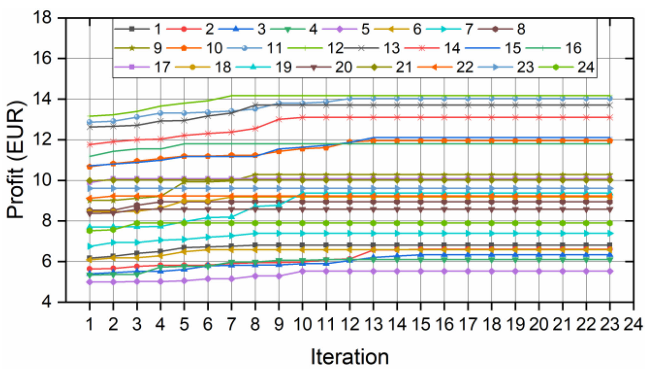

Fig. 13. The convergency curve of all prosumers in 24 time slots.

E. Analysis of Practical Feasibility and Computation Cost

The solving algorithm is implemented using a computer with Intel Core i7-8700K CPU 3.7GHz, 32G memory, TITAN XP GPU, 12G memory, and PYTHON3.6 with PyTorch was used as the testing environment for the algorithm. Interfaces can be designed in UEMS for prosumers to input their social attributes data and load metering data. The load metering data can directly download from the smart meters, and the social attributes data can be collected by asking prosumers when they use the UEMS or imported by the organizations that design the social survey. Then the STGCN can be implemented in each prosumer's UEMS in a distributed way, so that the privacy social attributes data is not required to share with others and the computation efficiency is improved. The stochastic game is also conducted in a distributed way with only a few KB's data exchanges between each prosumer and ESP. The data is the prosumer's load strategies that are set during the game process, not include the historical load metering data. Therefore, the privacy of prosumers can be guaranteed, and energy sharing can be conducted in the existing smart grid infrastructure. The information sharing is realized by the communication system based on the wireless channels in private 4G/5G network with Virtual Private Network (VPN), as well as equipment support of ESP and prosumers' UEMS.

The convergency of the solving algorithm is illustrated before, the iterative process including DIA stops until the subjective utility and load strategies of all prosumers are unchanged. In all the game stages, the maximum convergence

TABLE III
CALCULATION TIME IN LARGE-SCALE APPLICATION

Number of prosumers	11	22	44	88	176	352
Calculation time (s)	1.14	2.92	5.72	28.59	98.84	472.5

iteration number is 15 in time slot 2, and the convergency curve of all prosumers in 24 time slots is shown in Fig. 13. Although the convergency results may not the ideal optimal solution because of the heuristic mechanism, the results will approach the ideal one, which possesses practical value in the engineering application.

The proposed framework is distributedly conducted to support the scalability in the large-scale application, where the STGCN learning process is implemented in each prosumer's UEMS. The learning time of each prosumer is 78.20s, and the calculation time does not increase with the increment of prosumers numbers because the number of UEMS increases with the prosumers' number. For the stochastic game process, the calculation time increase with the prosumers number, and the computation complexity is $O(n)$. However, the calculation process for the stochastic game is fast, the maximum iteration time is 1.14s for the current 11 prosumers, which is far less than one hour. The calculation time for many large-scale scenarios with more prosumers is listed in Table III. Although the calculation time increase with the prosumers number, it still within the second level even the prosumers number reach 352. Therefore, the time consumption is also enough to conduct the energy scheduling in the large-scale application.

VI. CONCLUSION

In this paper, we propose a data-driven stochastic game model with social attributes for P2P energy sharing. We first show the SP through STGCN, and the learning process is based on the social survey data and load metering data The results indicate the probability distribution of social preferences is irregular and exhibit a similar characteristic of previous load distribution. The comparison of the STGCN network with other neural networks is also included. Then we show the optimal results of dynamic prices, load strategies, and utility through numerical verification, and study the influence of social attributes compared with the rational scenario. It is concluded that prosumers' load strategies approach to rationally optimum load to obtain greater utilities while maintaining the original distribution characteristics due to their social preference. The convergence and sensitivity analysis of the solution algorithm and the scalability analysis of the method is provided. Future research will deeply explore the influence of prosumers' social attributes on energy sharing by modeling and accurate learning methods.

REFERENCES

[1] Y. Parag and B. K. Sovacool, "Electricity market design for the prosumer era," *Nat. Energy*, vol. 1, no. 4, pp. 1–6, 2016, doi: 10.1038/nenergy.2016.32.

[2] N. Liu, X. Yu, C. Wang, C. Li, L. Ma, and J. Lei, "Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3569–3583, Sep. 2017.

[3] T. Morstyn, N. Farrell, S. J. Darby, and M. D. McCulloch, "Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants," *Nat. Energy*, vol. 3, no. 2, pp. 94–101, 2018.

[4] W. Tushar, C. Yuen, H. Mohsenian-Rad, T. Saha, H. V. Poor, and K. L. Wood, "Transforming energy networks via peer-to-peer energy trading: The potential of game-theoretic approaches," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 90–111, Jul. 2018.

[5] A. Ghosh, V. Aggarwal, and H. Wan, "Strategic prosumers: How to set the prices in a tiered market?" *IEEE Trans. Ind. Informat.*, vol. 15, no. 8, pp. 4469–4480, Aug. 2019.

[6] S. Cui, Y.-W. Wang, Y. Shi, and J.-W. Xiao, "A new and fair peer-to-peer energy sharing framework for energy buildings," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 3817–3826, Sep. 2020.

[7] W. Tushar *et al.*, "A coalition formation game framework for peer-to-peer energy trading," *Appl. Energy*, vol. 261, no. 1, Mar. 2020, Art. no. 114436.

[8] N. Liu, X. Yu, C. Wang, and J. Wang, "Energy sharing management for microgrids with PV prosumers: A Stackelberg game approach," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1088–1098, Jun. 2017.

[9] W. Tushar *et al.*, "Grid influenced peer-to-peer energy trading," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1407–1418, Mar. 2020.

[10] N. Liu, M. Cheng, X. Yu, J. Zhong, and J. Lei, "Energy-sharing provider for PV prosumer clusters: A hybrid approach using stochastic programming and Stackelberg game," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6740–6750, Aug. 2018.

[11] L. Chen, N. Liu, and J. Wang, "Peer-to-peer energy sharing in distribution networks with multiple sharing regions," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6760–6771, Nov. 2020.

[12] J. C. Do Prado and W. Qiao, "A stochastic decision-making model for an electricity retailer with intermittent renewable energy and short-term demand response," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2581–2592, May 2019.

[13] S. Cui, Y.-W. Wang, and J.-W. Xiao, "Peer-to-peer energy sharing among smart energy buildings by distributed transaction," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6491–6501, Nov. 2019.

[14] Z. Zhang, R. Li, and F. Li, "A novel peer-to-peer local electricity market for joint trading of energy and uncertainty," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1205–1215, Mar. 2020.

[15] T. Morstyn and M. D. McCulloch, "Multiclass energy management for peer-to-peer energy trading driven by prosumer preferences," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 4005–4014, Sep. 2019.

[16] P. T. Grogan and S. A. Meijer, "Gaming methods in engineering systems research," *Syst. Eng.*, vol. 20, no. 6, pp. 542–552, 2017.

[17] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Manage. sci.*, vol. 35, no. 8, pp. 982–1003, 1989.

[18] W. Tushar *et al.*, "A motivational game-theoretic approach for peer-to-peer energy trading in the smart grid," *Appl. Energy*, vol. 243, no. 1, pp. 10–20, Jun. 2019.

[19] S. R. Etesami, W. Saad, N. B. Mandayam, and H. V. Poor, "Stochastic games for the smart grid energy management with prospect prosumers," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2327–2342, Aug. 2018.

[20] Y. Xue and X. Yu, "Beyond smart grid—Cyber–physical–social system in energy future," *Proc. IEEE*, vol. 105, no. 12, pp. 2290–2292, Dec. 2017.

[21] W. Tushar, T. K. Saha, C. Yuen, P. Liddell, R. Bean, and H. V. Poor, "Peer-to-peer energy trading with sustainable user participation: A game theoretic approach," *IEEE Access*, vol. 6, pp. 62932–62943, 2018.

[22] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, Mar. 1979.

[23] G. El Rahi, S. R. Etesami, W. Saad, N. B. Mandayam, and H. V. Poor, "Managing price uncertainty in prosumer-centric energy trading: A prospect-theoretic Stackelberg game approach," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 702–713, Jan. 2019.

[24] L. Chen, N. Liu, C. Li, and J. Wang, "Peer-to-peer energy sharing with social attributes: A stochastic leader–follower game approach," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2545–2556, Apr. 2021.

[25] R. K. Jain, J. Qin, and R. Rajagopal, "Data-driven planning of distributed energy resources amidst socio-technical complexities," *Nat. Energy*, vol. 2, Jul. 2017, Art. no. 17112, doi: 10.1038/nenergy.2017.112.

[26] J. Wu *et al.*, "Multi-agent modeling and analysis of EV users' travel willingness based on an integrated causal/statistical/behavioral model," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 6, pp. 1255–1263, Nov. 2018.

[27] Z. Liu, Q. Wu, L. Christensen, A. Rautiainen, and Y. Xue, "Driving pattern analysis of nordic region based on national travel surveys for electric vehicle integration," *J. Mod. Power Syst. Clean Energy*, vol. 3, no. 2, pp. 180–189, Jun. 2015.

[28] L. S. Shapley, "Stochastic games," *Proc. Nat. Acad. Sci.*, vol. 39, no. 10, pp. 1095–1100, 1953.

[29] P. Suppes, "The role of subjective probability and utility in decision-making," in *Proc. 3rd Berkeley Symp. Math. Stat. Probab.*, vol. 5, 1956, pp. 61–73.

[30] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, 1998.

[31] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Basar, "Dependable demand response management in the smart grid: A Stackelberg game approach," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 120–132, Mar. 2013.

[32] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.

[33] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *Proc. AAAI Conf. Artif. Intell.*, Vol. 33, pp. 922–929, 2019.

[34] X. Feng, J. Guo, B. Qin, T. Liu, and Y. Liu, "Effective deep memory networks for distant supervised relation extraction," in *Proc. Int. Joint Conf. Artif. Intell.*, 2017, pp. 4002–4008.

[35] R. A. Horn and C. R. Johnson, "The hadamard product," in *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994, pp. 298–381.

[36] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent.*, Toulon, France, 2017. [Online]. Available: https://openreview.net/forum?id=SJU4ayYgl

[37] I. Goodfellow, Y. Bengio, and A. Courville, "Convolutional networks," in *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, pp. 378–415.

[38] Y. Chen, S. Mei, F. Zhou, S. H. Low, W. Wei, and F. Liu, "An energy sharing game with generalized demand bidding: Model and properties," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2055–2066, May 2020.

[39] R. Bellman, "The structure of dynamic programming processes," in *Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 2010, pp. 81–87.

[40] J. Zhan, W. Liu, C. Y. Chung, and J. Yang, "Switch opening and exchange method for stochastic distribution network reconfiguration," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2995–3007, Jul. 2020.

[41] Commission for Energy Regulation (CER), *CER Smart Metering Project-Electricity Customer Behaviour Trial, 2009-2010 [Dataset]*, 1st ed., Irish Soc. Sci. Data Archive, Dublin, Ireland, 2012. [Online]. Available: www.ucd.ie/issda/CER-electricity

[42] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 8, no. 9, pp. 1735–1780, 1997

[43] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. NIPS Workshop Deep Learn.*, Montréal, QC, Canada, 2014.

Liudong Chen (Graduate Student Member, IEEE) is currently pursuing the master's degree with the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China.

His research interests include energy management, smart grids, optimization theory, and machine learning.

Nian Liu (Member, IEEE) received the B.S. and M.S. degrees in electric engineering from Xiangtan University, Hunan, China, in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2009.

He is a Professor with the School of Electrical and Electronic Engineering, North China Electric Power University. He is a Member of the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources and a Member of Standardization Committee of Power Supply and Consumption in Power Industry of China. He was a Visiting Research Fellow with RMIT University, Melbourne, VIC, Australia, from 2015 to 2016. He has authored or coauthored more than 160 journal and conference publications and has been granted for more than ten patents of China. His major research interests include multienergy system integration, microgrids, cyber-physical energy system, and renewable energy integration. He is an Editor of IEEE Transactions on Smart Grid, IEEE Transactions on Sustainable Energy, IEEE Power Engineering Letters, and *Journal of Modern Power Systems and Clean Energy*.

Liangying Liu (Graduate Student Member, IEEE) is currently pursuing the master's degree with the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China.

Her research interests include distribution system optimization, smart grids, and machine learning.

Xinghuo Yu (Fellow, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1982 and 1984, respectively, and the Ph.D. degree in control theory and engineering from Southeast University, Nanjing, China, in 1988.

He is an Associate Deputy Vice-Chancellor, a Distinguished Professor, and a Vice-Chancellor's Professorial Fellow with the Royal Melbourne Institute of Technology University, Melbourne, VIC, Australia. His current research interests include variable structure and nonlinear control, complex and intelligent systems, and smart energy systems.

Dr. Yu was a recipient of number of awards and honors for his contributions, including the 2013 Dr.-Ing. Eugene Mittelmann Achievement Award of IEEE Industrial Electronics Society, the 2018 M. A. Sargent Medal from Engineers Australia, and the 2018 Australasian AI Distinguished Research Contribution Award from Australian Computer Society. He is the Junior Past President of IEEE Industrial Electronics Society for 2020 and 2021. He was an Associate Editor for the IEEE Transactions on Automatic Control, IEEE Transactions on Circuits and Systems—Part I: Regular Papers, IEEE Transactions on Industrial Electronics, and IEEE Transactions on Industrial Informatics.

Yusheng Xue (Life Member, IEEE) received the B.S. degree in electrical engineering from Shandong University, Jinan, China, in 1963, and the M.S. degree in electrical engineering from the State Grid Electric Power Research Institute (SGEPRI), Nanjing, China, in 1981, and the Ph.D. degree in electrical engineering from the University of Liege, Belgium, in 1987. He has been a Professor since 1993, a Member of the Chinese Academy of Engineering since 1995, the Chief Engineer of SGEPRI or NARI Group from 1994 to 2008, where he has also been the Honorary President since 2009. He is also an Adjunct Professor with several universities in China. His research interests include power system stability control, security, and economic operation. He is the Editor-in-Chief of *Journal of Automation of Electric Power System* (in Chinese, EI indexed) and that of *Journal of Modern Power Systems and Clean Energy* (in English, SCI IF 2.848 in 2019).