Received January 2, 2020, accepted January 16, 2020, date of publication January 20, 2020, date of current version January 28, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.2968000

Differentiated Incentive Strategy for Demand Response in Electric Market Considering the Difference in User Response Flexibility

DI LIU[1], YI SUN[1], BIN LI[1], XIE XIANGYING[2], AND LU YUDONG[3]
[1]School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China
[2]State Grid Electronic Commerce Company, Ltd., Beijing 100053, China
[3]Electric Power Research Institute, Zhejiang Electric Power Company, Ltd., Hangzhou 310014, China
Corresponding author: Bin Li (18500777272@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51777068, in part by the Fundamental Research Funds for the Central Universities under Grant 2019QN108, and in part by the Science and Technology Projects from the State Grid Corporation under Grant 52110418001C.

ABSTRACT In demand response programs, load service entity (LSE) can aggregate user as an independent entity to participate in the day ahead energy market, and complete the response target through incentive in the next day. In order to reduce the incentive cost of LSE, a differentiated incentive mechanism that consider the differences in user response flexibility is proposed in this paper. Then, a user response behavior model is established by using the long short-term memory (LSTM) network, with aim of accurately predicting users' response. Subsequently, an optimization strategy combining particle swarm (PSO) and LSTM is proposed, so that the response target can be accurately completed with low cost. Simulation experiments verified that the cost of LSE is close to the theoretical minimum, and can be reduced by 20% compared with the optimal result under unified incentive mechanism. Moreover, it also verified that the proposed strategy has high response accuracy and good stability.

INDEX TERMS Demand response (DR), differentiated incentive, user response behavior model, LSTM-PSO optimization strategy.

NOMENCLATURE

Abbreviation

LSE	load service entity
ISO	independent system operator
LSTM	long short-term memory
PSO	particle swarm optimization
URBM	user response behavior model
ACO	ant colony optimization
GA	genetic algorithm
SA	simulated annealing algorithm

Variables and parameters in modeling

t	time
T	temperature
L	user's load
I_0	basic incentive
γ	incentive factor
R_b^t	response reported by LSE in day ahead market
P_b^t	unit price reported by LSE in day ahead market
R_c^t	response target obtained from day ahead market
P_c^t	unit price obtained from day ahead market
R_i^t	response of user
C_i^t	user's response cost
α_i^t, β_i^t	parameters of the user's cost function
ε_i^t	noise of user's response
Y_i^t	revenue of the user
$R_{i-\max}^t$	maximum responsiveness of the user
Y_{LSE}^t	revenue of the LSE
$I_{unified}^{t*}$	incentive in unified incentive mechanism
$C_{differentiated}$	cost of LSE under differentiated incentive mechanism

The associate editor coordinating the review of this manuscript and approving it for publication was Pierluigi Siano.

$C_{unified}$	cost of LSE under unified incentive mechanism
μ_i, ν_i	coefficient of temperature influence degree
A^t	Weather in t timeslot
S	Season
D	Day type

Variables and parameters in LSTM network

f_t	forgetting gate
σ	sigmoid function
W	weight
b	bias
h_t	output
x_t	input
i_t	input gate
$\tilde{C}_t$	candidate value
C_t	CELL
o_t	output gate

Variables and parameters in PSO-LSTM algorithm

λ	a strictly increasing sequence
s	the number of iterations
$H(x)$	unbalance function
θ	multi-level assignment function
φ	level of the penalty function
P_i	optimal position of particle i
P_g	optimal position of particle swarm
V_i	speed of particle i
X_i	particle position after each iteration update

I. INTRODUCTION

With the massive penetration of new energy power generation, the flexibility of the generation is gradually decreasing. It is very important to aggregate demand side resources to maintain the flexibility of the power system [1]. According to the International Energy Agency (IEA), the demand response potential usually accounts for about 15% of the peak demand. By 2050, the EU's response potential may exceed 150GW [2]. Currently, distributed small users are gradually participating in the demand response. IEA proposed to integrate home users and participate in the power market in its report [2]. Similarly, the PJM power market also proposed to expand the scope of demand response to participate in the power market [3].

In PJM, distributed small users can participate in the demand response of the electricity market through the load service entity (LSE). LSE participates in the day-ahead market bidding as an entity, bids for the amount of load reduction and obtained the response target on the next day after the market clearing. Then the LSE need to organize users to complete the response target by means of direct load control (DLC) [4], dynamic price [5], contract [6], or incentives [7], etc. Since distributed small users are mainly home users, it is difficult to achieve DLC due to device limitations or privacy reasons [8], so in this paper, we focus on the incentive method.

A. LITERATURE REVIEW

Many literatures are dedicated to improve the economics of demand response. Reference [9] presented a fast distributed algorithm, to minimize the cost of aggregator while maximizing the users' comfort level. In [10], household load is divided into non-shiftable, shiftable appliances and electric vehicle, the total energy procuring cost is minimized by optimizing these loads. Reference [11] proposed a dynamic energy management framework based on highly-resolved personal energy consumption models, to re-shape the aggregate demand. In these studies, the load information of each appliance of the user needs to be obtained.

However, due to the limitation of the load information collection equipment, or the privacy issues, the user's load information of each appliance is often difficult to obtain, this will cause the uncertainty of user's response for LSE. User response functions need to be established when developing demand response strategies, many literatures have studied on this issue.

In [12], the author assumed that the response is linearly proportional to the incentive payment, and proposed a reward scheme for utilities. Considering the economic characteristics of users, [13] established a consumer's cost function using quadratic function, and then develop a joint online learning and pricing algorithm, to obtain the appropriate price for all the consumers in each time slot. Reference [14] used Stackelberg game theory to analyze the user's decision in demand response, and also used the quadratic function to establish the user's cost function. Reference [15] analyzed the impact of incentive-based demand response on microgrid operation, and [16] implemented incentive-based demand response by establishing stochastic energy cost function in microgrid.

At the same time, it is found that there are great differences between users through the analysis of user behavior [17]. Considering differences when developing incentives can improve the benefits in demand response, and there are already some literatures on this issue. Reference [18] divided the load into three categories and calculated the demand response cost respectively. Reference [19] designed three different reward schemes for different comfort level of users, but completed information is needed to calculate users' comfort level. In [20], a classification algorithm is employed to divide consumers into different categories, and a pricing model is formulated as a nonlinear programming problem, aiming to minimize the overall operation cost.

Based on the above analysis, demand response incentive strategies that consider user differences under incomplete information conditions need to be further studied:

1) When modeling the user's response behavior, most of the existing research established a static model, which is independent of time. However, due to the influence of the external environment (such as temperature, lighting, etc.), the user's response behavior may have different performances at different time, that is, the user's response elasticity is time-dependent [21]. Therefore, using a static model may result in larger error.

2) In the existing incentive mechanism, the difference of users has not been fully utilized. Most of the existing literatures clustered users first, and then designed incentive mechanisms for different types of users. But the differences among users of the same type have yet to be explored. At the same time, clustering may bring the issue of unfairness between different types of users. For example, A user who is clustered into category A thinks that the mechanism of category B is better, and he should be clustered into category B. In this case, the user may feel unfair.

B. NOVELTY AND CONTRIBUTION OF THIS PAPER

In many cases, LSE have specific demand response targets at a particular time [13]. For example, the LSE participates in day-ahead market to obtain specific response target, or obtains response target from demand response calls. Therefore, in order to accurately complete the target response, LSE needs to know the response behavior of users under different incentives. In this paper, a user response behavior model based on LSTM network is established. The model can reflect the relationship between user response behavior and external environmental factors, thus can describe user response behavior more accurately, compared with the existing static model. Then, a differentiated incentive mechanism and strategy is proposed to minimize incentive costs on the premise of achieving response target.

The main contributions of this paper are summarized as follows:

1) A differentiated incentive mechanism consists of basic incentive and incentive factor is proposed in this paper. In each time slot, the same basic incentive and incentive factor are announced to all users, which guarantees the fairness. Basic incentive ensures that users participating in demand response can get revenues, and incentive factor makes the user's revenue increase faster with the increase of response, thus increase users' motivation to participate in demand response. The incentive cost of the LSE under the differentiated incentive mechanism is always lower than the existing unified incentive mechanism is proved in this paper.

2) The user's response behavior under different external environments and incentives is analyzed, and a time-correlated user response behavior model (URBM) is established, which based on Long Short-Term Memory (LSTM) network. On the one hand, URBM can reflect the time correlation of user response behavior, thereby improving the accuracy of the model. On the other hand, it can flexibly select the input parameters (price, differentiated incentive, unified incentive, etc.) according to the requirements, without changing the structure of the model. In this paper, the simulation experiment showed that the model has a good performance under the differentiated incentive mechanism.

3) The Starkelberg game theory is used to analyze the decision-making behavior of users and LSE. Then a LSTM-PSO optimization strategy is established, by combining LSTM with particle swarm optimization (PSO). The incentive cost can be reduced to near the theoretical minimum, on the premise of achieving the response target. At the same time, the algorithm occupies less computing resources, and can support the implementation of large-scale engineering applications under the support of parallel computing.

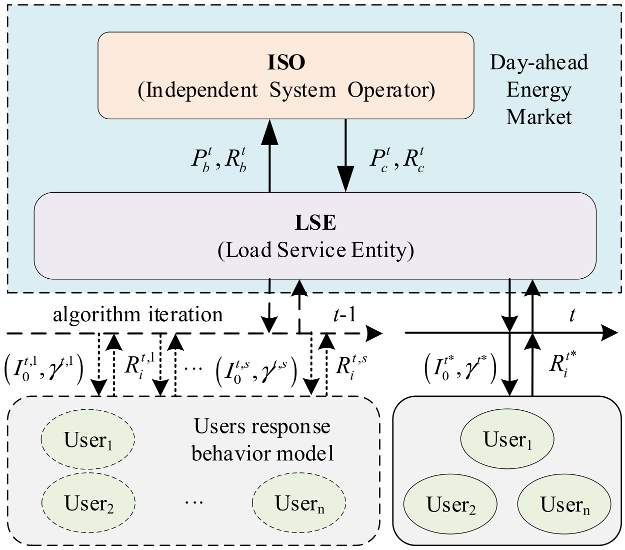

FIGURE 1. Implementation architecture of LSE aggregated distributed small users participation in demand response.

C. PAPER ORGANIZATION

The remainder of this paper is organized as follows. Section II described the system framework and proposed differentiated incentive mechanism, then the decision-making behavior of LSE and users is analyzed by using Stackelberg game theory, and the superiority of the proposed mechanism is proved mathematically. Section III established the user response behavior model. In Section IV, the LSTM-PSO optimization strategy is proposed. In Section V, case studies are performed and the results are discussed. Finally, conclusions and future work are presented in Section VI.

II. SYSTEM FRAMEWORK AND DIFFERENTIATED INCENTIVE MECHANISM

A. SYSTEM FRAMEWORK

In the electricity market, LSE can aggregate the demand side resources as an independent entity to participate in demand response bidding in the day-ahead energy market, such as day-ahead demand response program (DADRP) in PJM.

As shown in Fig.1, in day-ahead energy market, LSE reports the amount of load R_b^t that can be reduced in the t-slot on the next day and the corresponding unit price P_b^t. After the clearance of the day-ahead energy market, LSE obtains the amount of power R_c^t actually needed to be reduced on the next day. The price P_c^t is determined by the Locational Marginal Price (LMP) in the day-ahead energy market.

On the implementation day, LSE collects user and environment data in t-1 time slot, and obtain the optimal incentive strategy through algorithm iteration. At the beginning of the t period, LSE announces the incentive, and users make their response decision.

After the users execute the response decision, only the users' actual load can be measured. Therefore, in order to obtain the real response of users, it is necessary to calculate the baseline load of users. At present, many power markets have official baseline load calculation methods [22], many literatures also focus on the calculation method of user baseline load [17]. All the above methods can be used to calculate the user's baseline load, so as to calculate the user's actual response.

B. DIFFERENTIATED INCENTIVE MECHANISM DESCRIPTION

The differentiated incentive mechanism proposed in this paper consists of two parts: basic incentive I_0 and incentive factors γ. Under this mechanism, the revenue of user is as follows:

$$E_i^t = \left(I_0^t + \gamma^t R_i^t\right) \cdot R_i^t \tag{1}$$

where, I_0^t and γ^t are the basic incentive and incentive factors respectively in time slot t, R_i^t is the response of user i in time slot t.

Changes in the user's electrical behavior will result in a loss of comfort. According to the theory of demand elasticity [23], the loss of user comfort will increase at a faster rate as the response depth increases. Usually, the user's cost function can be approximated as follows [13]:

$$C_i^t = \frac{1}{2}\beta_i^t \left(R_i^t + \varepsilon_i^t\right)^2 + \alpha_i^t \left(R_i^t + \varepsilon_i^t\right) \tag{2}$$

where C_i^t is the user's response cost function, R_i^t is the response of the *i-th* user to the incentive in time slot t, α_i^t and β_i^t are two parameters of the user's cost function. ε_i^t is the noise caused by the randomness and uncertainty of the *i-th* user's response behavior at time slot t. In this paper, the random noise is assumed to obey the Gaussian distribution [13]. The user's revenue from the response in time slot t is:

$$\begin{aligned} Y_i^t &= \left(I_0^t + \gamma^t \left(R_i^t + \varepsilon_i^t\right)\right) \cdot \left(R_i^t + \varepsilon_i^t\right) - C_i^t \\ s.t.\ & 0 \leq R_i^t \leq R_{i-\max}^t \end{aligned} \tag{3}$$

where Y_i^t is the revenue of the user and $R_{i-\max}^t$ is the maximum responsiveness of the user in the time slot t.

C. STACKELBERG GAME ANALYSIS IN DIFFERENTIATED/UNIFIED INCENTIVE MECHANISM

A Stackelberg game is one type of extensive game that studies a situation with leader and followers, leader first make its strategy, and followers take actions subsequently. In proposed mechanism, LSE plays the leader role, and announces incentive first, and then users decide their response correspondingly.

The total revenue of LSE in proposed mechanism can be expressed as follows:

$$\begin{aligned} Y_{LSE}^t &= R_c^t \cdot P_c^t - \sum_{i=1}^{n} \left(I_0^t + \gamma^t R_i^t\right) R_i^t \\ s.t. \sum_{i=1}^{n} R_i^t &= R_c^t \end{aligned} \tag{4}$$

LSE's goal is to maximize revenue, so the optimal decision of LSE can be expressed as follows:

$$\begin{aligned} \left(I_0^{t*}, \gamma^{t*}\right) &= \arg\max_{I_0^t, \gamma^t} R_c^t \cdot P_c^t - \sum_{i=1}^{n} \left(I_0^t + \gamma^t R_i^{t*}\right) R_i^{t*} \\ s.t. \sum_{i=1}^{n} R_i^{t*} &= R_c^t \end{aligned} \tag{5}$$

where R_i^{t*} is the optimal response of the user i under the basic incentive I_0^t and the incentive factor γ^t.

Suppose user i offered by the basic incentive I_0^{t*} and incentive factor γ^{t*} in differentiated incentive mechanism, they will react based on their objective function. By substituting (2) into (3), the user's objective function can be obtained as follows:

$$\begin{aligned} \max_{R_i^t} Y_i^t &= \left(I_0^{t*} + \gamma^{t*}\left(R_i^t + \varepsilon_i^t\right)\right) \cdot \left(R_i^t + \varepsilon_i^t\right) \\ &\quad - \frac{1}{2}\beta_i^t \left(R_i^t + \varepsilon_i^t\right)^2 - \alpha_i^t \left(R_i^t + \varepsilon_i^t\right) \\ s.t. 0 &\leq R_i^t \leq R_{i-\max}^t \end{aligned} \tag{6}$$

The first derivative and second derivative of (6) can be expressed as follows:

$$\frac{dY_i^t}{dR_i^t} = I_0^{t*} + 2\gamma^{t*}\left(R_i^t + \varepsilon_i^t\right) - \beta_i^t\left(R_i^t + \varepsilon_i^t\right) - \alpha_i^t \tag{7a}$$

$$\frac{d^2 Y_i^t}{\left(dR_i^t\right)^2} = 2\gamma^{t*} - \beta_i^t \tag{7b}$$

From (7a) and (7b), the optimal decision R_i^{t*} of users in demand response can be obtained as follows:

$$R_i^{t*} = \begin{cases} \frac{\alpha_i^t - I_0^t}{} 2\gamma^t - \beta_i^t - \varepsilon_i^t 2\gamma^t - \beta_i^t < 0 \\ R_{i-\max}^t 2\gamma^t - \beta_i^t \geq 0 \end{cases} \tag{8}$$

In each demand response, LSE wants to complete the response target at the lowest cost to maximize its own revenue. As mentioned above, differentiated incentive mechanism can reduce the incentive cost of LSE, compared with the existing unified incentive mechanism, the mathematical proof is as follows.

Suppose user i offered by the incentive $I_{unified}^{t*}$ in unified incentive mechanism, they also will react based on their objective function. the user's objective function can be obtained as follows:

$$\max_{R_i^t} Y_i^t = I^t \cdot \left(R_i^t + \varepsilon_i^t\right) - \frac{1}{2}\beta_i^t \left(R_i^t + \varepsilon_i^t\right)^2 - \alpha_i^t \left(R_i^t + \varepsilon_i^t\right)$$
$$s.t. 0 \le R_i^t \le R_{i-\max}^t \tag{9}$$

From (9), the user's optimal decision can be obtained as:

$$R_{i,unified}^{t*} = \frac{I_{unified}^t - \alpha_i^t}{\beta_i^t} - \varepsilon_i^t \tag{10}$$

Suppose that the response target is the same under two kinds of mechanisms, that is $R_i^{t*}=R_{i,unified}^{t*}$, the cost of LSE under the two incentive mechanisms are as follows:

$$C_{unified} = \left(R_{i,unified}^{t*} + \varepsilon_i^t\right) \cdot I_{unified}^t \tag{11a}$$
$$C_{differentiated} = \left(R_i^{t*} + \varepsilon_i^t\right) \cdot \left(I_0^{t*} + \gamma^{t*}\left(R_i^{t*} + \varepsilon_i^t\right)\right) \tag{11b}$$

By substituting (8) and (10) into (11b) and (11a), the incentive cost of LSE can be expressed as:

$$C_{unified} = \beta_i^t \cdot \left(R_{i,unified}^{t*} + \varepsilon_i^t\right)^2 + \alpha_i^t \cdot \left(R_{i,unified}^{t*} + \varepsilon_i^t\right) \tag{12a}$$
$$C_{differentiated} = \beta_i^t \cdot \left(R_i^{t*} + \varepsilon_i^t\right)^2 + \alpha_i^t \cdot \left(R_i^{t*} + \varepsilon_i^t\right) - \gamma^t \cdot \left(R_i^{t*} + \varepsilon_i^t\right)^2 \tag{12b}$$

Since γ^t is a value that is always greater than zero, $C_{differentiated} < C_{unified}$ is always established, that is in the case of the same response target, the cost of the LSE in the differentiated incentive mechanism will always be low.

III. USER RESPONSE BEHAVIOR MODEL

As can be seen from (8), if LSE knows the parameters α_i^t and β_i^t, the response R_i^{t*} of the user i to the incentive $I_0^t + \gamma^t R_i^t$ at time t can be obtained. Then, LSE can obtain the optimal decision I_0^{t*} and γ^{t*} for each demand response by solving (5). But for the LSE, the parameters α_i^t and β_i^t of each user in different time slot t are unknown, the information that LSE can obtain is the incentives users received and corresponding response in the historical demand response.

In the existing research, some literatures used linear regression to estimate user parameters α_i^t and β_i^t [13], but the assumption of this method is that the user's parameters α_i^t and β_i^t are similar in different external environments, otherwise it will cause a large error. In fact, the user's parameters will be different in different external environments. Therefore, LSE needs to find a more accurate method, which can reflect the impact of environmental factors on user response behavior, to simulate the response of users under different incentives and make optimal incentive strategy accordingly.

The parameters of LSTM network can be trained by using the user's historical data (Incentives received by users, their response, and external environmental data, etc.), and then predict the user's response behavior in different external environments without needing to know the values of α_i^t and β_i^t.The prediction result of the LSTM network depends only on the historical data. Even if the values of α_i^t and β_i^t are changed, the LSTM network can fit well. Therefore, the LSTM network is very suitable for predicting the user's response behavior.

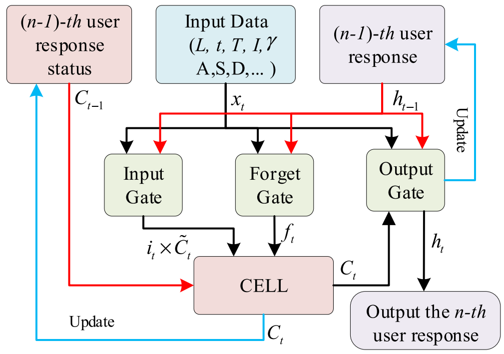

FIGURE 2. LSTM-based user response behavior model (URBM) architecture.

According to the actual operation of the PJM power market, the user's load curve has time-correlation characteristics and is affected by external environmental factors [24], correspondingly, the user's response behavior also has similar characteristics [25]. Therefore, the user's response behavior has historical similarity and is affected by the current environment at the same time. It is necessary to find an appropriate method to simulate and predict the user's response behavior. Long short-term memory (LSTM) network can transfer the user's historical characteristics on the time axis [26], and at the same time, it can output different predictions according to the current input [27], which is a good method to solve the above problems. LSTM-based user response behavior model (URBM) architecture is shown in Fig.2 [28].

As shown in Fig.2, a LSTM network consists of one CELL, one Input Gate, one Forget Gate and one Output Gate. The parameters of the LSTM network are trained based on historical data, and the predicted values are outputted based on the input data. According to the previous analysis, the user's response behavior has time-correlation characteristics and is affected by environmental factors. Therefore, the user's current load L, time t, temperature T, basic incentive I and incentive factor γ can be chosen as input data.

Among them, the current load L contains information of the maximum response potential. The value L in t slot can be accurately predicted by the data collected in t-1 slot [17]. the time t contains information of time-correlation characteristics, and the temperature T contains information of external influence factors, which can also be accurately predicted by using data in t-1 slot [29]. In addition, factors such as weather, season, day type (working day / non-working day), etc., will also affect the user's response behavior. All these inputs will affect the parameters α_i^t and β_i^t of the user in demand response. Therefore, the predicted response of the user in different incentives can be expressed as follows:

$$R_i^{t*}\left(I_0^t, \gamma^t\right) = URBM\left(L_i^t, t, T^t, I_0^t, \gamma^t, A^t, S, D, \ldots\right) \tag{13}$$

The Forgot Gate is used to determine what information is stored in the CELL. The model is expected to remember strongly related historical information and discard weakly related historical information. The update equation for the Forgot Gate is as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{14}$$

where σ is the sigmoid function, W_f is the weight of the Forgetting Gate, b_f is the bias of the forgetting gate, h_{t-1} is the output of the previous moment, and x_t is the input of the current moment.

The Input Gate determines what value will be used to update the information in the CELL. First, the input vector is processed by the activation function sigmoid, and the candidate value vector is generated by the tanh function. The input equation for the input gate is as follows:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{15}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{16}$$

where W_i, b_i are the input weight and bias, respectively, W_C, b_C are the weight and bias of the candidate value vector, respectively, and $\tilde{C}_t$ is the candidate value vector.

After the input information is processed by the Input Gate and the Forget Gate, the update rule of the stored data in the CELL can be obtained. The equation for updating the CELL is as follows:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{17}$$

After updating the state of the CELL, it is possible to determine what to output based on the content of the CELL and the current input, that is, the expected response of the user. The update equation for the Output Gate is as follows:

$$o_t = \sigma\left(W_o\,[h_{t-1}, x_t] + b_o\right) \tag{18}$$

$$h_t = o_t \times \tanh\left(C_t\right) \tag{19}$$

where, W_O, b_O are the input weight and bias respectively.

The LSTM-based user response behavior model is established based on TensorFlow. The flowchart is shown in Fig.3.

Since the training of the model takes a long time, LSE can train and store the user response behavior model in advance, and directly call the trained model when needed. Before each demand response is implemented, LSE optimizes to obtain the optimal incentive strategy by using the user response behavior model, and sends the incentive to the user to implement the demand response, then collects the actual response data of the user for the correction training of the model. After each n times of demand response, LSE retrains the user response behavior model based on the updated historical data to ensure the accuracy of the model.

IV. LSTM-PSO OPTIMIZATION STRATEGY

After the user response behavior model established, LSE needs to optimize the optimal incentive strategy according to the target response R_c^t of each demand response.

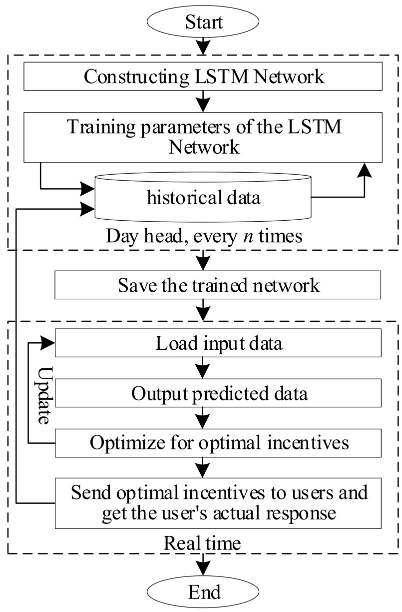

FIGURE 3. Training flow chart of user response behavior model based on LSTM.

TABLE 1. Comparison of intelligent optimization algorithms.

Algorithm	convergence speed	convergence ability
PSO	fast	May fall into local optimal solution
ACO	slow	Strong robustness, but greatly affected by the initial parameters
GA	general	The convergence is stable, but existing problems of Hanming cliff
SA	slow	The global optimal solution may not be obtained if the temperature drops too fast

If the LSE knows all the parameters, the model can be transformed into a nonlinear programming problem. But due to the lack of an accurate mathematical model of user response behavior, intelligent algorithm can be used to optimize the best incentive. Several intelligent algorithms are compared in this paper, including particle swarm optimization (PSO), ant colony optimization (ACO), genetic algorithm (GA) and simulated annealing algorithm (SA), as shown in Table 1.

From Table 1, it can be seen that the PSO algorithm has the fastest convergence speed, and the algorithm is simple and easy to implement for LSE. Although PSO may fall into local optimum, it can be avoided as much as possible by adjusting parameters. Therefore, considering the efficiency and implementation difficulty of the algorithm, PSO is chosen in this paper as the basic optimization algorithm. Then the LSTM-PSO optimization strategy is established by combining it with the user response behavior model.

From (4), the goal of LSE is to maximize the benefit in each demand response. At the same time, the incentive strategy of LSE should satisfy the constraint that the total user response is equal to the target response. Therefore, the fitness function of the LSTM-PSO optimization strategy can be composed of the benefit Y_{LSE}^{t} of the LSE and the penalty function considering the power balance constraint, as follows:

$$F(x) = Y_{LSE}^{t} - \lambda(s) H(x) \tag{20}$$

Among them, Y_{LSE}^{t} is the benefit of LSE, which can be calculated by using the prediction result of the user response behavior model in (13), as follows:

$$Y_{LSE}^{t} = R_c^t \cdot P_c^t - \sum_{i=1}^{n} \left(I_0^t + \gamma^t \cdot R_i^{t*}\left(I_0^t, \gamma^t\right)\right) \cdot R_i^{t*}\left(I_0^t, \gamma^t\right) \tag{21}$$

$\lambda(s)$ is a strictly increasing sequence. The form of $\lambda(s)$ in this paper is as follows:

$$\lambda(s) = s \cdot \sqrt{s} \tag{22}$$

where, s is the number of iterations of the LSTM-PSO optimization strategy.

$H(x)$ is obtained by the amount of imbalance between the user response and the target response, expressed as follows:

$$H(x) = \theta\left(\left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right|\right)\left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right|^{\varphi} \tag{23}$$

where θ is a multi-level assignment function, φ is the level of the penalty function, and the values of θ and φ can be determined by the following rules:

$$\begin{aligned} &when\ 0 < \left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right| \le 0.001\ \varphi = 1, \theta = 10 \\ &when\ 0.001 < \left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right| \le 0.1\ \varphi = 1, \theta = 20 \\ &when\ 0.1 < \left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right| < 1\ \varphi = 1, \theta = 100 \\ &when\ 1 \le \left|\sum_{i=1}^{n} R_i^{t*} - R_c^t\right| \varphi = 2, \theta = 300 \end{aligned} \tag{24}$$

After the fitness function is established, the optimal values of basic incentive I_0^{t*} and incentive factor γ^{t*} can be obtained by using the PSO algorithm. The current optimal position of particle i is updated as follows:

$$P_i^{k+1} = \begin{cases} P_i^k, F\left(x^{k+1}\right) \ge F\left(x^k\right) \\ X_i^{k+1}, F\left(x^{k+1}\right) < F\left(x^k\right) \end{cases} \tag{25}$$

The optimal position update equation for the entire particle swarm is as follows:

$$P_g^{k+1} \in \left\{P_1^k, P_2^k, \ldots, P_m^k\right\} = \max\left\{P_1^k, P_2^k, \ldots, P_m^k\right\} \tag{26}$$

The speed and position of each particle in each iteration is updated as follows:

$$\begin{aligned} V_i^{k+1} &= \omega \times V_i^k + c_1 \times r_1 \times \left(P_i - X_i^k\right) \\ &\quad + c_2 \times r_2 \times \left(P_g - X_i^k\right) \end{aligned} \tag{27}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \tag{28}$$

where i is the number of the particles in the particle group, k is the number of iterations, r_1 and r_2 are random numbers distributed between 0 and 1. and c_1 and c_2 are the acceleration factors, which determine the rate of change in the velocity of the particle in its optimal direction and global optimal direction, respectively.

In order to ensure that the particles do not escape the search space in the iteration process, it is necessary to set constraints on the particle's speed, as follows:

$$-V_{\max} \le v_i \le V_{\max} \tag{29}$$

The algorithm steps of LSTM-PSO optimization strategy are as follows:

Step 1: Initialize a particle swarm of size m, randomly set the initial position P_i^0 and the initial velocity v_i^0 of the particle within the search range;

Step 2: Transmitting the position information of each particle, that is, the values of basic incentive I_0^{t*} and incentive factor γ^{t*}, into the user response behavior model, then calculate the fitness function F_i based on the results predicted by the model, set $P_j^0\left(F_j^0 = \max\left\{F_1^0, F_2^0, \ldots, F_m^0\right\}\right)$ as the optimal position P_g^0 of the particle swarm;

Step 3: Update the position of each particle according to (25);

Step 4: Compare The fitness function value F_i^s of the optimal position of each particle and the value F_g^s of the optimal position fitness function of the particle swarm, if $F_i^s > F_g^s$, replace F_g^s with F_i^s;

Step 5: Update the speed and position of each particle according to (28), (29), and (30);

Step 6: If the maximum number of iterations is reached, terminate and exit, if not, continue to loop from step 3.

V. CASE STUDIES

A. SIMULATION SETTING

Suppose LSE has 100 user resources to participate in demand response and obtains the response target R_c^t =100kW in slot t of the next day from the electricity market with the price $P_c^t = 20¥/kW$.

According to (8), the response decision of the user i is determined by the values of α_i^t and β_i^t. In the existing research, there are many types of daily load characteristics of users [30], and they all show the periodic characteristics [31], that is, for the same type of users, they have similar load characteristics in the same time period every day [32].

In order to reflect the periodic characteristics of the user, one day is divided into four time periods, and the response

TABLE 2. User response behavior parameter

Combination	(S_1,D_1,A_1,t_1)	(S_1,D_1,A_1,t_2)	···	(S_4,D_2,A_4,t_4)
α_{i-0}^t	[0,5]	[0,5]	[0,5]	[0,5]
β_{i-0}^t	[10,15]	[10,15]	[10,15]	[10,15]

characteristics of user i are assumed to be the same every 6 hours. It should be pointed out that this paper only uses this division as an example to generate user data, the actual division of time in which users have similar response characteristics may be different, but the proposed algorithm can adapt.

At the same time, weather, season, day types and other factors will also affect the user's response flexibility. In the simulation experiments, the weather is divided into four categories: sunny, cloudy, rain and snow. The season includes spring, summer, autumn, and winter, and the day types are divided into working days and non-working days.

In order to simulate the complexity of user response in the real environment as much as possible, this paper assumed that users have different response flexibility under the combination of different influencing factors. There are four types of seasons, two day types, four types of weather, and four types of time segments, and there are 128 different combinations in total. In each combination, the basic values of α_{i-0}^t and β_{i-0}^t of user i are randomly generated, as shown in Table 2.

Among all loads of users, heating, ventilation and air conditioning (HVAC) load accounts for a large proportion [33], which are greatly affected by the external temperature [34]. And according to [35], temperature is one of the most important factors affecting the load behavior of users. Therefore, In order to reflect the influence of external temperature on user response behavior, it is necessary to add a temperature correction factor to correct the basic values of α_{i-0}^t and β_{i-0}^t.

The greater the deviation between the ambient temperature and the user's comfort temperature, the more the user's load demand [36]. Therefore, the adjusted parameters of users are as follows:

$$\begin{cases} \alpha_i^t = \left(1+\mu_i \cdot \dfrac{\left|T_t - T_i^0\right|}{T_i^0}\right) \cdot \alpha_{i-0}^t \\ \beta_i^t = \left(1+\nu_i \cdot \dfrac{\left|T_t - T_i^0\right|}{T_i^0}\right) \cdot \beta_{i-0}^t \end{cases} \tag{30}$$

Among them, T_i^0 is the optimum temperature of user i, T_t is the external temperature in time slot t, and μ_i and ν_i are the influence degree parameters of the basic values of α_{i-0}^t and β_{i-0}^t respectively, which are used to indicate the sensitivity of the user i to the external temperature. The random error ε_i^t of the user is generated according to the normal distribution with 0.02 as the mean square error and 0 as the expectation.

In this case, the values of all the parameters in (30) are shown in Table 3.

TABLE 3. The value of the temperature correction factor

Parameter	Value (°C)	Parameter	Value
T_i^0	26	μ_i	[0,1]
T_t	[-5,38]	ν_i	[0,1]

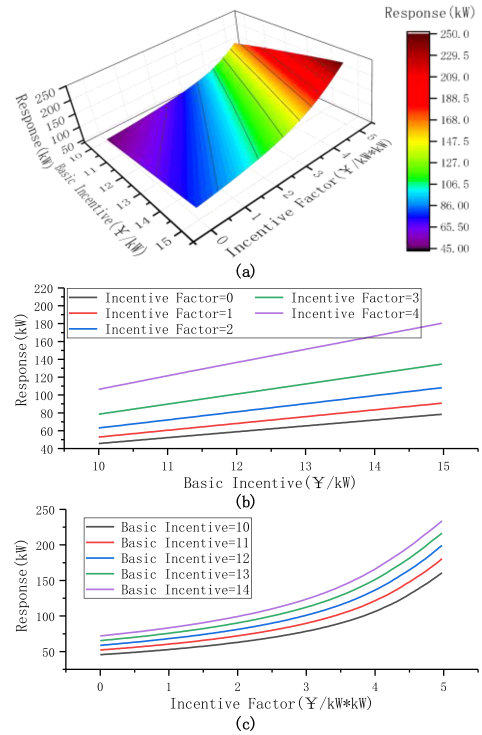

FIGURE 4. Relationship between incentive and response of user group.

The hidden layer unit of the user response behavior model is set to 20, and the batch method is used to train the model. The batch size is set to 60 and the learning rate is set to 0.0006.

When using the LSTM-PSO optimization strategy for iterative optimization, the number of particle swarms is set to 20, the number of iterations is set to 100, the values of c_1 and c_2 were set to 2, the search scope for the basic incentive and the incentive factor is set to [10,15] and [0,5], respectively, and the maximum speed is set to 15% of the variable range.

In order to verify the effect of the proposed differentiated incentive strategy, the strategy in [13] is used as the comparison algorithm, called unified incentive method in this paper. In fact, other literatures are similar, only have unified incentives in their strategy, such as [14].

B. RESULT ANALYSIS

1) ANALYSIS OF USER RESPONSE BEHAVIOR UNDER DIFFERENT INCENTIVES

Incentive directly determines the response of the user group, so the relationship between incentive and response is first analyzed, as shown in Fig.4.

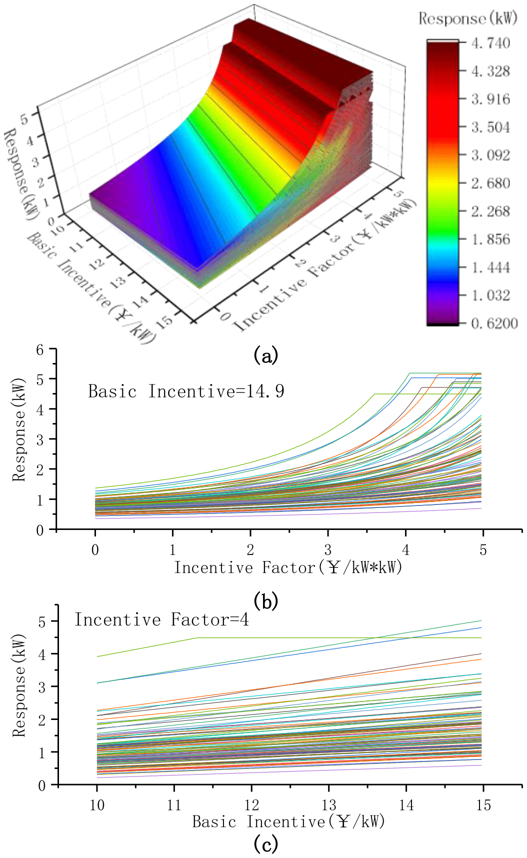

FIGURE 5. Relationship between incentive and response of single user.

The different combinations between the basic incentive I_0^t and the incentive factor γ^t can achieve the same response, as shown by contour lines in Fig.4(a), but the cost of LSE varies with different combinations. The response of user group increases with the increase of incentives. From Fig.4(b) it can be seen that the response of the user group rises linearly with basic incentive rises, and the slope of the rise is different under different incentive factors. At the same time, under the same basic incentive, the sensitivity of the user group's response increases with the increase of the incentive factor as shown in Fig.4(c). Therefore, it is necessary to find the optimal combination to reduce the incentive cost of LSE.

Fig.5 showed the relationship between incentive and response of each user. Where, each layer in Fig.5 (a) represents one user, and the lines of each color in Fig.5 (b) and (c) represent one user. As can be seen from Fig.5, the response characteristic of single user is similar to the user group, but the sensitivity of different users to the basic incentive I_0^t and the incentive factor γ^t is quite different. As the incentive rises, some users' responses are truncated and no longer rise, because the user has reached the maximum responsiveness. Therefore, LSE needs to develop incentive strategies within a reasonable range.

Furthermore, the relationship among unit cost, basic incentive and incentive factor is analyzed. As can be seen from

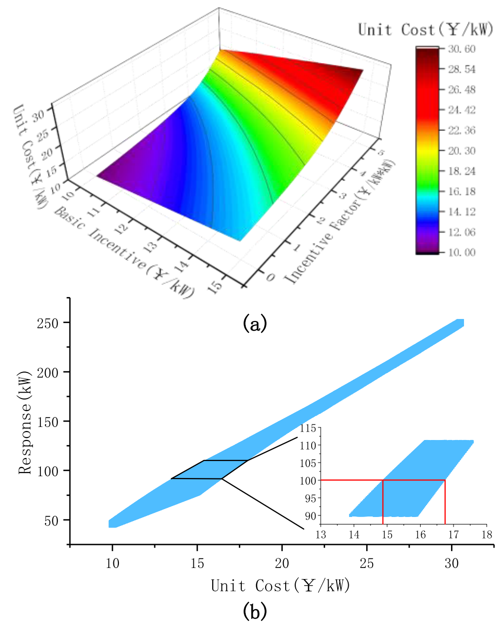

FIGURE 6. Relationship among unit cost, response, basic incentive and incentive factor.

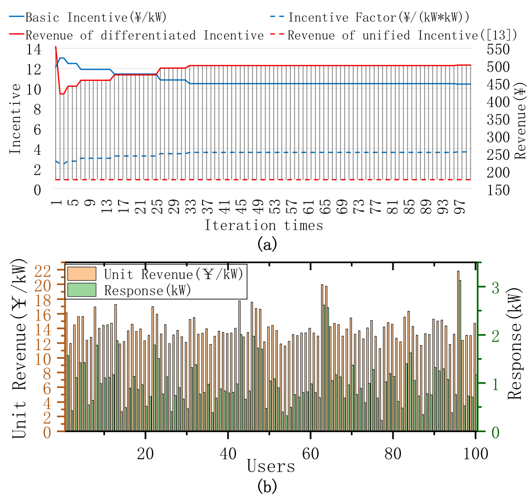

FIGURE 7. Incentives and revenues of LSE in each optimization iteration, and responses and benefits for each user.

Fig.6, the unit response cost also increases as the basic incentive I_0^t and the incentive factor γ^t increase. However, it can be seen from the trend of the contour line (the same unit cost) that the change in unit cost is different from the change in the response, that is, the unit cost of response is not strictly increased as the response increases.

Although unit costs are positively correlated with responses, for each identified response, there will be different unit costs due to different incentive combinations of the basic incentive I_0^t and the incentive factor γ^t. The purpose of the LSTM-PSO optimization strategy is to make the unit response cost as close as possible to the lower boundary on the premise of completing the target response.

2) ANALYSIS OF OPTIMIZATION RESULTS OF LSTM-PSO STRATEGY

For the case presented in this paper, the optimization iteration results are shown in Fig.7. Fig.7(a) showed the results of 100 iterations of the LSTM-PSO optimization strategy. Among them, the blue solid line and the blue dotted line represent the changes of the basic incentive and the incentive factor in the iteration process respectively. The red solid line and the red dotted line represent the revenue of LSE under the differentiated incentive mechanism and the unified incentive mechanism respectively.

It should be noted that both methods are optimized by PSO algorithm. The state space of unified incentive, which have only one variable, is small, so it converges to the optimal result in the first iteration.

The orange and blue curves represent the change in the basic incentive I_0^t and the incentive factor γ^t with iterations, respectively. After 100 iterations, the LSE's revenue reached 501.63 using the differentiated incentive method, and its incentive cost is 1498.37. In comparison, the LSE's revenue is 175.34 using unified incentive method, and its incentive cost is 1824.66. Compared with the unified incentive method, the proposed differentiated incentive method helps the LSE reduce the incentive cost by more than 20%. In addition, the unit incentive cost of differentiated incentive method after 100 iterations is 14.98, which is very close to the theoretical optimum about 14.90 (See Fig.6(b)).

Fig.7(b) showed the response of each user and its unit revenue, it can be seen that the larger the response is, the higher the unit revenue user obtained under the differentiated incentive mechanism proposed in this paper.

3) ANALYSIS OF THE ACCURACY AND STABILITY OF THE ALGORITHM

Then, the accuracy of the proposed method for the prediction of user group response is analyzed. Whether the total response of the user group can be close to the target response depends mainly on the performance of the user response behavior model and the penalty function.

As can be seen from Fig. 8(a), at the beginning of the iteration, since the error between the users' response and the target response is large, the loss of the fitness function value caused by the penalty function is large. As the iteration progresses, the loss of the fitness function value caused by the penalty function decreases rapidly and then stabilizes in a small range. It is shown that the penalty function established in this paper can well constrain the LSTM-PSO optimization strategy to search for the optimal solution in the domain that satisfies the balance constraint of the response.

As can be seen from fig.8(b), the predicted response of the user group has an error of about 4.2 kW from the target response of 100 kW at the beginning of the iteration. As the iteration proceeds, the predicted response of the user group can always fit near the target response and the error is within ±1.1kW. The actual response in the figure is assumed

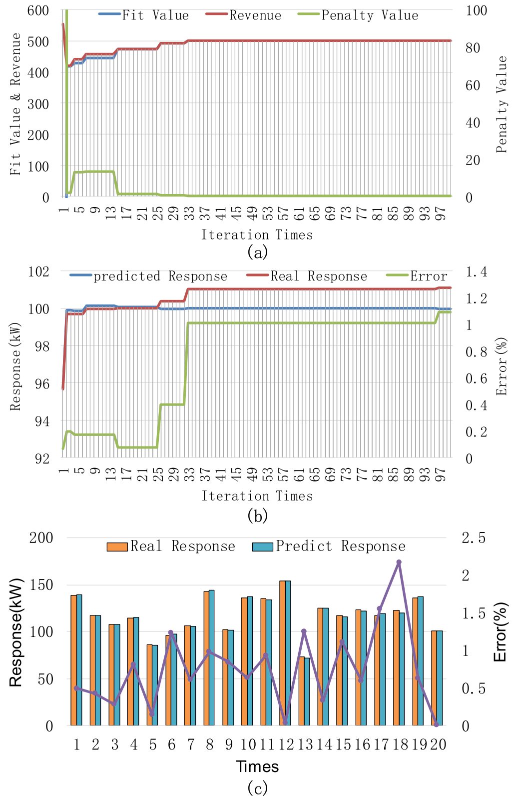

FIGURE 8. **The error between the user response and the response target, and the prediction result of 20 experiments of the user response behavior model.**

to know that all the parameters of the user model. It can be seen from the figure that the user response behavior model has a good performance. In 100 iterations, the user's response can be accurately predicted, and the maximum error is within 1.1%.

In order to verify the accuracy of the user response behavior model in depth, 20 random simulation experiments were performed in this paper, and the results are shown in Fig.8(c). It can be seen from the figure that the user response behavior model established in this paper has stable performance. Under different external environments and incentives, it can accurately predict the user's response behavior, and the maximum error is less than 2.5%.

It should be noted that due to the uncertainty of users' behavior, the prediction of user response behavior cannot be completely accurate. Therefore, there will be some deviations between the users' actual response and the target response.

In some power market, such as PJM, as long as the user's actual response is higher than the target response, it can be considered as completing the response task [3]. Therefore, LSE can set the response target in the optimization process to be higher than the actual response target (For example, the response target obtained from the electricity market is R_c^t, the response target can be set to $1.1 \times R_c^t$ during the optimization process), thereby ensuring that the response task can be completed. At present, the actual response in PJM market

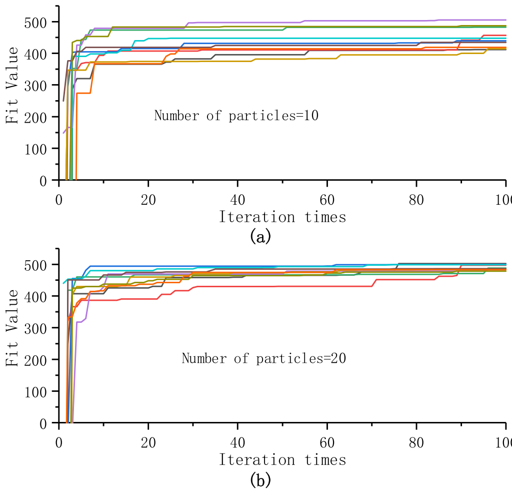

FIGURE 9. Convergence performance under different particle numbers.

is often greater than the target response [37]. But this will undoubtedly increase the incentive cost of LSE.

In addition, the LSE can also use other adjustable resources to make up for user response errors, such as load resources that can be directly controlled. But the cost of these adjustment methods will be higher than the incentive-based demand response. First, many load devices lack the function of remote control, and adding remote control module requires a lot of investment. Second, many users are unwilling to authorize LSE to control their own devices due to privacy or other issues, unless LSE pays a high price. So only by making the actual response of the user as close as possible to the target response, can it ensure that the LSE obtains a higher revenue.

Then, the paper verified the stability of convergence of the proposed optimization strategy. As mentioned above, the PSO algorithm is fast in calculation, but may fall into a local optimum. The number of particles is one of the important parameters of the algorithm's convergence. As long as the number of particles is large enough, the result can be prevented from falling into a local optimum as much as possible. In this paper, 10 particles and 20 particles are selected respectively for 10 times simulation experiments, the results are shown in Fig.9.

It can be seen from the figure that when the number of particles is 10, the optimization result is unstable after 100 iterations due to the small number of particles, but all the results are still better than the unified excitation method. When the number of particles is increased to 20, the results after 100 iterations can stably converge to the optimal value. Therefore, the optimization strategy proposed in this paper can help LSE get ideal optimization results stably as long as the number of particles is large enough.

4) EFFICIENCY ANALYSIS OF ALGORITHM

The operation efficiency of the algorithm has a great impact on the large-scale implementation of the algorithm. Therefore, the running time and memory usage of different modules of the proposed algorithm are recorded, and the running efficiency of the algorithm in large-scale implementation is analyzed.

TABLE 4. Algorithm module operation efficiency analysis.[1]

Algorithm module	Running time	Memory usage	Illustration
PSO	0.4s	5 MiB	Time for each iteration
URBM (Load graph)	5.1s	450 MiB	Only need to be loaded once per iteration
URBM (Load parameters and calculate results)	0.3s	5 MiB	Parallel computing can be used when the number of users is large

The running time and memory usage of each algorithm module are shown in Table 4. It can be seen that in the algorithm, the most computing resources are consumed by the prediction of user response behavior. In each iteration, it takes 5. 1s to load the graph, but because the structure of each user's graph is the same, in large-scale engineering implementation, graph needs to be loaded only once in each iteration, no matter how many users participate in the demand response.

At the same time, in the large-scale engineering implementation, parallel computing can be used to improve the efficiency of the algorithm, so it is necessary to make statistics on the memory usage of each algorithm module to analyze the feasibility of large-scale parallel computing. The parameter loading and result calculation of each user in the URBM model only use 5MiB memory, therefore, ordinary servers can support parallel computing for a large number of users.

The above analysis showed that the LSTM-PSO algorithm proposed in this paper can support large-scale engineering implementation of demand response. Since the calculation of each user's response behavior occupies less memory resources, parallel computing technology can be used in the project implementation, which can ensure that the algorithm's running time does not increase significantly, even with a large number of users.

VI. CONCLUSION

In order to make full use of users' differences to reduce the cost of LSE in demand response, this paper proposed a differentiated incentive mechanism and established a LSTM-PSO strategy. Compared with the existing unified incentive mechanism, revenue per unit response of user is no longer unified, but depend on users' response. The mathematical proof and simulation experiment in this paper showed that the incentive cost can be reduced by 20% compared with the optimal result under unified incentive mechanism. At the same time,

[1]Program environment is python, and URBM is established based on Tensorflow. Machine configuration is: core i7-9750H CPU, 16GB memory, and NVIDIA Geforce GTX 1660 Ti (6 GB) graphics card.

the accuracy and stability of the algorithm are also verified in the simulation experiments.

In future works, distributed energy storage and renewable energy resources on the user side can be considered when developing optimization strategy. At the same time, different LSEs can exchange energy, so appropriate mechanisms need to be proposed to achieve the Nash equilibrium between different LSEs.

REFERENCES

[1] A. Nikoobakht, J. Aghaei, M. Shafie-Khah, and J. P. S. Catalao, "Assessing increased flexibility of energy storage and demand response to accommodate a high penetration of renewable energy sources," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 659–669, Apr. 2019.

[2] IEA Publications, Paris, France. (2016). *Repowering Markets*. [Online]. Available: http://www.iea.org/publications/freepublications/publication/REPOWERINGMARKETS.PDF

[3] P JM Interconnection, Audubon, PA, USA. (Jun. 2017). *Demand Response Strategy*. [Online]. Available: http://www.pjm.com/-/media/library/reports-notices/demand-response/20170628-pjm-demand-response-strategy.ashx?la=en

[4] F. Elghitani and E. El-Saadany, "Smoothing net load demand variations using residential demand management," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 390–398, Jan. 2019.

[5] P. Jacquot, O. Beaude, S. Gaubert, and N. Oudjane, "Analysis and implementation of an hourly billing mechanism for demand response management," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4265–4278, Jul. 2019.

[6] D. G. Dobakhshari and V. Gupta, "A contract design approach for phantom demand response," *IEEE Trans. Autom. Control*, vol. 64, no. 5, pp. 1974–1988, May 2019.

[7] Y. Jia, Z. Mi, Y. Yu, Z. Song, L. Liu, and C. Sun, "Purchase bidding strategy for load agent with the incentive-based demand response," *IEEE Access*, vol. 7, pp. 58626–58637, 2019.

[8] A. Ghasemkhani, L. Yang, and J. Zhang, "Learning-based demand response for privacy-preserving users," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 4988–4998, Sep. 2019.

[9] S. Mhanna, A. C. Chapman, and G. Verbic, "A fast distributed algorithm for large-scale demand response aggregation," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2094–2107, Jul. 2016.

[10] S. Pal and R. Kumar, "Electric vehicle scheduling strategy in residential demand response programs with neighbor connection," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 980–988, Mar. 2018.

[11] M. Muratori and G. Rizzoni, "Residential demand response: Dynamic energy management and time-varying electricity pricing," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1108–1117, Mar. 2016.

[12] D. H. Vu, K. M. Muttaqi, A. P. Agalgaonkar, and A. Bouzerdoum, "Customer reward-based demand response program to improve demand elasticity and minimise financial risk during price spikes," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 15, pp. 3764–3771, Aug. 2018.

[13] P. Li, H. Wang, and B. Zhang, "A distributed online pricing strategy for demand response programs," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 350–360, Jan. 2019.

[14] M. Yu, S. H. Hong, Y. Ding, and X. Ye, "An incentive-based demand response (DR) model considering composited DR resources," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1488–1498, Feb. 2019.

[15] T. Khalili, A. Jafari, M. Abapour, and B. Mohammadi-Ivatloo, "Optimal battery technology selection and incentive-based demand response program utilization for reliability improvement of an insular microgrid," *Energy*, vol. 169, pp. 92–104, Feb. 2019.

[16] T. Khalili, S. Nojavan, and K. Zare, "Optimal performance of microgrid in the presence of demand response exchange: A stochastic multi-objective model," *Comput. Electr. Eng.*, vol. 74, pp. 429–450, Mar. 2019.

[17] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalao, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, Nov. 2018.

[18] A. Jafari, T. Khalili, H. G. Ganjehlou, and A. Bidram, "Optimal integration of renewable energy sources, diesel generators, and demand response program from pollution, financial, and reliability viewpoints: A multi-objective approach," *J. Cleaner Prod.*, vol. 247, Feb. 2020, Art. no. 119100, doi: 10.1016/j.jclepro.2019.119100.

[19] Z. Ni and A. Das, "A new incentive-based optimization scheme for residential community with financial trade-offs," *IEEE Access*, vol. 6, pp. 57802–57813, 2018.

[20] H. Yang, J. Zhang, J. Qiu, S. Zhang, M. Lai, and Z. Y. Dong, "A practical pricing approach to smart grid demand response based on load classification," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 179–190, Jan. 2018.

[21] Z. Guo, W. Li, A. Lau, T. Inga-Rojas, and K. Wang, "Trend based periodicity detection for load curve data," in *Proc. IEEE Power Energy Soc. General Meeting*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.

[22] P JM Interconnection, Audubon, PA, USA. (Nov. 2016). *Step-by-Step REST Examples for CBL Calculations*. [Online]. Available: https://www.pjm.com/~/media/etools/dr-hub/cbl-calculations-step-by-step.ashx

[23] A. Asadinejad, A. Rahimpour, K. Tomsovic, H. Qi, and C.-F. Chen, "Evaluation of residential customer elasticity for incentive based demand response programs," *Electric Power Syst. Res.*, vol. 158, pp. 26–36, May 2018.

[24] J. Warner-Freeman. (Jan. 2019). *Markets Report*. PJM Interconnection, Audubon, PA, USA. [Online]. Available: https://www.pjm.com/-/media/committees-groups/committees/mc/20190122-webinar/20190122-item-06a-markets-report.ashx?la=en

[25] S.-J. Kim and G. B. Giannakis, "An online convex optimization approach to real-time energy pricing for demand response," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2784–2793, Nov. 2017, doi: 10.1109/tsg.2016.2539948.

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

[27] A. Graves, "Long short-term memory," in *Supervised Sequence Labelling with Recurrent Neural Networks*, 4th ed. Berlin, Germany: Springer, 2012, pp. 37–45.

[28] C. Olah. *Understanding LSTM Networks*. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[29] X. Zhang, S.-C. Tan, and G. Li, "Development of an ambient air temperature prediction model," *Energy Buildings*, vol. 73, pp. 166–170, Apr. 2014.

[30] X.-J. Wang, L. Chen, and W.-Q. Tao, "Research on load classification based on user's typical daily load curve," in *Proc. IEEE Conf. Energy Internet Energy System Integr. (EI2)*, Beijing, China, Nov. 2017, pp. 1–4.

[31] L. Qiuyu, C. Qiuna, L. Sijie, Y. Yun, Y. Binjie, W. Yang, and Z. Xinsheng, "Short-term load forecasting based on load decomposition and numerical weather forecast," in *Proc. IEEE Conf. Energy Internet Energy System Integr. (EI2)*, Beijing, China, Nov. 2017, pp. 1–5.

[32] H. Zhou, W. Wang, Y. Wang, S. C. Wang, and H. Jiang, "Determination of similar days in load forecast based on grey incidence theory," in *Proc. Int. Power Eng. Conf. (IPEC)*, Singapore, Dec. 2007, pp. 163–166.

[33] N. Lu, "An evaluation of the HVAC load potential for providing load balancing service," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1263–1270, Sep. 2012.

[34] X. Chen, J. Wang, J. Xie, S. Xu, K. Yu, and L. Gan, "Demand response potential evaluation for residential air conditioning loads," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 19, pp. 4260–4268, Oct. 2018.

[35] M. Jie, G. Ciwei, C. Xiao, and Y. Yongxian, "A customer baseline load prediction and optimization method based on non-demand-response factors," in *Proc. China Int. Conf. Electr. Distrib. (CICED)*, Xi'an, China, Aug. 2016, pp. 1–5.

[36] M. Muratori, M. C. Roberts, R. Sioshansi, V. Marano, and G. Rizzoni, "A highly resolved modeling technique to simulate residential power demand," *Appl. Energy*, vol. 107, pp. 465–473, Jul. 2013.

[37] P JM Interconnection, Audubon, PA, USA. (Aug. 2018). *Load Management Performance Report 2017/2018*. [Online]. Available: https://www.pjm.com/-/media/markets-ops/dsr/2017-2018-dsr-activity-report.ashx?la=en

DI LIU was born in Henan, China, in 1990. He received the B.S. and M.S. degrees from North China Electric Power University (NCEPU), Beijing, China, in 2013 and 2017, respectively, where he is currently pursuing the Ph.D. degree in electric engineering. His research interests include machine learning applications in demand response and the optimization and data analytics of power and energy systems.

YI SUN was born in Liaoning, China, in 1972. He received the M.S. degree in communication and information system and the Ph.D. degree in electric information technology from North China Electric Power University, Beijing, China, in 2009 and 2014, respectively. He is currently a Professor of information and communication engineering with North China Electric Power University. His main research interests include smart power consumption, demand response, and power system communication technology.

BIN LI was born in Beijing, China, in 1983. He received the B.S. and Ph.D. degrees from the State Key Laboratory of Information Photonic and Optical Communications (IPOC), Beijing University of Posts and Telecommunications (BUPT), in 2005 and 2010, respectively. He was a joint training doctor with Yuan-Ze University (YZU), Taiwan, who focused on the antenna design, antenna measurement, electromagnetic scattering, and asymptotic high frequency techniques. After graduated from IPOC, BUPT, he worked in the Wireless and Optical Networking Research Group, Bell Labs Research China. He joined North China Electric Power University (NCEPU), in 2011, where he is currently an Associate Professor with the Communication Technology Research Center (CTRC), School of Electric and Electronic Engineering. His research interests include hybrid optical and wireless communication networks, next generation Internet, electric power communication, network routing, and signaling technology.

XIE XIANGYING was born in Beijing, China, in 1979. He is currently a Senior Engineer working with the State Grid Electronic Commerce Company, Ltd. His main research interests include energy Internet, and big data.

LU YUDONG was born in Hangzhou, Zhejiang, China, in 1977. He is currently a Senior Engineer working with the Electric Power Research Institute, Zhejiang Electric Power Company, Ltd. His main research interests include distributed power, big data, and machine learning.