

# LSTM-based Energy Management for Electric Vehicle Charging in Commercial-building Prosumers

Huayanran Zhou, *Student Member, IEEE*, Yihong Zhou, Junjie Hu, *Member, IEEE*, Guangya Yang, *Senior Member, IEEE*, Dongliang Xie, Yusheng Xue, and Lars Nordström, *Senior Member, IEEE*

**Abstract**—As typical prosumers, commercial buildings equipped with electric vehicle (EV) charging piles and solar photovoltaic panels require an effective energy management method. However, the conventional optimization-model-based building energy management system faces significant challenges regarding prediction and calculation in online execution. To address this issue, a long short-term memory (LSTM) recurrent neural network (RNN) based machine learning algorithm is proposed in this paper to schedule the charging and discharging of numerous EVs in commercial-building prosumers. Under the proposed system control structure, the LSTM algorithm can be separated into offline and online stages. At the offline stage, the LSTM is used to map states (inputs) to decisions (outputs) based on the network training. At the online stage, once the current state is input, the LSTM can quickly generate a solution without any additional prediction. A preliminary data processing rule and an additional output filtering procedure are designed to improve the decision performance of LSTM network. The simulation results demonstrate that the LSTM algorithm can generate near-optimal solutions in milliseconds and significantly reduce the prediction and calculation pressures compared with the conventional optimization algorithm.

**Index Terms**—Building energy management system (BEMS), electric vehicle (EV), long short-term memory (LSTM), recurrent neural network, machine learning, prosumer.

## NOMENCLATURE

### A. Indices

$f$  Mapping between input and output by LSTM

$\mathcal{H}, H$

Set of time steps of rolling horizon, and length of rolling horizon

$l$

Time-dependent hypothesis

$N, i$

Number of charging piles installed in a parking lot, and index of each charging pile

$T, t, \tau$

Number of time steps in one day, index of each time step, and length of each time step

$\wedge, \sim$

Indicators of estimated and filtered values

### B. Variables

$M_g^{im}, M_g^{ex}$

Complementary 0-1 binary variables to ensure that  $P_g$  has only one state at any time step

$M_{EVi}^c, M_{EVi}^d$

Complementary 0-1 binary variables to ensure that  $P_{EVi}$  has only one state at any time step

$P_{EVi}$

Charging or discharging power of EV connected to the  $i^{\text{th}}$  charging pile, denoted as the  $i^{\text{th}}$  EV

$P_{EVi}^c, P_{EVi}^d$

Charging (non-negative) and discharging (non-positive) power of the  $i^{\text{th}}$  EV

$P_g$

Exchanged power between commercial building and grid

$P_g^{im}, P_g^{ex}$

Imported (non-negative) and exported (non-positive)  $P_g$

$s_i$

State of charge (SOC) of the  $i^{\text{th}}$  EV

### C. Parameters

$\eta_i^c, \eta_i^d$

Efficiencies of charging and discharging processes

$\bar{b}, \underline{b}$

Upper and lower limits of  $P_g$

$C_i$

Battery capacity of the  $i^{\text{th}}$  EV

$c_{TOU}, c_F$

Time-of-use (TOU) tariff and feed-in tariff

$P_d$

Electrical demand of commercial building (non-negative)

$P_{PV}$

PV output (non-negative)

$\bar{P}_{EVi}^c, \bar{P}_{EVi}^d$

Rated values of charging and discharging power

$S_i^{dep}$

Desired SOC of the  $i^{\text{th}}$  EV at departure time

$S_i^{\max}, S_i^{\min}$

Upper and lower limits of the  $i^{\text{th}}$  EV SOC

$t_i^{arr}, t_i^{dep}$

Arrival time and departure time of the  $i^{\text{th}}$  EV

$t^{ps}, t^{pe}$

Start time and end time of peak period of TOU tariff

Manuscript received: July 21, 2020; accepted: March 10, 2021. Date of Cross-Check: March 10, 2021. Date of online publication: May 7, 2021.

This work was supported by the National Natural Science Foundation of China (No. 51877078), the State Key Laboratory of Smart Grid Protection and Operation Control Open Project (No. SGNR0000KJJS1907535), and the Beijing Nova Program (No. Z201100006820106).

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

H. Zhou, Y. Zhou, and J. Hu (corresponding author) 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: zhouhuayanran@163.com; zhouyihongc@126.com; junjiehu@ncepu.edu.cn).

G. Yang is with the Center for Electric Power and Energy, Technical University of Denmark, Lyngby, Denmark (e-mail: gyy@elektro.dtu.dk).

D. Xie and Y. Xue are with the State Grid Electric Power Research Institute, Nanjing, China (e-mail: xiedongliang@sgepri.sgcc.com.cn; xueyusheng@sgepri.sgcc.com.cn).

L. Nordström is with the Division of Electric Power and Energy Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden (e-mail: larsno@kth.se).

DOI: 10.35833/MPCE.2020.000501



## I. INTRODUCTION

TO mitigate environmental pollution, energy crisis, and climate change, the development of distributed photovoltaics (PVs), electric vehicles (EVs), and other distributed energy resources (DERs) has become the focus of societal attention in recent years [1], [2]. With the growing number of DERs at the load side, an increasing number of power users have changed from conventional consumers to prosumers (consumers with power generation capacity) [3]. As pointed out in [4], the active participation of prosumers at the end-user energy side is important for the global movement towards a society with sustainable renewable energy. Commercial buildings equipped with EV charging piles and solar PV panels are one of the typical prosumers in the grid. Under the incentives of electricity price, the energy use of these buildings can be optimized by the building energy management system (BEMS) [5]. Thus, the energy management of commercial buildings has significant potential for electricity cost saving, load leveling, and distributed generation consumption [6], [7].

The energy management undertaken by the BEMS is a sequential decision-making process that depends on the characteristics of the DERs. The existing research usually obtains optimal building energy management results by formulating an optimization model. This kind of method is model-driven and strictly follows the physical laws. Owing to the different charging and discharging efficiencies of EVs, the EV-related BEMS optimization model usually needs to introduce complementarity constraints, which guarantee that the charging and discharging of each EV are mutually exclusive [8]. Therefore, the BEMS optimization problem is non-convex and is hard to solve. Many methods have been proposed to solve the problem, including mixed-integer linear programming (MILP) [9], [10], intelligent algorithms [11], [12], iterative-based smoothing methods [13], and exact penalty methods [14]. However, to guarantee the coupling constraints of each EV's state of charge (SOC) over adjacent time steps, the aforementioned BEMS optimization models [9]–[14] need to be solved in a finite time horizon. As such, with the increasing number of dispatchable EVs and considered time steps, the BEMS faces a considerable amount of calculation pressure. In addition, the online solution of the optimization model depends on accurate prediction in the horizon, which is also difficult to obtain.

With the rapid growth of advanced computing infrastructures in recent years, machine learning methods appear to be suitable for overcoming the limitations of optimization models. Instead of building a physical model with complex constraints, these methods can acquire the tacit knowledge and formulate a mapping between the input and output by performing successive transformations of historical data. Subsequently, they can make decisions quickly in online execution, which greatly reduces the calculation pressure of BEMS in online execution. To achieve an efficient home-based demand response, a deep reinforcement learning (DRL) method based on a neural network and  $Q$ -learning algorithm is developed in [15]. By introducing a deep neural network to approximate the action-value function, such DRL-based

methods can avoid the drawback of dimensionality, which is the main limitation of conventional reinforcement learning (RL) method [16], [17]. The deep policy gradient algorithm, as part of the DRL method, is proposed in [18] to optimize the usage time and interruption frequency of multiple household electric devices in milliseconds. Reference [19] introduces a value-based DRL algorithm with a dueling deep  $Q$ -network structure to generate an interruption control signal of the aggregated interruptible load. Different from the aforementioned studies [15], [19], several deep learning (DL) algorithms are introduced in [20] to approximate the storage related home energy optimization strategy. The result indicates that many DL algorithms exhibit good performance in online execution. In [21], recurrent neural network (RNN)-based DL algorithms are applied to scheduling the energy profiles of household battery from historical optimization results. The performance of the algorithms has low sensitivity to the prediction accuracy. References [15]–[21] have provided valuable insights for our study.

However, it is still challenging to adapt the existing machine learning methods to the coordinated scheduling of EVs in commercial buildings. Firstly, considering the different characteristics of each EV such as charging demand, scheduling time limitation, and capacity, the training efficiency and generalization ability of the aforementioned machine learning methods have to be improved. Secondly, most of the aforementioned references ignore the influence of the temporal correlation in the BEMS problem, and still rely on prediction in online implementation. However, the temporal correlation needs to be considered, as the SOC of EV is time-coupling over adjacent time steps. Thirdly, as the machine learning model is a data-driven model, the internal logic and physical concepts of the model are not clear and the output results completely depend on the generalization ability of the model. There is no guarantee that the output of these machine learning methods is always within the physical limitations and meets the scheduling requirements.

In this study, a long short-term memory (LSTM) RNN-based machine learning algorithm is proposed to quickly solve the online BEMS scheduling problem. Compared with the aforementioned studies, this study provides the following contributions.

- 1) An LSTM-based system scheduling structure is constructed. In this structure, the training and execution of the LSTM network can be separated. The LSTM can be trained offline to acquire generalization ability and quickly generate each EV's scheduling result online in a fully decentralized manner.

- 2) An LSTM-based BEMS model is proposed. As one of the most advanced DL architectures for time-series prediction problems, the LSTM has powerful memorization capability. Thus, the LSTM can map the temporal correlation of the BEMS problem well based on historical data and there is no need for additional prediction.

- 3) A preliminary data processing rule and an additional output filtering procedure are designed to enhance the decision performance of machine learning. The optimal scheduling of the EVs can be learned better by the LSTM after the

preliminary data processing. Filtering is performed to remove unsafe, unsatisfactory, and fluctuating LSTM scheduling results.

The remainder of this paper is organized as follows. In Section II, the structure and task of a BEMS are introduced. Subsequently, the MILP-based optimization method, which is used to provide the training dataset for the proposed machine learning method, is reviewed in Section III. Section IV presents the proposed LSTM-based machine learning algorithm. Simulation results and discussions are presented in Section V. Finally, the conclusion is given in Section VI.

## II. PROBLEM FORMULATION

This section describes the system structure and main task of the BEMS.

As shown in Fig. 1, the commercial-building prosumer considered in this study is equipped with PV solar panels and a parking lot with EV charging piles, and there are no specific requirements for the capacities of PV and EV charging or the load level. Thus, the exchanged power between the commercial building and the grid at the point of common coupling (PCC) is determined by the electrical demand, PV output, and charging/discharging power of each EV, which can be given as (1).

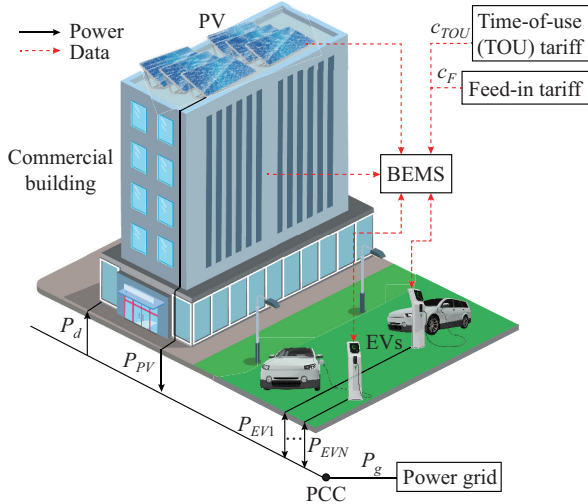


Fig. 1. BEMS structure and illustration of power and information flows in it.

$$P_g = P_d - P_{PV} + \sum_{i=1}^N P_{EVi} \quad (1)$$

Considering that the electricity price incentives include the TOU tariff and feed-in tariff, the commercial-building prosumers are known to have significant potential for electricity cost savings, as well as load leveling. In order to maximize the benefits of the commercial-building prosumers, the BEMS is installed to schedule  $P_g$ . It is usually expected that the PV output can be fully utilized to meet the electric power demand. Thus, the only schedulable resources considered here are the EVs. The main task of the BEMS is to coordinate the charging and discharging power of EVs  $P_{EV}$  with the PV output  $P_{PV}$  and electrical demand  $P_d$  under the incen-

tives of TOU tariff  $c_{TOU}$  and feed-in tariff  $c_F$ . Note that the EVs' parking time in commercial office buildings is relatively long enough to provide large scheduling flexibility, so we take commercial office buildings as the research object in this paper.

The BEMS can collect and store the information of  $P_d$ ,  $P_{PV}$ ,  $c_{TOU}$ ,  $c_F$ , EV behavior such as  $t_i^{arr}$  and  $t_i^{dep}$ , and EV characteristics such as  $\eta_i^c$ ,  $\eta_i^d$ ,  $C_i$ , and  $S_i^{dep}$ . In the conventional optimization-model-based BEMS, the power control signal is sent to each charging pile after a centralized optimization calculation. However, in the proposed LSTM-based machine learning method, the EV power control can be realized in a fully decentralized manner.

## III. MILP-BASED BEMS MODEL

In this section, we formulate a daily BEMS optimization model based on conventional MILP as (2)-(16) to provide the training dataset for the proposed machine learning method.

$$\min_{P_{EVi}^c, P_{EVi}^d} \sum_{t=1}^T (c_{TOU}(t)P_g^{im}(t) + c_F(t)P_g^{ex}(t)) \quad (2)$$

s.t.

$$P_g(t) = P_g^{im}(t) + P_g^{ex}(t) = P_d(t) - P_{PV}(t) + \sum_{i=1}^N P_{EVi}(t) \quad \forall t \in T \quad (3)$$

$$M_g^{ex}(t) \underline{b} \leq P_g^{ex}(t) \leq 0 \quad \forall t \in T \quad (4)$$

$$0 \leq P_g^{im}(t) \leq M_g^{im}(t) \bar{b} \quad \forall t \in T \quad (5)$$

$$M_g^{ex}(t) + M_g^{im}(t) = 1 \quad \forall t \in T \quad (6)$$

$$M_g^{ex}(t) \in \{0, 1\}, M_g^{im}(t) \in \{0, 1\} \quad \forall t \in T \quad (7)$$

$$P_{EVi}(t) = P_{EVi}^d(t) + P_{EVi}^c(t) \quad \forall i \in N, \forall t \in T \quad (8)$$

$$-M_{EVi}^d(t) \bar{P}_{EVi}^d \leq P_{EVi}^d(t) \leq 0 \quad \forall i \in N, \forall t \in T \quad (9)$$

$$0 \leq P_{EVi}^c(t) \leq M_{EVi}^c(t) \bar{P}_{EVi}^c \quad \forall i \in N, \forall t \in T \quad (10)$$

$$M_{EVi}^c(t) \in \{0, 1\}, M_{EVi}^d(t) \in \{0, 1\} \quad \forall i \in N, \forall t \in T \quad (11)$$

$$M_{EVi}^c(t) + M_{EVi}^d(t) = 1 \quad \forall i \in N, \forall t \in T \quad (12)$$

$$P_{EVi}^c = P_{EVi}^d = 0 \quad \forall i \in N, t \notin [t_i^{arr}, t_i^{dep}] \quad (13)$$

$$s_i(t) = s_i(t-1) + \left( \eta_i^c P_{EVi}^c(t) + \frac{P_{EVi}^d(t)}{\eta_i^d} \right) \frac{\tau}{C_i} \quad \forall i \in N, \forall t \in T \quad (14)$$

$$S_i^{\max} \leq s_i(t) \leq S_i^{\min} \quad \forall i \in N, \forall t \in T \quad (15)$$

$$s_i(t) \geq S_i^{dep} \quad \forall i \in N, t = t_i^{dep} \quad (16)$$

where the first term and second term of (2) are the payment to import electricity and the revenue of export electricity to the grid, respectively. The power balance at the PCC is ensured by (3), in which  $P_g$  is divided into  $P_g^{im}$  and  $P_g^{ex}$ . Equations (4)-(7) are set to ensure that  $P_g^{im}$  and  $P_g^{ex}$  are mutually exclusive in one time slot. In (8), the EV power is divided into charging and discharging power. Equations (9)-(12) ensure that  $P_{EVi}^c$  and  $P_{EVi}^d$  are mutually exclusive in one time slot. Equation (13) shows that the EVs do not have the scheduling capability in non-connected periods. Equation (14) formulates the relationship between the power and SOC



of the  $i^{\text{th}}$  EV. Equations (15) and (16) represent the SOC limit and charging demand of the  $i^{\text{th}}$  EV, respectively.

Solving this model relies on full-day complete information. This model can be used in offline scheduling. In this study, we use this model to generate the training dataset (historical data) for the proposed machine learning method (see Section IV) and reference results for the proposed machine learning method in practical application (see Section V). For convenience, we will call this model MILP with complete information later in this paper.

Moreover, a rolling-horizon MILP-based model optimization method (see Appendix A), which we call MILP with incomplete information in this paper, is applied to online scheduling problems. However, this model still faces a computational burden in execution because of the difficulty of solving integer variables and the high prediction dependency in each solving process. A comparison between the online scheduling results obtained by using the MILP with incomplete information model and those obtained by the proposed machine learning method is presented in Section V.

#### IV. LSTM RNN-BASED BEMS MODEL

In this section, an LSTM RNN-based machine learning method with significantly less computational burden and no prediction dependency on the power of PV, EV, or other electrical demand of the building is presented to address the BEMS problem. First, preliminaries of the method are introduced in Section IV-A. Subsequently, in Section IV-B, the control structure of LSTM-based system is constructed. Some details of the network training and execution procedure are provided in Section IV-C and Section IV-D, respectively. In Section IV-E, a filtering procedure is designed to enhance the decision performance of LSTM network. Finally, the overall procedure and structure of the proposed LSTM-based algorithm are summarized in Section IV-F.

##### A. Preliminaries of Method

As described in Section II, the main task of the BEMS is to schedule the charging and discharging of the numerous EVs parked in the commercial building. If we can obtain the SOC of all EVs at time step  $t+1$ , the charging or discharging power for each EV at time step  $t$  can be calculated as:

$$P_{EV_i}(t) = \begin{cases} \frac{s_i(t+1) - s_i(t)}{\tau\eta_i^c} C_i & s_i(t+1) - s_i(t) \geq 0 \\ \frac{s_i(t+1) - s_i(t)}{\tau/\eta_i^d} C_i & s_i(t+1) - s_i(t) < 0 \end{cases} \quad (17)$$

Thus, we can generate the EV power scheduling by predicting the SOC at the next time step.

Considering that the SOC has temporal correlation and is related to its previous state, its prediction can be regarded as a time-series prediction problem [22]. Generally, the aim of the time-series prediction problem is to predict the one-step-ahead or multi-step-ahead outputs based on several previous states and several previous inputs of the system. In this study, to achieve control of the EVs during the entire optimization horizon, we repeat the one-time-step-ahead predictions based on the previous prediction at each time step.

As an updated variant of the RNN, the LSTM is one of the most advanced DL architectures for time-series prediction problems [23], [24]. The LSTM network can model the temporal relationship of the time series using feedback connections to the internal nodes, i.e., the LSTM units in the LSTM layer. The unique LSTM unit structure, including forget gate, input gate, and output gate, greatly enhances the memorization capability of the LSTM network in the multi-step-ahead time-series prediction problem. In this way, the time coupling constraints of the SOC of each EV over adjacent time steps can be mapped well by LSTM, and the BEMS problem can be solved in a myopic way but still with temporal correlation. Therefore, the LSTM network can be applied to modelling the BEMS task and reduce the prediction difficulty.

Generally, an LSTM model is built based on the training of a dataset. The LSTM network consists of an input layer, an LSTM layer, a fully connected layer, and an output layer. Before the training process, the number of LSTM layers, number of nodes in each layer, and learning stopping criteria must be specified. The first step in training is presenting the data of previous states and previous inputs to the input layer. The weights of the network are then continuously adjusted according to the error between the network output and the output value in the training dataset until the algorithm converges.

After training, the LSTM can map inputs and outputs well. Once we input a new set of data to the input layer, the LSTM network can generate the corresponding output with temporal correlation.

##### B. LSTM-based System Structure

Next, we will apply the LSTM network to the energy management system.

To achieve a fast and optimal schedule for the BEMS, we use a system structure in which the training and execution of the LSTM network are separated.

The network training is done at the BEMS level. After training, the LSTM in BEMS has the generalization ability to map states (inputs) to optimal decisions (outputs), and the trained LSTM network parameters can be obtained. But the execution is done at the charging piles level. We assign each charging pile with an LSTM network which has the same network structure as the LSTM in BEMS. The LSTM in each charging pile can copy the trained network parameters from the BEMS so that all the EVs connected to one building could be scheduled in a fully decentralized manner.

Based on the system structure, the LSTM in the BEMS can be trained offline, and the LSTM in each charging pile can carry out the scheduling online. Thus, the LSTM network solution process can be divided into offline training and online execution. This structure-based energy management has no prediction dependency on the PV output, electrical demand, or other system data, and takes only milliseconds to generate power scheduling outputs, as validated in Section V. Note the offline training will be re-implemented after a certain period at the BEMS level to retrain a new model based on up-to-date data, and the new model will be

updated to the charging pile accordingly.

### C. Offline Training Details

This subsection introduces the details of the offline training in practical application. First, we need to decide on the input and output of the LSTM network. Second, the training dataset is specified. A preliminary data processing rule is then proposed to make the training more targeted. Finally, details of the hyperparameter setting and the final structure are presented.

#### 1) Input and Output

As described in Section IV-A, we have selected the LSTM as the tool for predicting the SOC of the EV connected to each charging pile. The output of each LSTM network is obviously the SOC at the next time step  $\hat{s}(t+1)$ .

The prediction of the SOC is a time-series prediction problem related to the previous state. Hence, the SOC at the current time step  $s(t)$  is an input. As already known, the SOC at the next time step is related to the PV output  $P_{pv}(t)$ , building electrical demand  $P_d(t)$ , and TOU tariff  $c_{TOU}(t)$  at the current time step. The feed-in tariff is constant in our case. Therefore, it is not an input to the network.

Thus, there are four nodes in the input layer corresponding to  $s(t)$ ,  $P_{pv}(t)$ ,  $P_d(t)$ , and  $c_{TOU}(t)$ , respectively, and there is one node in the output layer corresponding to  $\hat{s}(t+1)$ .

Owing to the powerful memorization capability of the LSTM (see Section IV-A), the relationship between the inputs and outputs in this paper can be described as (18), where the output state of time step  $t+1$  are related to the input and the state of the previous  $l$  time steps.

$$\begin{aligned} \hat{s}(t+1) = & f[s(t), s(t-1), \dots, s(t-l+1), c_{TOU}(t), c_{TOU}(t-1), \dots, \\ & c_{TOU}(t-l+1), P_{pv}(t), P_{pv}(t-1), \dots, P_{pv}(t-l+1), P_d(t), \\ & P_d(t-1), \dots, P_d(t-l+1)] \end{aligned} \quad (18)$$

#### 2) Training Dataset

The training dataset consists of historical data of inputs and outputs. The historical data of  $P_{pv}$ ,  $P_d$ ,  $c_{TOU}$ , EV behavior such as  $t_i^{arr}$  and  $t_i^{dep}$ , and EV characteristics such as  $\eta_i^c$ ,  $\eta_i^d$ ,  $C_b$ , and  $S_i^{dep}$  collected by the BEMS are used to build this dataset. Based on these historical data, the historical optimal output  $s_i$  in each charging pile can be generated by solving the MILP with complete information discussed in Section III. Subsequently, the historical data of  $P_{pv}$ ,  $P_d$ ,  $c_{TOU}$ , and the corresponding optimal  $s_i$  ( $i=1, 2, \dots, N$ ) are stored in the training dataset in chronological order.

Considering that the ranges of the aforementioned historical data are different and the LSTM is sensitive to the data scale, these data are scaled to  $[0, 1]$  by the min-max normalization method.

Since the LSTM network-based building energy management is a data-driven technology, to ensure the execution performance of the LSTM network, all possible situations of the system should be considered. It implies that a training dataset with massive data is required to realize good generalization ability in different situations.

#### 3) Preliminary Data Processing

Generally, the data sequences in most time-series prediction problems have the characteristic of temporal correlation.

To solve the time-series prediction problems, the neural network can be simply trained by sending the data sequences to the network without further processing.

However, in this paper, we should consider two aspects. First, the EV connected in a specific charging pile could be different on different days. Second, the EVs would not stay in the same building parking lot and keep charging the whole day. It means that the SOC data on one day can also be time-discrete. If we train the network with the training dataset directly, the LSTM would just try to fit the historical data without learning to make feasible predictions, as there is no temporal correlation in some data sequences. Therefore, it would be difficult to train the LSTM network, and even though this network reaches a good training target, it would only lead to an overfitting problem and have poor performance in application. The poor result of such training mode, i.e., LSTM with unprocessed data, is discussed in Section V.

To prevent the LSTM networks from overfitting the temporal correlations between the SOC of different EVs connected to the same charging pile, a preliminary data processing is necessary. First, the training dataset is divided into several segments according to the arrival and departure times of each independent EV connected to the  $i^{\text{th}}$  charging pile. Subsequently, we remove the time steps in which there is no EV in the  $i^{\text{th}}$  charging pile. Thus, the time steps in each segment are time-continuous. In the training process, we can use each segment data to train the LSTM network to realize a good fitting performance.

#### 4) Setting of Network Structure and Hyperparameters

To train this network, we select Adam [25] as the optimizer of the proposed network because of its computational efficiency and good performance. We set the learning rate of Adam to be 0.01 to decrease the convergence time, and the other parameters are set as the default values. After comparing the accuracy of the prediction and complexity of the network, we select two LSTM layers with 20 nodes, and the overall structure of the LSTM network proposed in this paper is depicted in Fig. 2. The input layer is used to scale the input values to the range of  $[0, 1]$  by the min-max normalization method. The fully-connected layer is used as the output layer, while the regression layer is used to calculate the mean square error (MSE), which could reflect the convergence of network in training.

### D. Online Execution Procedure

After the LSTM in each charging pile copies the trained network parameters from the BEMS, the LSTM in each charging pile can generate the power scheduling for each connected EV. When there is an EV connected to a charging pile, the charging pile detects the actual SOC of that EV and updates it in the input vector. Meanwhile, the LSTM in the charging pile can obtain data including  $P_{pv}(t)$ ,  $P_d(t)$ , and  $c_{TOU}(t)$  at the current time step from the BEMS. As given in (18), all the inputs have been successfully updated so that the network can predict the SOC at the next time step, after which the predicted SOC will be used to generate the power scheduling of the current time step based on (17).

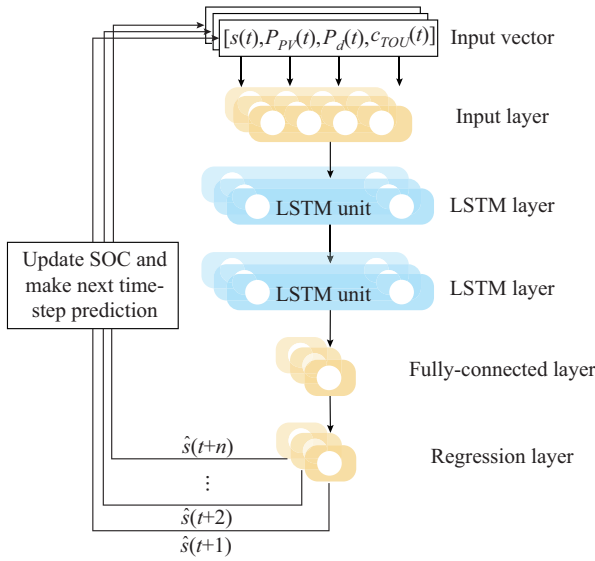


Fig. 2. Final structure of each LSTM neural network proposed in this paper.

### E. Filtering

After applying the LSTM, the SOC of the EVs at the next time step can be predicted and the charging or discharging power control can be realized. In fact, as a data-driven model, the LSTM-based BEMS generates the outputs, which completely depends on the generalization ability of the LSTM network. Owing to the lack of strict physical constraints and the limited generalization ability of LSTM, the outputs generated by the LSTM may have two issues. On the one hand, these outputs may violate some constraints of the EVs including the maximum SOC  $S_i^{\max}$ , the minimum SOC  $S_i^{\min}$ , the maximum charging and discharging rates  $\bar{P}_{EV}^c$  and  $\bar{P}_{EV}^d$ , and EV charging demand  $S_i^{dep}$ . On the other hand, there may be some fluctuations among the outputs due to the limited generalization ability of the LSTM. For example, when the trends of the SOC curve slightly fluctuate, from (18), we can observe that these slight fluctuations have direct impacts on power and thus result in errors. When the errors of many EVs occur simultaneously, the accumulated error in power can be significant even if the accuracy of the predicted SOC curve is high, which affects the electricity cost as well as the load leveling effect.

Considering that the interpretability of machine learning mechanism is poor, and the internal logic and physical concepts of machine learning models are not clear enough, it is hard to completely avoid these issues from the LSTM network. Therefore, it is necessary to design a filter outside the LSTM network to guarantee that the output of these machine learning methods is within the physical limitations and meet the scheduling requirements. Here, we propose a filtering process with two parts. In the first part, we illustrate our criteria to avoid violation of the operational constraints of the EV. In the second part, we aim to filter out the fluctuations due to the limited generalization ability of the LSTM networks.

Note that the filtering process is used only to avoid abnormal outputs of the LSTM networks, which do not always oc-

cur, and the result still mainly relies on the LSTM networks.

#### 1) The First Part of Filter

In the first part, we ensure that the outputs of the LSTM do not violate the constraints of the EVs. Here, five criteria are set up to determine whether these violations occur, and they are formulated as (19)-(23).

$$\hat{s}_i(t+1) \leq S_i^{\max} \quad \forall i \in N \quad (19)$$

$$\hat{s}_i(t+1) \geq S_i^{\min} \quad \forall i \in N \quad (20)$$

$$\frac{\hat{s}_i(t+1) - s_i(t)}{\tau} C_i \leq \bar{P}_{EV_i}^c \eta_i^c \quad \forall i \in N \quad (21)$$

$$\frac{\hat{s}_i(t+1) - s_i(t)}{\tau} C_i \geq -\bar{P}_{EV_i}^d / \eta_i^d \quad \forall i \in N \quad (22)$$

$$t + \frac{S_i^{dep} - s_i(t)}{\bar{P}_{EV_i}^c \eta_i^c / C_i} < t_i^{dep} \quad \forall i \in N \quad (23)$$

If (19) or (20) is not satisfied, the SOC of the EV will be set to the maximum SOC  $S_i^{\max}$  or the minimum SOC  $S_i^{\min}$ , respectively. If (21) or (22) is violated, the SOC at the next time step will be recalculated based on charging or discharging at the maximum rate. Formula (23) is used to assess whether the SOC of the EV can reach  $S_i^{dep}$  at  $t_i^{dep}$  if the EV is maintained charging at  $\bar{P}_{EV_i}^c$  from the current time step, and if (23) is not satisfied, this EV will be maintained charging at  $\bar{P}_{EV_i}^c$  until  $t_i^{dep}$ . Note that in the scheduling process, the uncertainties are mainly reflected in EV plug-in time and their charging demand. These uncertainties are strongly influenced by users' subjective initiative. The probabilistic model of historical data-based EV travel prediction is not sufficient to guarantee accuracy. Therefore, this paper assumes that interactions are available between aggregators and EV users to obtain the possible travel patterns of electric vehicles including  $S_i^{dep}$  and  $t_i^{dep}$ .

The first part forms the basic aspect to ensure the feasible power scheduling of the EV.

#### 2) The Second Part of Filter

In the second part, we filter out the fluctuations caused by the lack of strict physical constraints and the limited generalization ability of the LSTM networks. For specificity, we classify the fluctuations into two components, i.e., abnormal charging and discharging behaviors. Our criteria for determining the abnormal behavior are elucidated in the following paragraphs.

First, when the TOU tariff is not at the peak, the discharging is considered abnormal, because it would reduce the scheduling flexibility during the peak period, i.e., the discharging may cause extra charging to occur in the peak time to satisfy the EV charging demand. Such discharging process leads to a higher electricity cost.

Second, during the peak time of the TOU tariff, the charging is considered abnormal because it leads to a higher cost, unless the current SOC needs inevitable charging to occur at the peak time to satisfy the EV charging demand. This is because the LSTM should evenly arrange inevitable charging over the entire period. Otherwise, to satisfy the SOC demands, the full-speed charging of numerous EVs near the departure time leads to an additional peak of the load curve,

which affects the load leveling. The criterion of detecting abnormal charging is described as follows: charging at  $\bar{P}_{EVI}^c$  from the current SOC to  $S_i^{dep}$  starts from the peak period of the TOU tariff. This criterion is also given in (24).

$$t_i^{dep} - \frac{S_i^{dep} - S_i(t)}{\bar{P}_{EVI}^c / C_i} \in [t^{ps}, t^{pe}] \quad \forall i \in N \quad (24)$$

Moreover, during the peak time of the TOU tariff, the discharging is considered abnormal if it requires  $\hat{s}_i(t+1)$  to have additional inevitable charging during another peak time of the tariff. This criterion is similar to (24). However, the SOC in the equation should be the predicted value. It can be given as:

$$t_i^{dep} - \frac{S_i^{dep} - \hat{s}_i(t+1)}{\bar{P}_{EVI}^c / C_i} \in [t^{ps}, t^{pe}] \quad \forall i \in N \quad (25)$$

After detecting the abnormal charging or discharging behavior, the SOC at the next time step remains unchanged.

The filter designed is applied to  $\hat{s}(t+1)$  to obtain a filtered  $\tilde{s}(t+1)$ , which is a feasible and reliable prediction of the SOC.

The entire filtering procedure is given in Algorithm 1. As the filtering algorithm is composed of simple logical judgment statements and assignment statements, it can be quickly implemented.

---

**Algorithm 1:** entire filtering procedure

---

```

1: for  $i = 1, 2, \dots, N$  do
  First part
2: if  $\hat{s}_i(t+1) > S_i^{\max}$ 
3:    $\hat{s}_i(t+1) = S_i^{\max}$ 
4: else if  $\hat{s}_i(t+1) < S_i^{\min}$ 
5:    $\hat{s}_i(t+1) = S_i^{\min}$ 
6: if  $(\hat{s}_i(t+1) - s_i(t))C_i / \tau > \bar{P}_{EVI}^c \eta_i^c$ 
7:    $\hat{s}_i(t+1) = s_i(t) + \bar{P}_{EVI}^c \eta_i^c \tau / C_i$ 
8: else if  $(\hat{s}_i(t+1) - s_i(t))C_i / \tau < -\bar{P}_{EVI}^d \eta_i^d$ 
9:    $\hat{s}_i(t+1) = s_i(t) + \bar{P}_{EVI}^d \eta_i^d \tau / C_i$ 
10: if  $t + (S_i^{dep} - s_i(t)) / (\bar{P}_{EVI}^c \eta_i^c / C_i) > t_i^{dep}$ 
11:    $\hat{s}_i(t+1) = s_i(t) + \bar{P}_{EVI}^c \eta_i^c \tau / C_i$ 
  Second part
12: if  $t \notin [t^{ps}, t^{pe}]$ 
13:   if  $\hat{s}_i(t+1) - s_i(t) < 0$ 
14:      $\hat{s}_i(t+1) = s_i(t)$ 
15:   else if  $t \in [t^{ps}, t^{pe}]$ 
16:     if  $\hat{s}_i(t+1) - s_i(t) > 0$ 
17:       if  $t_i^{dep} - (S_i^{dep} - s_i(t)) / (\bar{P}_{EVI}^c \eta_i^c / C_i) \notin [t^{ps}, t^{pe}]$ 
18:          $\hat{s}_i(t+1) = s_i(t)$ 
19:     else if  $\hat{s}_i(t+1) - s_i(t) < 0$ 
20:       if  $[t_i^{dep} - (S_i^{dep} - \hat{s}_i(t+1)) / (\bar{P}_{EVI}^c \eta_i^c / C_i)] \in [t^{ps}, t^{pe}]$ 
21:          $\hat{s}_i(t+1) = s_i(t)$ 
22: end for

```

---

### F. Overall Implementation Procedure of LSTM-based BEMS Model

Based on the analysis, the overall implementation procedure of our proposed LSTM-based BEMS model is illustrated in Fig. 3.

At the offline stage, the BEMS collects the historical data of all the charging piles, PV generation, building electrical demand, and the TOU tariff. Based on those historical data,

the BEMS can obtain the historical optimal EV power scheduling results by solving the MILP with complete information. All the historical data and optimal solutions are stored in the LSTM training dataset. After the preliminary dataset processing, the LSTM in the BEMS can be trained and the trained LSTM network parameters can be obtained. Considering that the data distribution may not be time-invariant, we implement such an offline training stage after a certain period to retrain a new model based on up-to-date data.

At the online stage, the LSTM in each charging pile can copy the trained network parameters from the BEMS. Subsequently, the SOC at the next time step of the connected EV can be predicted by the LSTM in each charging pile based on the real-time input data as shown in (18). Then, the predicted SOC goes through a filtering process to enhance the decision performance of LSTM network. Afterwards, the charging and discharging control can be generated according to (17). The decentralized online scheduling procedure repeats to generate power control of the connected EV until the EV departs. The EV-related data are also stored and would be used at the offline training stage, as mentioned before.

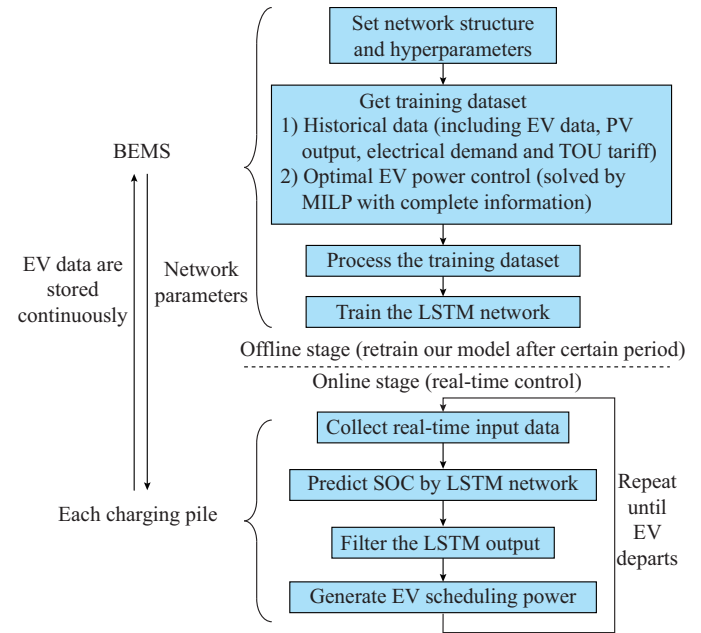


Fig. 3. Overall implementation procedure of LSTM-based BEMS model.

### V. CASE STUDY AND SIMULATION RESULTS

In this section, we provide a series of simulations to demonstrate the application of the proposed LSTM method in decision-making of EVs on commercial-building energy management. All simulations are run on a computer with an Intel® Core i7-7500U CPU @2.70 GHz-2.90 GHz and 8 GB of RAM. The LSTM algorithm is implemented on MATLAB R2019a platform, and the MILP model is solved using the YALMIP toolbox together with the intlinprog solver.

#### A. Data and Parameters

The case of a medium-scale office building with 500 kW PV onsite generation and 100 EV charging piles is studied



[26]. Considering that in the office building, the regular electrical demand and EV availability are relatively higher on weekdays, and the weekday energy management in turn has larger regulation potential, we therefore only focus on the BEMS problem on weekdays in our simulation. The BEMS problem at weekends and holidays can also be solved by training the corresponding LSTM network by the data of those day-types. We have selected the data on 250 weekdays in a year for our study. The electrical demand, PV output, TOU, and feed-in tariffs are shown in Fig. 4(a)-(c), respectively. The parking activities of the EVs on weekdays are from [7], which are displayed in Fig. 4(d). The first 220 days are selected as the training data of the LSTM network. The data of the 221<sup>st</sup> day to the 250<sup>th</sup> day are selected as the validation data to test the performance of our LSTM method.

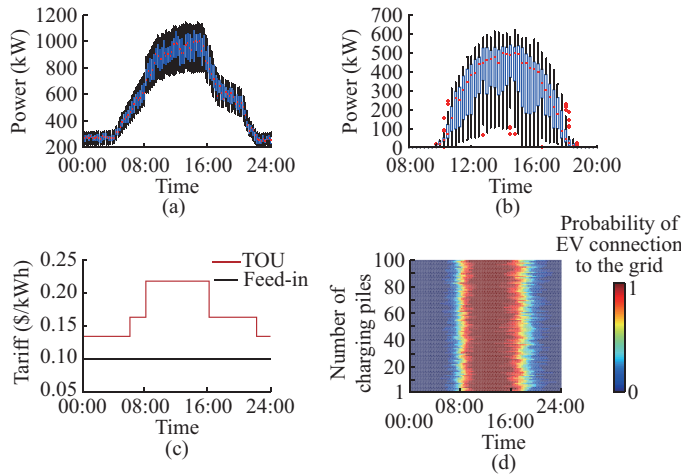


Fig. 4. Simulation data. (a) Electrical demand on weekdays of one year. (b) PV output on weekdays of one year. (c) TOU tariff and feed-in tariff. (d) EV availability on weekdays of one year.

Note that Fig. 4(a) and (b) is boxplots of the electrical demands and PV outputs data over 250 days. There are 96 boxes in the plots and each box represents the values at that time-step over 250 days. Each box is represented by a blue rectangle in the figure. The bottom and top edges of the box indicate the 25<sup>th</sup> and 75<sup>th</sup> percentiles, respectively. The red central mark “—” in that box indicates the median. The black dashed lines extended from the boxes are called whiskers, which extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the red “+” symbol.

The scheduling horizon for BEMS optimization in one day is from 00:00 to 24:00, and we take 1 time step as 15 min; thus, we can divide one day into 96 time steps. In the rolling-horizon optimization, considering the trade-off between the prediction accuracy and solving efficiency, one rolling-horizon is 4 time steps.

We assume that only one EV is connected to a charging pile in the office building during a weekday. Each EV can be charged at any rate from 0 to 3.3 kW with different currents. The average battery capacity of each EV is 30 kWh and the average initial SOC is 0.6. The charging and discharging efficiencies are both 95%. Each EV should be

charged to 0.85 SOC minimum before departure.

To verify the accuracy and efficiency of the proposed LSTM-based algorithm, we solve the case using five different methods: ① MILP with complete information (see Section III); ② MILP with incomplete information (see Appendix A); ③ the LSTM algorithm (see Section IV); ④ the LSTM without filter 2, which is based on the complete LSTM algorithm but adding only the first part of the filtering process; and ⑤ the LSTM with unprocessed data, which does not have preliminary data processing. In addition, the non-scheduling case is also simulated, which means that once the EV is connected to the charging pile, it is quickly charged until it is full and then disconnected.

### B. Learning Performance of LSTM Network

Firstly, the learning performance of the LSTM network is evaluated. The loss function of the LSTM network during iterative training is shown in Fig. 5. In this study, the loss function is the mean square error (MSE) between the network output and the output value in the training dataset. After 100 episodic iterations, the loss function has reached a very small value. After 500 episodic iterations, the loss function has been reduced to less than  $10^{-4}$ . Even though there are some slight fluctuations, the overall trend of the loss function proves the convergence of the algorithm. After convergence, the network has acquired the generalization ability and can be applied to online scheduling.

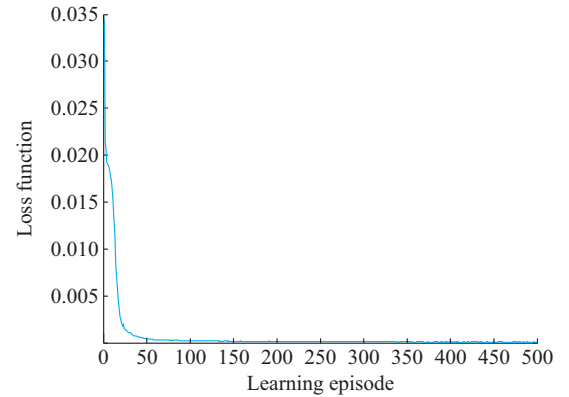


Fig. 5. Loss function of LSTM network during training process.

### C. Scheduling Performance of LSTM Algorithm in One Day

#### 1) SOC Scheduling Results

In this subsection, we present the energy management results of a day, e.g., the 250<sup>th</sup> day. First, the SOC scheduling results of a single EV are analyzed. Subsequently, the corresponding power scheduling results of the whole building are demonstrated and the solution time of the different methods is compared.

Since the optimal SOC obtained by the MILP with complete information is the learning target of the LSTM network, the SOC scheduling result is an important aspect of the BEMS performance. Figure 6 presents the SOC scheduling results of 8 randomly selected EVs obtained by using different methods.

The SOC of the LSTM is the closest to that of the MILP



with complete information. The second-closest one is that of the LSTM without filter 2, whereas that of the LSTM with unprocessed data is the most divergent.

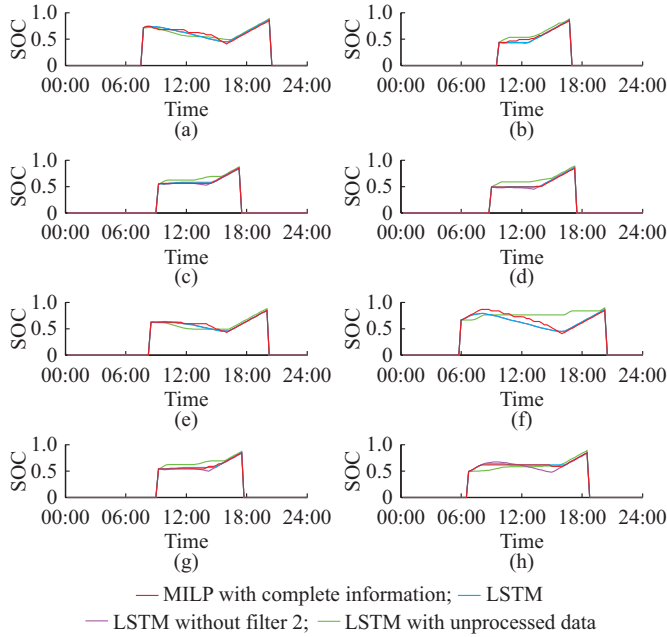


Fig. 6. SOC scheduling results of 8 randomly selected EVs using different methods. (a) EV 5. (b) EV 16. (c) EV 22. (d) EV 6. (e) EV 34. (f) EV 56. (g) EV 80. (h) EV 98.

In fact, we can observe that the SOC of the LSTM without filter 2 is almost coincident with that of the LSTM in Fig. 6(a), (e), and (f). This phenomenon indicates that the abnormal outputs, which we want to avoid by adding the filter 2, do not always occur, and the final SOC result still mainly relies on the LSTM networks. However, as can be seen in Fig. 6(b)-(d), (g), and (h), the SOC result of the LSTM without filter 2 is more fluctuant than that of the LSTM, particularly in Fig. 6(h). Such SOC fluctuation is not good for saving electricity cost and load leveling of the commercial building, as indicated in (17). Thus, this fluctuation must be removed, and this can be accomplished by the second part of the filter.

As for the solution of the LSTM with unprocessed data, the large errors mainly result from the lack of preliminary data processing before training. The unprocessed training dataset contains many time steps when the SOC is 0, i.e., there is no EV plugged in the charging pile. The SOC in those redundant time steps are learned by the LSTM network. The preliminary data processing allows the LSTM to focus more on the SOC information when an EV is available (connected to the charging pile), resulting in a better solution.

Given that the optimal solution obtained by the MILP with complete information cannot be applied to the BEMS online scheduling, we can use the near-optimal solution obtained by the LSTM network in online scheduling, because it is close to that of the MILP with complete information.

### 2) Power Scheduling Results

To illustrate the effect of BEMS load leveling, the corresponding power scheduling results of the commercial build-

ing using different methods in comparison with the base electrical demand are plotted in Fig. 7. The peak shaving effect of scheduling can be clearly observed in Fig. 7 regardless of the methods. However, the degree is different.

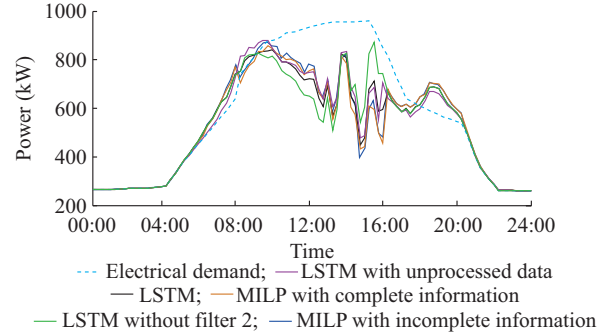


Fig. 7. Power scheduling results of commercial building using different methods.

Similar to Fig. 6, the LSTM achieves the closest power scheduling performance to that of the MILP with complete information. Owing to the lack of preliminary data processing, the power scheduling result using the LSTM with unprocessed data deviates more from the optimal solution obtained by the MILP with complete information than that of the LSTM. Moreover, because of the limited generalization ability of the LSTM network, the power scheduling result of the LSTM without filter 2 shows strong volatility, especially during the peak tariff period between 08:00 and 16:00. This is because the power depends on the difference between the adjacent SOC, and thus, the fluctuation of the SOC of different EVs leads to a disordered power scheduling result in a time step. Therefore, from the perspective of power leveling, the addition of filter 2 is very necessary.

The power scheduling result of the MILP with incomplete information also deviates slightly from the optimal solution of the MILP with complete information. This deviation is mainly because of the prediction inaccuracy in the rolling-horizon optimization. From the results of the power scheduling, both the LSTM and MILP with incomplete information methods can obtain near-optimal solutions.

As the MILP with incomplete information is a commonly used online scheduling method at present, a more detailed comparison between the MILP with incomplete information and LSTM is presented in the following subsection in terms of solution time and electricity cost.

### 3) Solution Time

Based on several simulations, we collect statistics on the online solution time of both the MILP with incomplete information and LSTM. For simplicity, the time of prediction in the MILP with incomplete information is not considered, and only the calculation time of optimization is counted.

Our simulation results show that the solution time of one LSTM network can reach a millisecond level, i.e., 0.002 s on average, including the LSTM output time and the filtering process. Considering the LSTM-based system structure, the LSTM in each charging pile will output results simultaneously; thus, the number of EVs has very limited impact on the output time.

The average solution time of the MILP with incomplete information is 1.81 s for one calculation. However, the solution time of the MILP increases exponentially with the increase of the EV numbers [8].

The difference in solution time between the two methods is mainly because the MILP problem involves numerous matrix inversions, which are time-consuming. By comparison, the LSTM does not require this procedure. Furthermore, the MILP with incomplete information must be combined with a prediction to obtain optimal scheduling results in practical application. Therefore, considering the additional prediction time, the solution time of the MILP is longer. The proposed LSTM does not bring huge computation or prediction burden to the BEMS, which makes the LSTM method more suitable for online scheduling.

#### D. Scheduling Performance of LSTM Algorithm in 30 Consecutive Weekdays

To further demonstrate the long-term performance and stability of the LSTM method, we present the BEMS results of the 221<sup>st</sup> day to the 250<sup>th</sup> day. First, an electricity cost comparison in a month is illustrated, and the power scheduling results on 30 consecutive weekdays are then compared.

##### 1) Electricity Cost

We compare the electricity cost of the commercial building on 30 consecutive weekdays with different methods and a non-scheduling case, as depicted in Fig. 8.

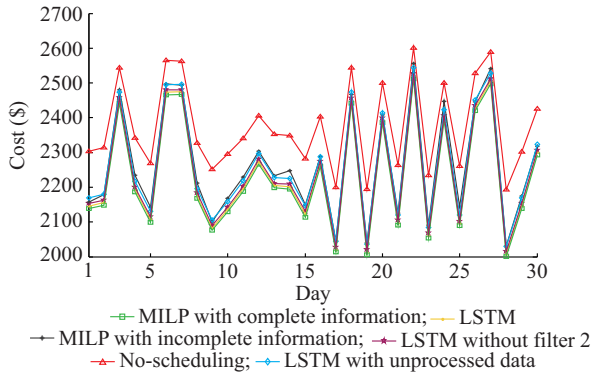


Fig. 8. Comparison of different methods in terms of electricity cost on 30 consecutive weekdays.

When there is no scheduling, the electricity cost is the highest, reaching \$71247.87 in total. The scheduled cost given by the MILP with complete information is the lowest, which is \$66875.47 in total, because the scheduling is a complete information-based full-horizon optimization. We can observe that the LSTM is the second-lowest cost with \$67095.45 in total, which is 0.33% higher than the lowest one.

According to the analysis of the power scheduling results in Fig. 7, although the fluctuation of the LSTM network has a negative impact on the power leveling, its effect on electricity costs is not evident. This is mainly because the majority of random power fluctuations during the peak tariff period can cancel out each other over the course of a day. As shown in Fig. 7, the power consumption of the LSTM with

unprocessed data is high in most time steps, resulting in a higher electricity cost as shown in Fig. 8. Therefore, the preliminary data preprocessing before LSTM training is important to reduce the electricity cost when using the LSTM method.

We can also observe from Fig. 8 that the popular rolling-horizon optimization method, namely the MILP with incomplete information, achieves a higher electricity cost than the LSTM method. In fact, the solution using the MILP with incomplete information depends largely on the prediction of the future state of the system such as the PV output prediction. The accuracy of the prediction can barely reach 100% (90% in this study), leading to a higher electricity cost. In contrast, the LSTM solution does not depend on the prediction of any system states. The training process enables the LSTM model to learn the mapping relationship between the inputs and the optimal output from the historical data obtained by the MILP with complete information. It can be concluded that in this case, the proposed LSTM method exhibits a better performance in terms of cost savings than the conventional MILP.

##### 2) EV Power Scheduling Results

Based on Fig. 9, we can infer that the summed power scheduling results of all the EVs (100 in total) have good stability by using the LSTM.

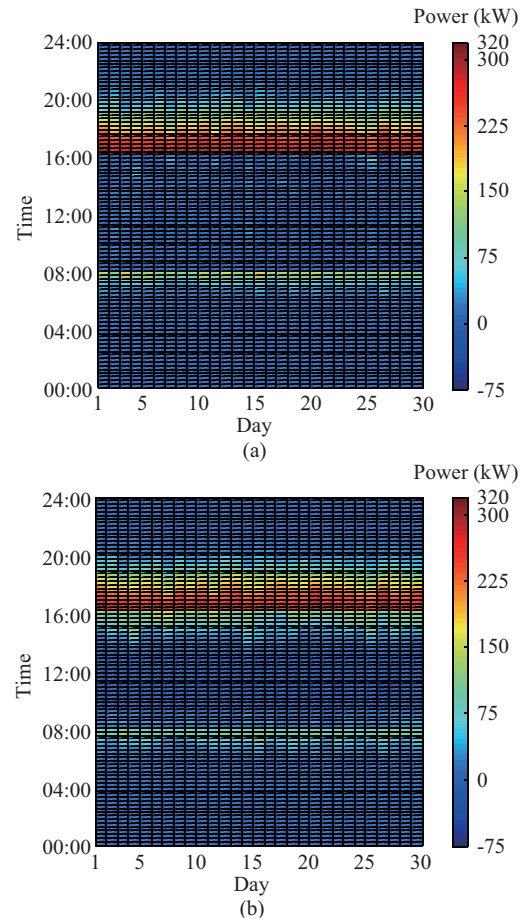


Fig. 9. Comparison of summed power scheduling results of all EVs (100 in total) on 30 consecutive weekdays. (a) Using MILP with complete information. (b) Using LSTM.

The power scheduling results of 30 consecutive days using the LSTM are slightly different from those of the MILP with complete information, particularly at around 08:00 (start time of the peak period of TOU tariff) and 16:00 (end time of the peak period of TOU tariff). Such differences are acceptable at the online execution stage, as what we have analyzed from the perspective of electricity cost and load leveling. Therefore, once the LSTM model is trained, it can run stably in the BEMS over a certain period of time.

## VI. CONCLUSION

In this study, to achieve effective cost saving and load leveling of the commercial-building prosumer, an LSTM-based machine learning method is proposed to schedule the EV charging and discharging considering the PV output and other electrical demands of the building. A preliminary data processing rule and an output filter have also been designed to improve the LSTM mapping performance. The proposed method can be separated into offline and online stages. At the offline stage, the LSTM network is trained to acquire the generalization ability to map from the states (inputs) to decisions (outputs) based on historical data. At the online stage, each EV's scheduling result can be quickly generated in a fully decentralized manner. The performance of the proposed LSTM method is compared with a commonly used MILP algorithm through a case study. The simulation results demonstrate that the electricity cost of the proposed LSTM method is close to the theoretical optimal solution, and better than that of the conventional MILP method, even without any prediction of the system data. Meanwhile, the solution time of the proposed LSTM method is of the order of milliseconds, which also means that it can help reduce the computational burden to a large extent.

Furthermore, the added preliminary data processing step can significantly help improve the accuracy of the network output, and the additional filtering has enhanced the load leveling effect. In general, the proposed LSTM-based algorithm can not only release the prediction and calculation pressures, but also achieve better results than the commonly used method in commercial building prosumer energy management.

Although the EV in the commercial building is used as an example to illustrate the effectiveness of the proposed method, one should note that the method can also be extended to optimize the power of other flexible loads in the commercial building such as heating ventilation and air conditioning (HVAC). The HVAC may be a more suitable dispatching resource in a shopping mall. At present, some related research works have been done on modeling of EV-HVAC using the virtual battery model. On this basis, it is easy to apply the method proposed in this paper to the energy management of a shopping mall by optimizing the power of HVAC.

## APPENDIX A

In this appendix, the formulation of MILP with incomplete information will be presented. In order to apply the MILP model to real-time scheduling, the rolling-horizon-optimization-based MILP model, which we call MILP with in-

complete information, is formulated in this paper. Solving this model relies on real-time information and some ultra-short-term prediction information of the rolling-horizon, i.e., incomplete information.

At the beginning of each time step, the BEMS carries out the optimization for the first coming time step  $t$  and the ultra-short-term prediction-based time horizon  $\mathcal{H}=\{t+1, t+2, \dots, t+H\}$ . For each optimization horizon  $[t, t+H]$ , the objective function (2) and constraints (3)-(16) will remain valid. The power scheduling of time step  $t$  and the time horizon  $\mathcal{H}$  are included in the optimization solutions. However, only the scheduling of time step  $t$  from the optimization solutions will be implemented. When the next time step  $t+1$  comes, the BEMS carries out the optimization again based on updated available information and new ultra-short-term prediction, and then only implements the scheduling of the time step  $t+1$ . The BEMS obtains the online scheduling by continuously implementing this process.

## REFERENCES

- [1] A. Samimi and M. Nikzad, "Complete active-reactive power resource scheduling of smart distribution system with high penetration of distributed energy resources," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 6, pp. 863-875, Nov. 2017.
- [2] J. Wang, H. Zhong, J. Qin *et al.*, "Incentive mechanism for sharing distributed energy resources," *Journal of Modern Power Systems and Clean Energy*, no. 7, no. 4, pp. 837-850, Jul. 2019.
- [3] M. Khorasany, Y. Mishra, B. Babaki *et al.*, "Enhancing scalability of peer-to-peer energy markets using adaptive segmentation method," *Journal of Modern Power Systems and Clean Energy*, no. 7, no. 4, pp. 791-801, Jul. 2019.
- [4] Y. Xue and X. Yu, "Beyond smart grid-cyber-physical-social system in energy future," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2290-2292, Dec. 2017.
- [5] Y. Song, Y. Ding, S. Pierluigi *et al.*, "Optimization methods and advanced applications for smart energy systems considering grid-interactive demand response," *Applied Energy*, vol. 259, pp. 1-3, Feb. 2020.
- [6] D. Azuatalam, A. C. Chapman, and G. Verbić, "Probabilistic assessment of impact of flexible loads under network tariffs in low-voltage distribution networks," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 4, pp. 951-962, Jul. 2021.
- [7] Z. Liu, Q. Wu, M. Shahidehpour *et al.*, "Transactive real-time electric vehicle charging management for commercial buildings with PV on-site generation," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4939-4950, Sept. 2019.
- [8] J. Hu, H. Zhou, Y. Li *et al.*, "Multi-time scale energy management strategy of aggregator characterized by photovoltaic generation and electric vehicles," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 727-736, Jul. 2020.
- [9] K. Sou, J. Weimer, H. Sandberg *et al.*, "Scheduling smart home appliances using mixed integer linear programming," in *Proceedings of 2011 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, USA, Dec. 2011, pp. 5144-5149.
- [10] K. Paridari, A. Parisio, H. Sandberg *et al.*, "Energy and CO<sub>2</sub> efficient scheduling of smart appliances in active houses equipped with batteries," in *Proceedings of 2014 IEEE International Conference on Automation Science and Engineering (CASE)*, Taipei, China, Aug. 2014, pp. 632-639.
- [11] T. Sousa, H. Morais, Z. Vale *et al.*, "Intelligent energy resource management considering vehicle-to-grid: a simulated annealing approach," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 535-542, Mar. 2012.
- [12] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 134-143, Sept. 2010.
- [13] M. Fukushima and G.-H. Lin, "Smoothing methods for mathematical programs with equilibrium constraints," in *Proceedings of International Conference on Informatics Research for Development of Knowledge Society Infrastructure*, Kyoto, Japan, Mar. 2004, pp. 206-213.



- [14] C. Hu, "Exactness of penalty functions for solving MPEC model of the transportation network optimization problems with user equilibrium constraints," in *Proceedings of 2006 International Conference on Management Science and Engineering*, Lille, France, Oct. 2006, pp. 2045-2049.
- [15] X. Xu, Y. Jia, Y. Xu *et al.*, "A multi-agent reinforcement learning based data-driven method for home energy management," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3201-3211, Jul. 2020.
- [16] F. Ruelens, B. J. Claessens, S. Vandael *et al.*, "Demand response of a heterogeneous cluster of electric water heaters using batch reinforcement learning," in *Proceedings of 2014 Power Systems Computation Conference (PSCC)*, Wroclaw, Poland, Aug. 2014, pp. 1-7.
- [17] H. Berlink and A. H. R. Costa, "Batch reinforcement learning for smart home energy management," in *Proceedings of 1st International Workshop on Social Influence Analysis/24th International Joint Conference on Artificial Intelligence (IJCAI)*, Buenos Aires, Argentina, Jul. 2015, pp. 2561-2567.
- [18] E. Mocanu, D. C. Mocanu, P. H. Nguyen *et al.*, "On-line building energy optimization using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3698-3708, Jul. 2019.
- [19] B. Wang, Y. Li, W. Ming *et al.*, "Deep reinforcement learning method for demand response management of interruptible load," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3146-3155, Jul. 2020.
- [20] C. Keerthisinghe, A. C. Chapman, and G. Verbic, "Energy management of PV-storage systems: policy approximations using machine learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 257-265, Jan. 2019.
- [21] K. Paridari, D. Azuatalam, A. C. Chapman *et al.*, "A plug-and-play home energy management algorithm using optimization and machine learning techniques," in *Proceedings of 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Aalborg, Denmark, Oct. 2018, pp. 1-6.
- [22] Y. Bao, T. Xiong, and Z. Hu, "Multi-step-ahead time series prediction using multiple-output support vector regression," *Neurocomputing*, vol. 129, pp. 482-493, Apr. 2014.
- [23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
- [24] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451-2471, Oct. 2000.
- [25] D. P. Kingma and J. Ba. (2014, Dec.). Adam: a method for stochastic optimization. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [26] S. M. Frank and P. K. Sen, "Estimation of electricity consumption in commercial buildings," in *Proceedings of 2011 North American Power Symposium*, Boston, USA, Aug. 2011, pp. 1-7.

**Huayanran Zhou** received the B.S. degree in electrical engineering and automation from North China Electric Power University, Baoding, China, in 2018. She 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 distributed energy resources energy management, transactive energy and their participations in prosumers energy management, and multi-energy system optimization.

**Yihong Zhou** received the B.S. degree in electrical engineering and its automation from North China Electric Power University, Beijing, China, in 2020. His research interests include machine learning, multi-agent system

and their application in prosumers energy management and smart grid.

**Junjie Hu** received the M.Sc. degree in control theory and control engineering from Tongji University, Shanghai, China, in 2010, and the Ph.D. degree in electrical engineering from the Technical University of Denmark, Copenhagen, Denmark, in 2014. He was a Postdoctoral Researcher with the Department of Electrical Engineering, Technical University of Denmark. He is currently an Associate Professor with the School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. His current research interests include distributed energy resources energy management, transactive energy and their participations in prosumers energy management, and multi-energy system optimization.

**Guangya Yang** received the Ph.D. degree in 2008 from the University of Queensland, St Lucia, Australia, in the field of electrical power systems. Currently, he is a Senior Power System Engineer with Ørsted and Senior Scientist with Technical University of Denmark, Copenhagen, Denmark. Since 2009, he has been developing and leading many industrial collaborative projects in Denmark. He is editorial board member of IEEE Transactions on Sustainable Energy, IEEE Transactions on Power Delivery, IEEE Access, and Journal of Modern Power System and Clean Energy. His research interests include power system dynamic security and protection, offshore wind power system design and control, and transactive energy applied to distributed energy resources.

**Dongliang Xie** received the Ph.D. degree in electrical engineering from Southeast University, China, in 2012. During 2011-2012, he was a Research Associate in Hong Kong Polytechnic University, Hong Kong, China. He is currently working for State Grid Electric Power Research Institute (SGE-PR), Nanjing, China. His research interests include analysis, simulation and control for smart grid architectures and essentials consisting of renewable power generation, power market and power system interactions, and demand elasticity.

**Yusheng Xue** received the Ph.D. degree from the University of Liege, Liege, Belgium, in 1987. He has been an Academician with the Chinese Academy of Engineering since 1995. He is currently the Honorary President of the State Grid Electric Power Research Institute (SGEPRI), Nanjing, China, an Adjunct Professor in dozens of Chinese universities, and an Adjunct Professor with the University of Newcastle, Newcastle, Australia. He is the Editor-in-Chief of the Automation of Electric Power System (in Chinese) and the Journal of Modern Power Systems and Clean Energy, as well as the Chairman of the Technical Committee of Chinese National Committee of CIGRE since 2005. His research interest includes power system automation.

**Lars Nordström** received the M.Sc. degree in electrical engineering from Royal Institute of Technology (KTH), Stockholm, Sweden, in 1996, and the Ph.D. degree in industrial information and control systems from KTH, in 2006. He is currently a Professor in information systems for power system control at KTH. He is head of the division of Electric Power and Energy Systems at KTH and Deputy Head of the School of Electrical Engineering and Computer Science, with specific responsibility for faculty development. His research and teaching are focused on issues at the crossroads of control, communication and power systems. His research interests include future architectures, functionality and quality aspects of information and communication systems used for power system control, operation, automation and protection.