Reusing Backup Batteries as BESS for Power Demand Reshaping in 5G and Beyond

Guoming Tang[1], Hao Yuan[2], Deke Guo[2,1], Kui Wu[3], Yi Wang[1,4*]
[1]Peng Cheng Laboratory, Shenzhen, Guangdong, China
[2]National University of Defense Technology, Changsha, Hunan, China
[3]University of Victoria, Victoria, BC, Canada
[4]Southern University of Science and Technology, Shenzhen, Guangdong, China

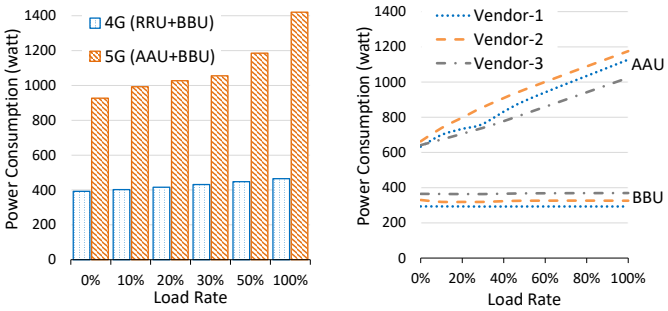
***Abstract*—The mobile network operators are upgrading their network facilities and shifting to the 5G era at an unprecedented pace. The huge operating expense (OPEX), mainly the energy consumption cost, has become the major concern of the operators. In this work, we investigate the energy cost-saving potential by transforming the backup batteries of base stations (BSs) to a distributed battery energy storage system (BESS). Specifically, to minimize the total energy cost, we model the distributed BESS discharge/charge scheduling as an optimization problem by incorporating comprehensive practical considerations. Then, considering the dynamic BS power demands in practice, we propose a deep reinforcement learning (DRL) based approach to make BESS scheduling decisions in real-time. The experiments using real-world BS deployment and traffic load data demonstrate that with our DRL-based BESS scheduling, the peak power demand charge of BSs can be reduced by up to 26.59%, and the yearly OPEX saving for $2,282$ 5G BSs could reach up to US$\$185,000$.**

***Index Terms*—5G cellular networks, backup batteries, BESS, power demand reshaping**

I. Introduction

To cope with the ever-increasing demands of mobile broadband, low-latency communications and IoT connections in the future wireless network, the worldwide mobile network operators are upgrading their network facilities and marching towards the 5G network at an unprecedented pace. Compared with the older generation networks like 4G/LTE, the signal range of a 5G base station (BS) is much shorter, resulting in the ultra-dense BS deployment (also know as *network densification* [1]), especially for fully signal coverage in the urban or "hotspot" areas. According to [2], by 2026 and in China only, over 14 million 5G BSs will be built.

To deploy and operate such a large-scale BSs, however, it needs an enormous investment, covering both capital expense (CAPEX) and operating expense (OPEX). The OPEX, in particular, is mainly attributed to the electricity consumed by the BSs. It has been well noticed that the newly constructed 5G BSs are with considerable power consumption. According to the power metering results from the mobile operators, as illustrated in Fig. 1, the power consumption of a 5G BS is typically 2∼3 times of that of a 4G BS. Considering the large-scale and ultra-dense deployment of 5G BSs in the near future, it could lead up to a tenfold increased electricity cost [3]. Consequently, the energy consumption of BSs could become a huge burden of the mobile operators, and how to cut down the energy cost is among the top priorities with the shift to 5G and beyond.



(a) Power Cons. (4G vs. 5G) (b) Power Cons. (5G AAU vs. BBU)

Fig. 1. Real-world power measurements of 4G and 5G BSs under varying load rates. The BS major components include RRU (for 4G), AAU (for 5G) and BBU (for both).

The energy cost of the mobile operator is typically composed by i) *energy charge* that is based on the total consumed electricity amount (in kWh) and ii) *demand charge* which is determined by the peak power demand (in kW) as a penalty. Particularly, the demand charge from severe unbalanced power consumption could account for a large proportion of the electricity bill, e.g., it could be up to 8x the energy charge for a commercial datacenter in Georgia, US [4]. As a result for the mobile operators, if they can cut down the demand charges, they can lower the energy cost accordingly, given that the total amount of consumed energy is constant. In the current cellular network, however, there seems no way to flatten the peak power demands, e.g., by peak power shaving strategy of workload shifting [5], as the real-time traffic demands from mobile users can hardly be shifted or even delayed.

Observations & Opportunities: When constructing a new BS, a backup battery with certain power capacity is usually deployed in accompaniment. It helps safeguard the BS's normal functioning against power outages for a reserved time (several to a dozen hours), as the power grid does not guarantee a hundred percent uptime. The backup batteries are

This work was supported in part by the National Key R&D Program of China (2019YFB1802600), the National Natural Science Foundation of China (No.61802421, No.U19B2024), the project of "FANet: PCL Future Greater-Bay Area Network Facilities for Large-scale Experiments and Applications" (No.LZC0019) and Guangdong Institute of Chinese Engineering Development Strategies (No.2019-GD-13). Corresponding authors: Yi Wang, Deke Guo.

necessary for nowadays BSs [6], especially those core and mission-critical BSs (e.g., colocating edge servers in mobile edge computing [7], [8]). Although a large-number of backup batteries have been or are going to be deployed, they are rarely used but lie idle in practice, especially in the urban areas where the grid power is ultra-stable.

To improve their utilization and cost efficiency, the backup batteries can be leveraged as natural candidates participating in the electricity regulation markets. Particularly, the backup battery can be transformed to and treated as a battery energy storage system (BESS). Then the distributed BESS can be utilized for BS power demand reshaping, e.g., discharging during the peak hours and recharging during the off-peak hours, and thus reduce the demand charges for the mobile operators. Meanwhile, the BESS's original duty of BS power backup can be barely affected.

Ideas & Challenges: Aiming at demand charge reduction, we exploit the potential benefits of operating the backup batteries as a distributed BESS for peak power demand shaving. Ideally, with strategic battery discharge/charge scheduling, the power demands of BSs could be reshaped and flattened as much as possible, and the incurred demand charge of the network system could thus be minimized.

To realize the above goals, however, we are faced with several challenges. Firstly, it is not straightforward to model the optimal battery discharge/charge scheduling problem in a distributed manner, especially considering the spatial-temporal varying power demands of all BSs in the network system. Secondly, as the lifetime of a battery is limited and shortened along with discharge/charge cycles, how to quantify the battery's degradation/replacement cost and balance it with the demand charge saving are critical and challenging. Furthermore, as the BESS's discharge/charge operations should obey certain physical laws (e.g., discharge/charge rates are limited) and are bounded by realistic constraints (e.g., high depth of discharge is prohibited), how to incorporate these factors with the BESS operating problem is non-trivial.

Contributions: By addressing the above challenges, we make the following contributions in this work.

- Based on the heterogeneous network (HetNet) architecture and distributed BESS scenario, we model the BESS discharge/charge scheduling as an optimization problem, which takes into account the practical considerations of BS power supply/demand as well as the backup battery specifications.
- After analysing the computational complexity of the optimization problem, we propose a deep reinforcement learning (DRL) based approach to solve the BESS discharge/charge scheduling problem. The DRL-based BESS scheduling approach accommodates all factors considered in the modeling phase and can adapt to the dynamic power demands from the BSs.
- Using a real-world BS deployment scenario and BS traffic load traces, we demonstrate that, leveraging the distributed BESS and our DRL-based scheduling approach, the demand charge of the mobile operator could be reduced by up to 26.59%, and corresponding yearly OPEX saving for $2,282$ 5G BSs could reach US$\$185,000$.

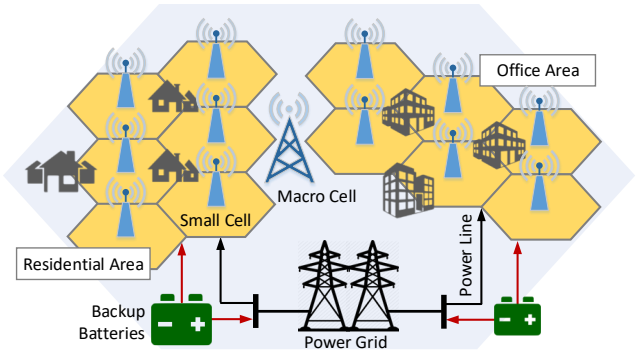


Fig. 2. Future 5G HetNet consisting of macro and small cells, and the backup power supplied by batteries. Ultra-dense small cells are deployed in the residential and office areas, where the amounts of end users and mobile traffic demands are large. The size of battery indicates its power capacity.

II. System Models

In this section, we present the system models applied and basic assumptions in this work. For clarity, the notations used in the modeling are explained in Table I.

A. Scenario Overview

As illustrated in Fig. 2, we consider the HetNet architecture for the future 5G network and beyond, where the macro BSs are deployed sparsely and mainly in charge of signal coverage in large-scale and the small BSs are densely deployed in the places where the mobile traffic demands are large, such as the residential and office areas. As the backup power of BSs, the backup battery[1] is equipped, either dedicated to one BS (e.g., the mission-critical and core BSs) or shared by multiple BSs (e.g., the densely deployed small BSs).

Given a targeted wireless network (5G or beyond), we assume that a number of M 5G BSs have been or are planning to be deployed. All the BSs can be denoted by a set $\mathcal{U} = \{1, 2, \cdots, M\}$. In the target network, we name the places where batteries are deployed as *battery points* (BPs). In practice, there exist natural BP candidates for backup battery deployments, such as the BS equipment cabinets/rooms (shown in Fig. 3) or the telecom central offices. Assume that the number of BPs is N and they are denoted by the set $\mathcal{V} = \{1, 2, \cdots, N\}$. As one battery can be potentially shared by multiple BSs, we have $N \leq M$ under our scenario.

B. BS Power Supply and Demand

As shown in Fig. 3, the power of BSs is directly supplied by the power grid and backuped by the battery. The backup battery can be installed at any BP, with a certain capacity, and dedicated to one BS or shared by multiple BSs nearby [9]. Particularly, those BSs sharing the same backup battery form a *virtual cell* (VC), as illustrated by Fig. 3.

[1]A backup battery typically consists of multiple battery cells, thus also known as a battery group. We use the word "battery" for simplicity.

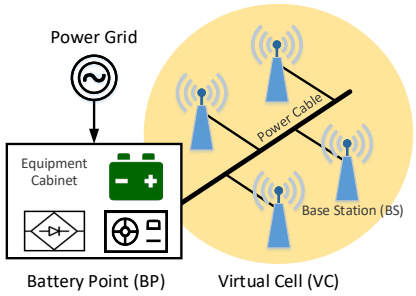


Fig. 3. Implementation of power supply and backup of a VC (containing four BSs) at a BP. The BP here is exemplified by an equipment cabinet, containing the rectifier (bottom-left of cabinet), power supply (bottom-right of cabinet) and backup battery (top-right of cabinet).

The battery capacity of each BP, referring to how fast and how much the BP can supply backup power to the BSs, is allocated based on the peak power demand of BSs within the associated VC. Particularly, to denote the battery capacity allocated at each BP, we define a *capacity vector* as follow:

$$c := [c_1, c_2, \cdots, c_N]^\intercal \tag{1}$$

in which c_n denotes the capacity of battery deployed at the n-th BP, typically in unit of ampere·hour or AH.

We consider a discrete time model, where the electricity billing cycle (e.g., one month) is evenly split into T consecutive slots with length of Δt and denoted by $\mathcal{W} = \{1, 2, \cdots, T\}$. Thus, at an arbitrary BP n, the power demand of associated BSs during the billing cycle can be defined by a *power demand vector*:

$$p_n := [p_n(1), p_n(2), \cdots, p_n(T)]^\intercal \tag{2}$$

where $p_n(t)$ is the power demand at BP n in time slot t. Note that the power demands can be obtained by power meter readings at each BP. Alternatively, referring to Fig. 1(b), the linear relationship between the traffic load and power consumption can be leveraged to estimate the power demands.

C. Battery Specification

At an arbitrary time slot t, we model the state of battery at BP n with a tuple:

$$\pi_n(t) := \langle SoH_n(t), SoC_n(t), DoD_n(t) \rangle \tag{3}$$

where the notations of SoH, SoC and DoD represent the *state of health*, *state of charge*, and *depth of discharge* of the battery, respectively. To be detailed: i) **SoH** indicates the health condition of the battery, by quantifying the battery's actual lifetime over time, as a percentage of a new battery's capability, ii) **SoC** gives how much energy the battery currently stores, as a percentage of its capacity, and iii) **DoD** indicates how much energy the battery has released, as a percentage of its capacity. Notice that SoC is actually the complement of DoD. We adopt both metrics by following the common practice in BESS modeling, which also makes later descriptions more clear.

For simplicity, we discretize the SoC of a battery into K equal-spaced states, e.g., ten SoCs states of $\{10\%, 20\%, \cdots, 100\%\}$. Among the discrete SoC states, we use SoC_{min} and SoC_{max} to indicate the minimum and maximum SoCs for the battery's discharging and recharging, respectively. Note that the minimum and maximum SoCs are recommended by the vendor and prevent the battery from over-discharging/charging.

TABLE I
SUMMARY OF NOTATIONS

Notation	*Description*
N, T	number of BPs and time slots, respectively
$\mathcal{V}, \mathcal{W}$	the set of total BPs and time slots, respectively
c_n	battery capacity allocated at BP n
$p_n(t)$	power demand of BSs at BP n at time t
$\pi_n(t)$	battery state of BP n at time t
b_n	discharged/charged power from/to battery at BP n
$\tilde{b}_n$	net discharged/charged power from/to battery at BP n
$A_{1:N}(t)$	aggregated power demand of whole system at time t
$A^{peak}_{1:N}$	aggregated peak power demand of the whole system
$\mathcal{J}^e(t)$	energy charge of the whole system at time t
$\mathcal{J}^d(t)$	demand charge of the whole system at time t
$\mathcal{J}^b(t)$	battery degradation cost of BESS at time t
α, β	discharging and charging efficiencies, respectively
$R+, R-$	max charge and discharge rates of battery, respectively
$\lambda_e, \lambda_d, \lambda_b$	prices of energy charge, demand charge, battery cost
$s_n(t)$	environment state of VC at BP n at time t
$a(t)$	action taken by the agent at time t
ϕ	mapping policy from environment states to actions
$R(t)$	reward function of the DQN
$Q, \tilde{Q}$	Q-values of the main net and target net, respectively
$\theta, \tilde{\theta}$	parameters of the main net and target net, respectively

III. Power Demand Reshaping via BESS Scheduling

The backup batteries deployed at the BSs can be regarded as a distributed BESS. By discharging during the peak power time and recharging during the off-peak power time, it is able to reshape the total power demands of BSs and reduce the demand charge of the energy cost.

To represent the battery discharge/charge schedule, we define the *backup power supply vector* for the battery at BP n:

$$b_n := [b_n(1), b_n(2), \cdots, b_n(T)]^\intercal \tag{4}$$

where $b_n(t)$ is a real number variable and can be i) **positive**: denoting the discharging power from the battery (instead of power grid) at BP n to the BSs during time slot t, ii) **negative**: indicating the charging power from the grid to the battery, or iii) **zero**: indicating that no discharging or charging happens.

A. Energy Cost with BESS

To quantify the energy cost (including energy charge and demand charge) of the targeted network system with BESS, we first assume the following electricity billing policy, which was widely applied in previous literature [4], [5], [10].

- Energy Charge: the total amount of consumed energy (in unit of kWh) of the network system multiplies the energy price (in unit of US$/kWh and denoted by λ_e).
- Demand Charge: the maximum power consumption (in unit of kW) of the network system during the billing cycle multiplies the peak power price (in unit of US$/kW and denoted by λ_d).

With the above billing policy, we then derive the energy charge and demand charge of the whole system (i.e., the targeted network system with BESS) for each time slot.

Given the power demands at each BP, $p_n(t)$, and the battery discharge/charge schedule, $b_n(t)$, we can formulate the *aggregated* power demand of the whole system for an arbitrary time slot t by:

$$A_{1:N}(t) = \sum_{n=1}^{N} \Big(p_n(t) - \tilde{b}_n(t) \Big) \tag{5}$$

where $\tilde{b}_n(t)$ is defined by:

$$\tilde{b}_n(t) = \begin{cases} \alpha \cdot b_n(t) & \text{, if } b_n(t) > 0 \\ b_n(t)/\beta & \text{, if } b_n(t) \leq 0 \end{cases} \tag{6}$$

and represents the *net power* released/recharged from/to the battery considering the power loss in AC-DC conversion and battery leakage, with $\alpha \in (0,1)$ and $\beta \in (0,1)$ denoting the discharging and charging efficiencies, respectively.

Then, the incurred **energy charge** of the whole system during time slot t can be represented by:

$$\mathcal{J}^e(t) = \lambda_e A_{1:N}(t) \Delta t. \tag{7}$$

And according to the billing policy, the incurred **demand charge** (for aggregated peak power demand) in time slot t can be represented by:

$$\mathcal{J}^d(t) = \max\Big\{ 0, \lambda_d \big(A_{1:N}(t) - A_{1:N}^{peak} \big) \Big\} \tag{8}$$

where $A_{1:N}^{peak}$ records the peak power demand of the whole system during the past $t-1$ time slots. After the time slot t, if $A_{1:N}(t) > A_{1:N}^{peak}$ (or $\mathcal{J}^d(t) > 0$), $A_{1:N}^{peak}$ will be updated to $A_{1:N}(t)$.

B. Battery Degradation Cost

Every cycle of discharge/charge does some "harm" to the battery and reduces its lifetime. The battery has to be replaced by a new one when its SoH drops down to a preset "dead" level, denoted by SoH_{dead} in this paper. In other words, there is an attached *degradation cost* for the battery with each discharge/charge cycle.

For any battery, we also know that it has a limited number of discharge/charge cycles, mainly determined by its chemical material (e.g., LA or LI) and usage pattern (e.g., DoD in each cycle). Thus, given the relationship between the DoD level and corresponding number of discharge/charge cycles, as illustrated in Fig. 7(a), we are able to evaluate the degradation cost for each specific discharge event.

Given the state of battery at BP n after time slot $t-1$, i.e., $\langle SoH_n(t-1), SoC_n(t-1), DoD_n(t-1) \rangle$, the SoH decrease of the battery during this time slot can be measured by:

$$\Delta SoH_n(t) = \begin{cases} \frac{1 - SoH_{dead}}{h(DoD_n(t-1) + \Delta DoD_n(t))} & \text{, if } b_n(t) > 0 \\ 0 & \text{, if } b_n(t) \leq 0 \end{cases} \tag{9}$$

where $h(\cdot)$ gives the total number of discharge/charge cycles that a battery can endure in its lifetime under an input DoD level (exemplified in Fig. 7(a)), and $\Delta DoD_n(t)$ gives the increase of DoD and can be calculated by:

$$\Delta DoD_n(t) = \frac{b_n(t) \Delta t}{c_n}. \tag{10}$$

With Eqt. (9), we attribute the accumulated SoH decrease over the charge/discharge cycle into each time slot of discharging.

With the above expression of SoH decrease at each BP, we can then formulate the **degradation cost** of the distributed BESS for any arbitrary time slot t:

$$\mathcal{J}^b(t) = \lambda_b \sum_{n=1}^{N} \Delta SoH_n(t) \tag{11}$$

in which λ_b is a coefficient converting the battery degradation to a monetary cost, with the unit of "US\$/SoH decrease".

C. Optimal BESS Operation Scheduling

1) Discharge/charge Rate Constraint: The maximum charging rate of the battery at a BP, denoted by R^+, gives the largest power that the battery can be recharged with during a time slot. The maximum discharging rate of the battery can be denoted by R^-, which shows the largest power that the battery can supply with in a time slot. Thus, the following constraint holds:

$$-R^+ \leq b_n(t) \leq R^- \tag{12}$$

and according to [11], R^- is usually 5$\sim$10 times larger than R^+ for LA batteries.

2) SoC Constraint: For any arbitrary time slot, the SoC of battery at each BP is expected to follow the constraint:

$$SoC_{min} \leq SoC_n(t) \leq SoC_{max} \tag{13}$$

where SoC_{min} and SoC_{max} are the specified minimum and maximum SoCs, respectively. Particularly, to safeguard the BP against potential power outages, SoC_{min} should be set following the required reserved time of the backup battery.

3) Battery State Updating: Given the battery state of BP n at time $t-1$, i.e., $\pi_n(t-1)$, the battery state at time t can be updated by:

$$\pi_n(t) \leftarrow \begin{cases} SoH_n(t) &= SoH_n(t-1) - \Delta SoH_n(t) \\ SoC_n(t) &= SoC_n(t-1) - \Delta DoD_n(t) \\ DoD_n(t) &= DoD_n(t-1) + \Delta DoD_n(t) \end{cases} \tag{14}$$

where $\Delta SoH_n(t)$ and $\Delta DoD_n(t)$ are given by Eqt. (9) and Eqt. (10), respectively, and $SoC_n(0), SoH_n(0)$ and $DoD_n(0)$ are the initial SoC, SoH and DoD of the battery, respectively.

For the billing cycle window $\mathcal{W}$, the following optimization problem can be formulated to find the optimal discharge/charge schedule for the distributed BESS.

$$\min_{b_n(t)} \quad \sum_{t=1}^{T} \Big(\mathcal{J}^e(t) + \mathcal{J}^d(t) + \mathcal{J}^b(t) \Big) \tag{15a}$$

$$\text{s.t.} \quad (12) \sim (14), \forall n \in \mathcal{V}, \forall t \in \mathcal{W}. \tag{15b}$$

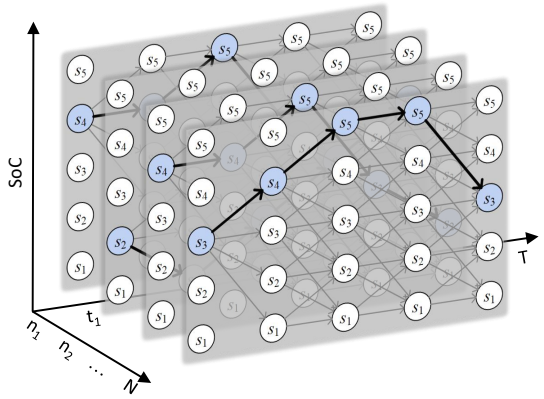


Fig. 4. Dynamic programming for discharge/charge scheduling of four batteries. One slice represents the scheduling process for one battery.

D. Problem Analysis

Dynamic programming (DP) has been applied to solve the battery discharge/charge scheduling problem for long [12], [13]. The main idea of the DP approach is to evaluate all possible discharging/charging sequences of the battery step by step and wrap up the optimal sequence by backward tracing. For example, as illustrated by one slice (for one battery case) in Fig. 4, given the initial state of a battery, all paths (costs) from current state to possible states at next time are calculated, from which the least cost path will be recorded. The above process goes on until the end of the billing cycle. The optimal path is then traced backwards from the last state, after all possible paths have been traversed. When applying the above DP approach for distributed BESS operating in our case, however, we have two major challenges.

1) High Computational Complexity: The computational complexity in our case is way higher than that of the conventional one. First, the number of batteries increases from one to hundreds or even thousands in our set scenario (as shown in Fig. 6). Given such a large-scale distributed BESS, its discharge/charge scheduling is much more complex than one battery case as given in [12], [13]. Second, the time window considered in our work is much longer than the aforementioned work. Although with the same time slot length (15-min), previous work designed the DP approach for a one-day scheduling, while we consider the real-world billing cycle of one-month. Such a (30x) time length extension results in a considerable complexity increase. The above two changes lead to a much larger searching space, compared with the one battery scheduling case, and thus prohibit any brute-force based searching approach for the optimal solution, including the DP approach introduced above. As a direct measurement, the complexity for the one-day scheduling of one battery is $\mathcal{O}(TK^2)$, where T is the number of time slots in one-day and K is the number of possible SoC changes for a battery. Nevertheless, the complexity in our case increases to $\mathcal{O}(T'K^2N)$, where $T'(\gg T)$ is the number of time slots in one-month billing cycle and N is the number of BPs among the whole network system.

2) Dynamic Power Demand: Another major defect of applying DP approach in our case is that, it assumes that the power demands are known in advance and searches for the optimal schedule in an offline way [12], [13]. This assumption is too strong for our problem, as the traffic demands of BSs as well as their power demands are indeed dynamic and cannot be obtained beforehand. Although there are works for traffic demands forecasting in wireless network [14], [15], the results are usually rough patterns and cannot be quite precise. As a consequence, the conventional DP approach for offline scheduling optimization is unable to be applied for our problem. A new approach to BESS scheduling, which can deal with the dynamic power demands and make (discharge/charge) decisions in real-time, is in great need.

Lyapunov stochastic optimization was applied for online operating of energy storage [16], [17], whereas it could not guarantee satisfactory outcomes for small and medium sized energy storage [18] nor be directly applied in a distributed manner as in our case. To tackle the above challenges, we propose an online discharge/charge scheduling approach based on deep reinforcement learning in the following section.

IV. A DRL-based Approach to Distributed BESS Scheduling

The constantly changing power demands of distributed BSs makes it hard to yield optimal BESS scheduling for model-based strategies, whereas it gives the opportunity to the learning-based methods. In this section, we present an online discharge/charge scheduling approach based on deep reinforcement learning (DRL), which exploits past experiences (e.g., historical BS power demands) for better decision-making by adapting to current state of environment (e.g., real-time BS power demands and BESS SoCs).

A. DRL based BESS Scheduling: Components & Concepts

A typical DRL framework consists of five key components: *agent*, *state*, *action*, *reward* and *policy*. Each component in our DRL-based BESS scheduling model is explained as follows.

- **Agent**: The role of the agent in BESS scheduling is to make decisions by interacting with the environment. Specifically, in our problem the agent determines the discharge/charge power at each BP in each time slot, according to the current states of power demand at each BS and SoC at each battery. The goal of the agent is to shave the peak power demand of the targeted network system, so as to minimize the total energy cost.
- **State**: The agent will be given a feedback state from the environment (i.e., the whole system) after taking a specific action. Specifically, at each time slot t, the state of the environment can be defined by a vector $s(t) = [A_{1:N}^{peak}, s_1(t), s_2(t), \ldots, s_N(t)]$, where $A_{1:N}^{peak}$ is the aggregated peak power demand in historical time slots and $s_n(t)(1 \leq n \leq N)$ is the state of VC at BP n and represented by $s_n(t) = \langle p_n(t), \pi_n(t) \rangle$.
- **Action**: The agent will take an action after observing the state of the environment. In our problem, the choice of action is a scheduling policy that indicates i) whether the battery should be charged or discharged and ii) how

much energy should be charged or released during each time slot. Therefore, the action taken at time t, denoted by $a(t)$, is equivalent to b_n defined in Eqt. (4). Notably the action space of agent is discretized according to the battery's discrete SoC states and restricted by the preset constraints of Eqts. (12) $\sim$ (14).

- **Policy**: The policy is defined by $\phi: \mathcal{S} \rightarrow \mathcal{A}$, where $\mathcal{S}$ and $\mathcal{A}$ represent the state and action spaces, respectively. Thus, the policy gives the mapping from environment states to specific actions. By $a(t) = \phi(s(t))$, it indicates the action to be taken at time t under the state $s(t)$. In our problem, we generate the policy by training a neural network during interacting with the environment.
- **Reward**: The agent takes an action at an arbitrary time t, $a(t)$, after observing the environment state $s(t)$. Then at the next time slot $t+1$, the agent will receive a reward from a reward function taking $a(t)$ and $s(t)$ as inputs. Thus the reward function is used to evaluate the effectiveness of the actions.

B. Reward Function Design

For an arbitrary time slot t, the agent first collects the state of the environment $s(t)$ and then takes action $a(t)$ accordingly. To evaluate the performance of action $a(t)$, we define the following reward function:

$$R(t) = V^e(t) + V^d(t) + V^b(t) \tag{16}$$

in which:

- $V^e(t) = -\mathcal{J}^e(t)$, measures the action reward of battery discharge/charge to the energy charge at time slot t;
- $V^d(t) = -\mathcal{J}^d(t)$, measures the action reward of battery discharge/charge to the demand charge at time slot t;
- $V^b(t) = -\mathcal{J}^b(t)$, measures the action reward of battery discharging to battery degradation at time slot t.

By setting the negative signs in front of $\mathcal{J}^e(t), \mathcal{J}^d(t)$ and $\mathcal{J}^b(t)$, the function is expected to reward any action that leads the decrease of i) energy charge, ii) demand charge, or iii) battery degradation.

Then, at the end of time slot t, the agent evaluates the performance of its action $a(t)$ using the defined reward function $R(t)$. During the following time slots, the agent aims to maximize the expected cumulative discounted reward: $\mathbb{E}[\sum_{t=1}^{\infty} \gamma^t R(t)]$, where $\gamma \in (0, 1]$ is a factor discounting future rewards.

Hence, the agent takes actions iteratively to approach the optimization objective as shown by Eqt. (15a), through maximizing its reward step by step. As the time goes, this process is expected to converge to an optimized action (i.e., discharge/charge schedule) for operating the distributed BESS.

C. Learning Process Design

In traditional reinforcement learning, Q-table is applied for the learning process (namely Q-learning). Due to the limited dimension of state vectors that the Q-table can handle, deep Q-network (DQN) was proposed recently by combining the neural network with Q-learning algorithm and successfully applied for high dimension state learning problems. Thus, in this work we leverage the DQN to learn the optimal BESS discharge/charge schedule. The proposed framework of BESS scheduling with DQN is illustrated in Fig. 5.

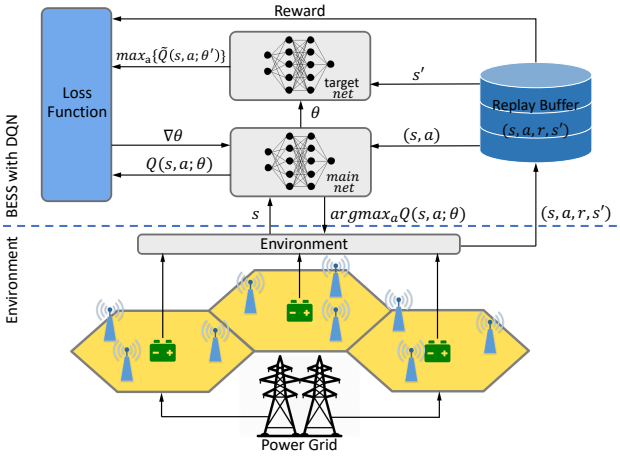


Fig. 5. The framework of BESS Scheduling with DQN.

- **Replay Buffer**: The method of experience replay is adopted in DQN to eliminate the correlation among input data and improve the learning effectiveness. Specifically, every time interacting with the environment, the agent stores its experience reflecting the performance of its action into the replay buffer, in form of $\langle s(t), a(t), R(t), s(t+1)\rangle$. In the following iterations, the agent will randomly choose (a mini-batch of) the stored experience from the replay buffer and then the algorithm will update the neural network parameters (θ in Fig. 5) by means of stochastic gradient descent (SGD).
- **Neural Networks**: Two neural networks, main net and target net, are constructed in our framework. The parameters in the main net are updated in real-time in response to the environment, while the target net copies the parameters from the main net every κ time slots. For an arbitrary time slot t, after receiving environment states, the main net generates the Q-value by:

$$Q(s(t), a(t)) \leftarrow \mathbb{E}\big[R(t) + \gamma \cdot \mathbb{E}[Q(s(t+1), a(t+1))]\big] \tag{17}$$

  where θ is the difference of network parameters between the main net and target net, and γ is the factor discounting the accumulative reward.
- **Loss Function**: The main net updates its network parameters through the following loss function, by minimizing the mean squared error (MSE) between its Q-value and the Q-value of target net:

$$L(\theta) \leftarrow \mathbb{E}[(\tilde{Q} - Q(s, a; \theta))^2] \tag{18}$$

  where θ is the main net parameters, and $\tilde{Q}$ is the Q-value of target net and calculated by:

$$\tilde{Q} \leftarrow R + \gamma \cdot \max_a\{Q(s, a; \tilde{\theta})\} \tag{19}$$

  in which $\tilde{\theta}$ is the target net parameters and updates every κ time slots by copying those of the main net.

Algorithm 1: DQN Training

Input: Power demands of BSs $p_n, 1 \leq n \leq N$
Output: Discharge/charge actions $a(t), 1 \leq t \leq T$
1 Initialize replay buffer (RB) to capacity N;
2 Initialize main network Q with random weights θ;
3 Initialize target network $\tilde{Q}$ with weights $\tilde{\theta} = \theta$;
4 **for** $iteration = 1 : MaxLoop$ **do**
5 **for** $t = 1 : T$ **do**
6 Get environment state $s(t)$;
7 $a(t) = \begin{cases} \text{argmax}_a Q(s(t), a(t); \theta), & \text{prob. } \epsilon \\ \text{random action, prob. } 1 - \epsilon \end{cases}$
8 Execute action $a(t)$ and receive $R(t)$ and $s(t+1)$;
9 Store $\langle (s(t), a(t), R(t), s(t+1) \rangle$ into RB;
10 Randomly sample a mini-batch of experience $\langle s(i), a(i), R(t), s(i+1) \rangle$ from RB;
11 $F(i) = \begin{cases} R(i), & \text{terminates at step } i+1 \\ R(i) + \gamma \cdot \max_a \{\tilde{Q}(s, a; \tilde{\theta})\}, & \text{else} \end{cases}$
12 Perform SGD on $(F(i) - Q(s, a; \theta))^2$ w.r.t. θ;
13 Set $\tilde{Q} = Q$ by every κ steps;
14 **end**
15 **end**

The DQN training process is depicted by the pseudo-code in Algorithm 1. First, the agent initializes the network parameters θ (Lines 1-3). Then, at each time slot t, the agent gets the environment state $s(t)$ and then selects an action $a(t)$ based on ϵ-greedy method (i.e., randomly selecting the action with the probability of $1 - \epsilon$), and chooses the action $\text{argmax}_a Q(s(t), a(t); \theta)$ (Lines 6-7). After taking the action and interacting with the environment, the agent receives the reward $R(t)$ and observes the next state $s(t+1)$ of the environment, and then stores the experience $\langle (s(t), a(t), R(t), s(t+1) \rangle$ into replay buffer (RB) (Lines 8-9). After interacting with the environment by randomly choosing (a mini-batch of) stored experience, the agent updates the main network parameters with SGD (Lines 10-12). The target network copies the parameters of the main network every κ time slots (Line 13).

Essentially, our DRL-based approach introduced above is different from all previous work performing offline BESS scheduling with fixed/known power demands. To be detailed, after training with adequate historical power demands data from the BSs, the DQN can make real-time decisions on discharge/charge scheduling of the BESS in a distributed manner. The output action from the DQN at each time slot (i.e., $a(t)$) gives the detailed power discharging or charging amount for each battery in the network system. According to our experimental results, the training process usually gets converged in several hundred iterations (refer to the convergence analysis by Fig. 9 in Sec. V-C).

V. Experimental Evaluations

A. Experiment Setup

1) BS and Traffic Demand Data: We consider a real-world BS deployment scenario in the downtown area of a metropolis in China. The locations of $2,282$ (4G/LTE) BSs are extracted from an open dataset on worldwide BS information [19],

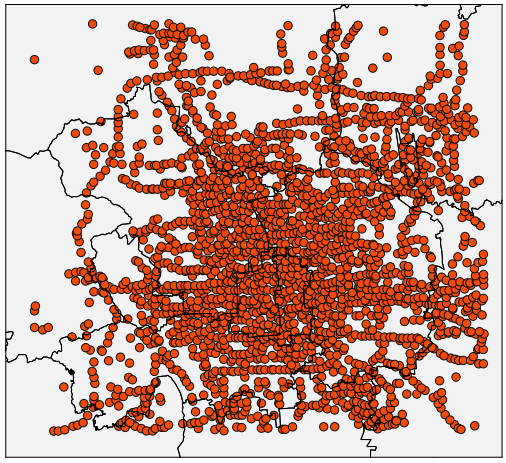


Fig. 6. The BS map of the downtown area of a metropolis in China. Each point shows the location of corresponding BS and the total number of BSs used in our experiments is $2,282$.

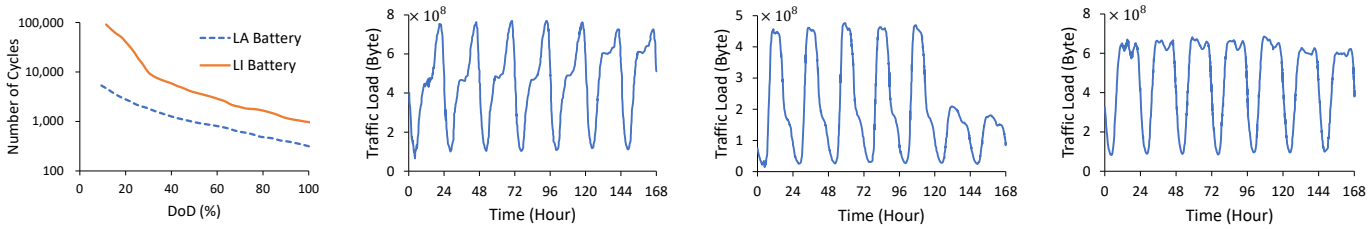
TABLE II
Parameter Settings

	Parameter	*Setting*
Billing Policy	billing cycle window $\mathcal{W}$	one month (30 days)
	[1]energy charge price λ_e	\$0.049/kWh
	[1]demand charge price λ_d	\$16.08/kW
	[2]battery cost λ_b	\$260/kWh(LA); \$271/kWh(LI)
Battery Config.	discharge efficiency α	75% (LA); 85% (LI)
	charge efficiency β	99.7% (LA); 99.9% (LI)
	max charge rate $R+$	16 MW (LA); 16 MW (LI)
	max discharge rate $R-$	8 MW (LA); 8 MW (LI)

[1]Prices of energy/demand charges in 2018, referring to the contract in [20]. Since we cannot find any public contract for energy/demand charges in China, the US prices are used only as a rough estimation.
[2]Battery capacity costs in 2018, referring to the data in [21].

and we mark them on the map in Fig. 6. Based on the traffic load patterns of BSs from the mobile operators in this metropolis, all the BSs considered in this scenario could be divided into five types, mainly determined by the areas where they are located [14]. Particularly, the traffic demands from BSs at the areas of *resident*, *office*, and *comprehensive* account for nearly ninety percentage of the total demands (resident: 17.55%, office: 45.72%, comprehensive: 24.81%). Thus, we only consider these three types of BSs in our experiments and their traffic demands within one-week period are illustrated in Fig. 7(b)-Fig. 7(d), respectively. We then refer to the relationship between the traffic demand and power consumption given by Fig. 1(b) and derive the power demands of corresponding BSs, assuming that the 4G/LTE BSs will be upgraded to 5G ones in the future.

2) Parameter Settings: For the BESS battery types, we consider two mainstream batteries as BESS on the current market: lead-acid (LA) and lithium-ion (LI). We first follow the common practice at reserved time setting for backup battery provision [6], and the battery capacity at each BP is set to support the peak power demand from corresponding VC for up to 12 hours. We then refer to [5], [20], [21] for parameter settings of electricity billing policy and BESS configurations. The parameter settings are summarized in Table II.



(a) DoD levels vs. #Lifetime cycles (b) Traffic loads (residential area) (c) Traffic loads (office area) (d) Traffic loads (comprehensive)

Fig. 7. (a) Relationship between DoD levels and battery lifetime (in number of discharge/charge cycles) for LA and LI batteries, respectively [22]. (b)-(d) Aggregated traffic loads of BSs at the residential, office and comprehensive areas, respectively, in one week period [14].

TABLE III
RESULTS SUMMARY (ONE BILLING CYCLE)

	Energy Charge ($)		Demand Charge ($)		Battery Cost ($)		Cost Saving ($)		Saving Ratio (%)	
	LA	LI	LA	LI	LA	LI	LA	LI	LA	LI
All w/o BESS	77,709	77,709	61,612	61,612	0	0	/	/	/	/
Part (10%) w/ BESS	77,825	77,780	60,067	59,940	400	39	1,029	1,562	0.74	1.12
Part (30%) w/ BESS	78,090	77,922	56,978	56,527	1,060	109	3,193	4,763	2.29	3.42
Part (50%) w/ BESS	78,459	78,019	53,889	53,246	2,020	181	4,954	7,874	3.56	5.65
All w/ BESS	79,022	78,309	46,165	**45,231**	4,151	376	9,984	**15,405**	7.17	**11.06**

3) Scenario Settings: We compare and analyze the overall energy cost (including energy charge, demand charge and batter cost), detailed scheduling results and return of investment (ROI) for the following scenarios:

- All w/o BESS: where all BSs are not equipped with BESS (the situation of current cellular networks);
- All w/ BESS: where all BSs are equipped with BESS and operated by our DRL-based scheduling strategy;
- Part w/ BESS: where only part ($10\%, 30\%$ and 50%) of the BSs are randomly chosen to equip with BESS and operated by the DRL-based scheduling approach.

B. General Performance at Cost Reduction with BESS

The results from all the set scenarios are summarized in Table III, from which we have the following major findings.

1) All w/ BESS: Compared with All w/o BESS, the cost reduction under this scenario is obvious. In one billing cycle (30 days), the overall cost savings are US$\$9,984$ and US$\$15,405$ for the LA and LI battery cases, respectively, corresponding to saving ratios of 7.17% and 11.06%, respectively. The demand charges are cut down significantly by 25.07% and 26.59% with LA and LI, respectively. Meanwhile, the energy charge slightly increases due to the energy loss in discharging and recharging. Thanks to the improvement of energy storage technologies, the battery degradation cost keeps at a well accepted level.

2) Part w/ BESS: The cost reduction grows with the increase of BESS deployment scale. For the LI battery case, the resulting cost saving ratios from 10%, 30% and 50% BESS deployment scales are 1.12%, 3.42% and 5.65%, respectively. The energy loss from discharge/charge cycles and battery degradation cost increase accordingly with the growth of BESS deployment scale.

With All w/ BESS, therefore, the yearly cost saving of the mobile operators could reach up to US$\$120,000$ and US$\$185,000$ in the LA and LI battery cases, respectively, merely for the $2,282$ 5G BSs in one city. Considering that the Telecom revenue growth is getting slower [23], such a cost reduction is a considerable saving for the mobile operators.

C. Case Studies of DRL-based BESS Scheduling

To study details of the BESS scheduling solutions from our DRL-based approach, we look into the BESS discharge/charge results at three BSs with different power demand patterns, in a one-day scheduling period. The results are illustrated in Fig. 8.

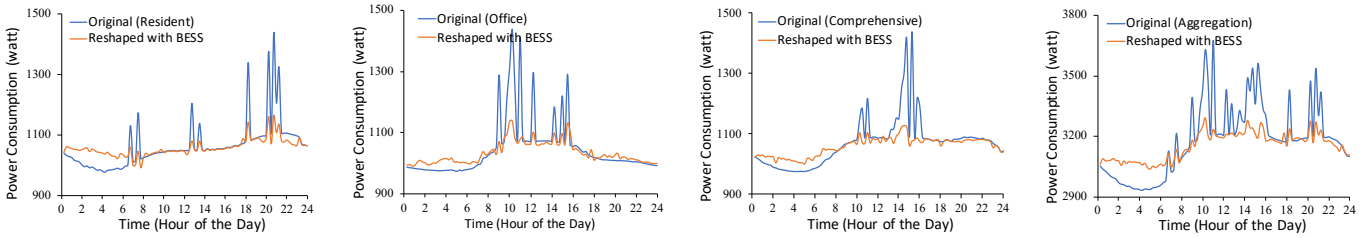
1) From an Overall View: The DRL-based approach can learn and adapt to the general power demand patterns of the BSs during the whole time period. This can be verified by the reshaped power demand curves, showing that the BESS charges itself in off-peak power time slots and releases power in peak power time slots. Such a scheduling ability might be possessed by previous model-based approaches, in dealing with the global trend of (offline) power demands.

2) From Local Views: The DRL-based approach can capture the dynamic changes precisely. We can see that, when the power demand increases/decreases sharply in a very short time, the BESS can make prompt response with a proper discharging/charging operation. This capability of tackling dynamic and uncertain power demands makes our approach superior to others in real-time BESS scheduling.

In addition, the convergence of DQN training process with our DRL-based approach is illustrated in Fig. 9 (for one arbitrary BS case). With one-week historical power demands data, the DQN in our case can be well trained after about 300 iterations, taking less than one hour on a commodity computer. The trained DQN is then applied for BESS scheduling in real-time, typically taking seconds for one decision-making.

D. ROIs of Different BESS Deployments

The return of investment (ROI) is a financial metric defined by the benefit (cost saving in our case) divided by the total investment. It measures the probability of gaining a return from an investment and has been widely used to evaluate the efficiency of an investment [24]. With the capacity costs of batteries (given in Table II), the total investments under



(a) Type-1 BS power reshaping. (b) Type-2 BS power reshaping. (c) Type-3 BS power reshaping. (d) Aggregated power reshaping.

Fig. 8. (a)-(c) Power demand reshaping (in one-day period) at three different BSs with types of resident, office and comprehensive, respectively. (d) Aggregated power demand reshaping for the three BSs in (a)-(c).

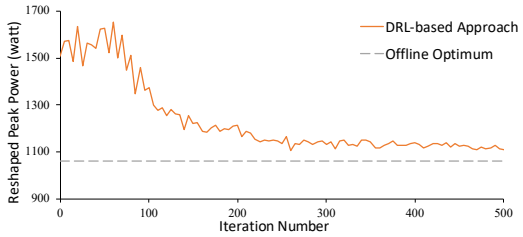


Fig. 9. Convergence of DQN training process.

different scenarios can be calculated. With cost saving values in Table III, the ROIs can thus be derived.

The ROIs for different scales of BESS deployment are shown in Table IV. We can find that with the LI battey, the ROI under scenario of All w/ BESS reaches 6.81%, indicating a relatively high investment efficiency. As the capacity cost of battery storage is estimated to decrease dramatically in the future (especially for the LI battery [25]), the ROI from large-scale BESS deployment could highly rise in 5G/B5G era.

VI. Related Work

In this section, we review relevant work on peak power demand shaving with battery energy storage system.

1) General System Peak Power Shaving with BESS: The optimal BESS discharge/charge scheduling problem was first investigated in [12], and a dynamic programming based approach was presented to tackle this problem for the general power system. A similar DP approach was applied in the later work of [13], in which the optimal sizing problem of BESS was also investigated. Although proved effective for the BESS operating in load reshaping and peak power shaving, the methods proposed are only applicable for one battery (as also presented in [26], [27]), which are different from the distributed batteries scenario in our case.

2) DC Peak Power Shaving with Centralized BESS: There are a great number of works leveraging the centralized BESS (i.e., the UPS) for data center (DC) peak power shaving. In [5], assuming that the power demands were all/partial known in a billing cycle, peak shaving methods with optimal DC battery control were presented to save the energy cost. Specifically, workload shifting strategies were leveraged in operating the battery, whereas in our problem, the workload (user traffic demands) cannot be shifted. The authors of [10] used energy storage in DC for peak demand charge reduction and regulation market participation jointly, and showed that the joint fashion with energy storage was superior than either

TABLE IV
ROIs of Scenarios with LA and LI Batteries, Respectively

Scenario	ROI (LA BESS)	ROI (LI BESS)
Part (10%) w/ BESS	4.31%	6.44%
Part (30%) w/ BESS	4.42%	6.53%
Part (50%) w/ BESS	4.47%	6.66%
All w/ BESS	4.58%	6.81%

separated one. The operating strategies in this work, however, were designed for one battery and based on known power demands, which are not suitable to our problem.

3) DC Peak Power Shaving with Distributed BESS: A distributed battery deployment scenario was considered for CDNs in [11], where a model for optimal battery provision was proposed to minimize the total power supply of distributed DCs. The battery discharge/charge scheduling, however, was not touched in this work. As the most relevant work to ours, [22] targeted a distributed battery scenario in the DC with fine-grained power backup at server level. It also purposed to reduce the peak power demand of the whole DC, whereas the BESS discharge/charge schedule was not optimized but chosen from pre-set policies, e.g., random discharging, least-recently-used (LRU) discharging and max-SoH discharging.

For peak power shaving of cellular networks in 5G and beyond, we are the first to (re)use the BS backup batteries as a distributed BESS system. The scale of batteries in our work is much larger than that in all previous work. Furthermore, the BESS operating schedule with our DRL-based approach is determined in a learning fashion, rather than those based on fixed models [11] or pre-set policies [22].

VII. Conclusions

To cut down the energy cost of mobile operators in the shift to 5G and beyond, we proposed to reuse the backup batteries of BSs as a distributed BESS for peak demanded power shaving. An optimization problem to minimize the total energy cost was established, incorporating the BESS discharge/charge scheduling with practical considerations in electricity billing policy and battery specifications. To solve the problem under the dynamic power demands, we proposed a DRL-based approach that accommodates all factors in the modeling phase and makes decisions in real-time. Using real-world BS deployment and traffic demand data, the experiments under different scenarios show that our solution can significantly cut down the peak power demand charge and total energy cost for the mobile operators.

REFERENCES

[1] J. Liu, M. Sheng, L. Liu, and J. Li, "Network densification in 5g: From the short-range communications perspective," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 96–102, 2017.

[2] HUAWEI, "5g telecom power target network," https://carrier.huawei.com/~/media/CNBGV2/download/products/network-energy/5G-Telecom-Energy-Target-Network-White-Paper.pdf, 2019.

[3] MTN Consulting, "Operators facing power cost crunch," https://www.mtnconsulting.biz/product/operators-facing-power-cost-crunch/, 2020.

[4] H. Xu and B. Li, "Reducing electricity demand charge for data centers with partial execution," in *ACM e-Energy*, 2014, pp. 51–61.

[5] M. Dabbagh, B. Hamdaoui, A. Rayes, and M. Guizani, "Shaving data center power demand peaks through energy storage and workload shifting control," *IEEE Transactions on Cloud Computing*, 2017.

[6] F. Wang, X. Fan, F. Wang, and J. Liu, "Backup battery analysis and allocation against power outage for cellular base stations," *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 520–533, 2018.

[7] G. Tang, D. Guo, K. Wu, F. Liu, and Y. Qin, "Qos guaranteed edge cloud resource provisioning for vehicle fleets," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5889–5900, 2020.

[8] H. Guo, J. Liu, J. Zhang, W. Sun, and N. Kato, "Mobile-edge computation offloading for ultradense iot networks," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4977–4988, 2018.

[9] G. Tang, Y. Wang, and H. Lu, "Shiftguard: Towards reliable 5g network by optimal backup power allocation," in *IEEE SmartGridComm*, 2020, pp. 1–6.

[10] Y. Shi, B. Xu, B. Zhang, and D. Wang, "Leveraging energy storage to optimize data center electricity cost in emerging power markets," in *ACM e-Energy*, 2016, pp. 1–13.

[11] D. S. Palasamudram, R. K. Sitaraman, B. Urgaonkar, and R. Urgaonkar, "Using batteries to reduce the power costs of internet-scale distributed networks," in *ACM SoCC*, 2012, pp. 1–14.

[12] D. K. Maly and K.-S. Kwan, "Optimal battery energy storage system (bess) charge scheduling with dynamic programming," *IEE Proceedings-Science, Measurement and Technology*, vol. 142, no. 6, pp. 453–458, 1995.

[13] A. Oudalov, R. Cherkaoui, and A. Beguin, "Sizing and optimal operation of battery energy storage system for peak shaving application," in *Lausanne Power Tech*. IEEE, 2007, pp. 621–625.

[14] F. Xu, Y. Li, H. Wang, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," *IEEE/ACM Transactions on Networking*, vol. 25, no. 2, pp. 1147–1161, 2016.

[15] L. Chen, D. Yang, D. Zhang, C. Wang, J. Li *et al.*, "Deep mobile traffic forecast and complementary base station clustering for c-ran optimization," *Journal of Network and Computer Applications*, vol. 121, pp. 59–69, 2018.

[16] Y. Guo and Y. Fang, "Electricity cost saving strategy in data centers by using energy storage," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1149–1160, 2012.

[17] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *ACM SIGMETRICS*, 2011, pp. 221–232.

[18] C.-K. Chau, G. Zhang, and M. Chen, "Cost minimizing online algorithms for energy storage management with worst-case guarantee," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2691–2702, 2016.

[19] OpenCellid, https://opencellid.org/, 2020.

[20] Dominion Energy South Carolina, Inc., "Rate 23 - industrial power service," https://etariff.psc.sc.gov/Organization/TariffDetail/150?OrgId=411, 2020.

[21] US Department of Energy, "Energy storage technology and cost characterization report," https://www.energy.gov/eere/water/downloads/energy-storage-technology-and-cost-characterization-report, 2019.

[22] B. Aksanli, T. Rosing, and E. Pettis, "Distributed battery control for peak power shaving in datacenters," in *IEEE IGCC*, 2013, pp. 1–8.

[23] Statista, "Forecast growth worldwide telecom services spending from 2019 to 2023," https://www.statista.com/statistics/323006/worldwide-telecom-services-spending-growth-forecast/, 2020.

[24] Wikipedia, "Return on investment," https://en.wikipedia.org/wiki/Return_on_investment, 2020.

[25] National Renewable Energy Laboratory (NREL), "Cost projections for utility-scale battery storage," https://www.nrel.gov/docs/fy19osti/73222.pdf, 2019.

[26] H. Wang and B. Zhang, "Energy storage arbitrage in real-time markets via reinforcement learning," in *IEEE Power & Energy Society General Meeting*, 2018, pp. 1–5.

[27] L. Yang, M. H. Hajiesmaili, R. Sitaraman, A. Wierman, E. Mallada, and W. S. Wong, "Online linear optimization with inventory management constraints," *ACM POMACS*, vol. 4, no. 1, pp. 1–29, 2020.