RESEARCH ARTICLE

Energy Cooperation Among Sustainable Base Stations in Multi-Operator Cellular Networks

**ANIKA TAHSIN[1], PALASH ROY[1,2], MD. ABDUR RAZZAQUE[1], (Senior Member, IEEE),
MD. MAMUN-OR-RASHID[1], MOHAMMAD SIRAJ[3],
SALMAN A. ALQAHTANI[4], (Member, IEEE), MD. RAFIUL HASSAN[5],
AND MOHAMMAD MEHEDI HASSAN[6], (Senior Member, IEEE)**

[1]Green Networking Research Group, Department of Computer Science and Engineering, University of Dhaka, Dhaka 1000, Bangladesh
[2]Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka 1207, Bangladesh
[3]Department of Electrical Engineering, College of Engineering, King Saud University, Riyadh 11543, Saudi Arabia
[4]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[5]College of Arts and Sciences, University of Maine at Presque Isle, Presque Isle, ME 04769, USA
[6]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Mohammad Mehedi Hassan (mmhassan@ksu.edu.sa)

ABSTRACT Energy Harvesting technology contributes significantly to green cellular networking by ensuring self-sustainability and extinguishing environmental hazards. Due to the imbalance between the harvested energy and traffic load of the base stations (BSs), energy cooperation has become a crucial requirement. However, the decision of optimal energy cooperation among the BSs in a multi-operator cellular network is a challenging task due to the consideration of various factors, such as cost, loss of energy, future information of traffic load, and harvested energy of the BSs, etc. The two conflicting objectives are minimizing the energy buying cost and the loss of energy while transferring through the power links. In this work, we present an optimal energy cooperation framework, formulated as a multi-objective linear programming (MOLP) problem which brings a trade-off between the two above-mentioned conflicting objectives considering the harvested energy and load of the BSs at future time slots. For the prediction of harvested energy of the BSs, we develop a Deep Q-Learning-based prediction method that intelligently increases measurement accuracy through continuous exploration and exploitation. The results of simulation experiments carried out in MATLAB depict that the proposed multi-operator energy cooperation framework outperforms state-of-the-art works in terms of cost, performance, and energy-loss reduction.

INDEX TERMS Energy harvesting, green cellular networking, energy cooperation, energy sustainability, deep Q-learning, optimization.

I. INTRODUCTION

The Cellular telecommunication has been playing a crucial role in the domain of information and communication technology (ICT). At present, the number of mobile phone users is drastically increasing and currently, it is nearly 7.8 billion [1], [2]. With the advent of mobile Internet services and the emergence of data-intensive applications, the volume of mobile data traffic in the cellular network is surging up gradually [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed M. A. Moustafa.

And to keep pace with them, the deployment of cellular base stations (BSs) is also being escalated.

The major part of energy consumption in the cellular network (around 60% - 80%) is held by BSs [4]. Therefore, the increase in the number of BSs contributes to higher energy consumption and increased operating expenditure (OPEX), a large part of which is dominated by electric bills [5]. Due to the shortage and unavailability of electricity in the rural areas, the mobile operators proceeded towards using diesel generators which is not considered a viable option because of its increased OPEX approximately by 10 times [6]. The

revenue of mobile operators is decreasing because of the ever-increasing OPEX generated from the maintenance of the increased number of BSs. On the other hand, electricity consumption in the cellular BSs contributes immensely to carbon emission thus causing environmental hazards [7]. Therefore, while dealing with the increased number of BSs, it is not possible to provide power to all of them from the electricity grid in a cost-effective way [8]. Therefore, the energy harvesting (EH) technology of green wireless cellular networks is anticipated as a promising technology that will be cost-effective and environmentally friendly.

Energy harvesting is an emerging technology that converts the ambient energy from the environment (from solar, wind, sound, kinetic energy, etc.) to electrical energy that will power autonomous electronic devices to operate or store the energy for later use [9]. This technology can enable low-power BSs to capture energy from the environment and serve its users with that energy. However, the availability of ambient environmental energy is uncertain in most cases, such as, solar energy will be unavailable during rainy days. Therefore, combining energy-harvesting capable BSs with BSs purchasing energy from electrical grids can be considered a promising solution that will diminish the OPEX of mobile operators and also cause less environmental hazard [10].

In this research work, we advocate cellular networks where energy harvesting and storage devices will be implemented inside the BSs. The BSs are of two types, i.e., off-grid BSs that only operate using the harvested energy and on-grid BSs that are also capable of purchasing energy from the electricity grid as well as using the harvested energy. Because of the imbalance in harvested energy and loads in the BSs, energy sharing among them has become a crucial requirement [11].

Energy cooperation among the BSs has been well studied in a few papers [12], [13], [14]. In [12], the authors have introduced the concept of energy sharing between two nodes in the networks. In [13], the authors have proposed optimal energy allocation among the BSs as a convex optimization problem with an objective to minimize the energy mismatches and the distance between the source and consumer BSs. However, they have not considered multi-operator cellular networks and any pricing while sharing energy. In [14], the authors have proposed an optimized energy transfer framework among the BSs considering energy pricing. However, they only tried to minimize the energy buying cost of the consumer BSs and they did not consider the minimization of the loss of energy while transferring energy from source BSs to consumer BSs.

Energy cooperation plays a vital role in ensuring self-sufficiency and promoting green cellular networking. However, taking the decision of optimal energy cooperation is a challenging task due to the consideration of various factors, such as cost, service performance, energy loss, etc. In a real-life cellular network scenario, usually, the BSs of different operators are deployed more closely in a certain geographical area or zone than the BSs of the same operator. Therefore,

in most cases, inter-operator BS distance is much less than that of intra-operator BSs. If two BSs of different operators share energy among themselves, the consumer BS will have to pay a price to the source BS. On the other hand, receiving energy from a source BS of the same operator encounters no energy buying cost.

While transferring energy from a source BS to a consumer BS through the sequence of power links between them, a certain amount of energy loss occurs due to the resistivity of the link and it is proportional to the length of the power link between them [15]. The more the distance between the source and consumer BSs, the more energy loss will be encountered. Therefore, to minimize energy loss, it will be better if consumer BSs receive energy from nearby source BSs. However, this will increase the energy buying cost of the consumer BSs as BSs of different operators are usually more adjacent and it requires energy buying cost receiving energy from a different operator BS. On the other hand, to minimize cost, it is preferable to choose the source BS of the same operator but the distance between source and consumer BSs rises in this case thus resulting in an increased loss of energy. Therefore, minimizing both the loss of energy and energy buying costs are two conflicting objectives. In this paper, we have developed a framework for the **M**inimization of **C**ost and **E**energy **L**oss namely MCEL for cellular networks. The main contributions of this paper are summarised as follows:

- We formulate an optimal energy cooperation framework for the BSs of multiple network operators as a multi-objective linear programming (MOLP) problem.
- We bring a trade-off between the above-mentioned two conflicting objectives, i.e., minimizing the energy buying cost of the consumer BSs and minimizing the loss of energy while transferring through power links.
- We have explored and exploited the deep Q-learning method for the prediction of harvested energy of the BSs in the future time slot while determining the energy demand/surplus of the BSs.
- We have simulated our proposed system in MATLAB and studied comprehensive performance analysis and compared the proposed system with other state-of-the-art works.

The rest of the paper is organized as follows. Section II presents some related works. Sections III and IV describe the system architecture and solution approach of our proposed system respectively. Section V represents the performance of the proposed solution and comparison with other state-of-the-art works. Finally, section VI presents the conclusion which summarizes our contribution and future scope of work.

II. RELATED WORKS
Due to the rapid increase in the number of BSs contributing to a higher carbon footprint, energy harvesting in the BSs is gaining much attention. Many international research projects have been working on green cellular networking and energy efficiency in cellular networking over the decade, such as

EARTH [16], ECONET [17], TREND [18] etc. A wide range of work has been done on energy management and redistribution among the BSs in a cellular network. In [19], an optimal energy management solution is given for the BSs with energy harvesting capability connected to the electrical grid where they tried to minimize the cost of energy buying from the grid. In [20], multiple electrical grid retailers have been encountered and energy management of the BSs has been presented to minimize cost and carbon footprint. However, they didn't consider any energy sharing among the BSs to meet the energy demands.

Energy cooperation among the BSs has been studied in a good number of works [21], [22], [23], [24], [25], [26]. However, in [21], the authors have not encountered energy harvesting at the BS sites and have not studied the uncertainty related to it. In [26], the case of energy sharing is studied for two sustainable BSs and the energy is transferred through the direct link between them. Several energy cooperation architectures have been presented in the literature. In [27], an energy-sharing scheme is given, where an entity called an aggregator, which works as a mediator between a group of BSs and the electrical power grid. When the BSs have surplus energy they inject power into the aggregator and in case of shortage, they can drag power from it. In this scheme, some amount of harvested energy can get wasted if none of the BSs drags it when it is injected. In [28], wireless power transfer (WPT) and packet power grid (PPG) energy-sharing architectures are presented. Their extensive analysis shows that the power transfer efficiency of WPT is not as good as PPG [29].

The authors of [13] have proposed an energy cooperation framework among the BSs and they have tried to minimize the distance between source and consumer BSs to minimize the loss of energy while transferring through the power link between them. They have also tried to minimize the mismatch between the shareable energy of source BSs and the energy demand of consumer BSs. However, they have not considered a multi-operator network scenario and have not encountered any energy pricing which is not practical in real life. They also do not consider the prediction of traffic load and harvested energy at future time slots of a certain BS. Rather, they have calculated the energy demand or shareable surplus energy of a BS based on a threshold battery state.

In [14], the authors have proposed an optimal energy-sharing framework that focuses on the minimization of the energy buying cost of the consumer BSs. However, they have not considered the multi-operator network scenario. They also don't consider the loss of energy while transferring energy from source to consumer BSs. For determining the energy demand of the BSs, they have used threshold energy levels rather than using any prediction of harvested energy and traffic load of the BSs in the next time slot. In [30], the authors have proposed an energy cooperation framework among the BSs such that the loss of energy due to energy transfer between source and consumer BSs is minimized.
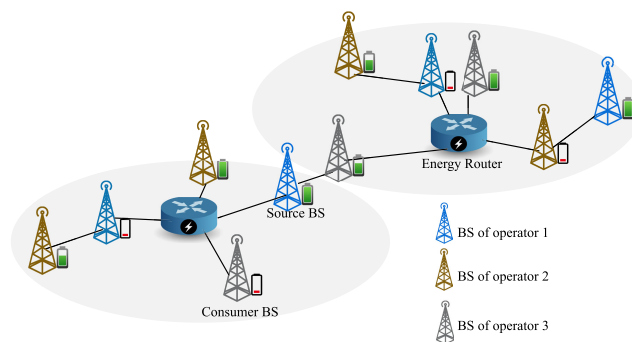

FIGURE 1. Energy cooperation architecture in multi-operator cellular networks.

They have presented a prediction approach for harvested energy and traffic load of the BSs to determine the energy demand of the BSs in future time slots. However, they do not consider multi-operator cellular networks and did not consider any energy pricing of the BSs.

None of the existing works considers the trade-off between the two conflicting objectives, i.e., to make a good balance between minimizing the energy buying cost of the BSs and minimizing the loss of energy due to energy transfer. In this paper, we present an optimal energy cooperation framework that brings a trade-off between these two conflicting objectives. We have considered the multi-operator cellular network scenarios and the prediction of harvested energy and traffic load of the BSs for determining the demand of consumer BSs and the available energy of the source BSs in future time slots.

III. SYSTEM MODEL AND ASSUMPTIONS

The system architecture of our proposed system is represented in Fig. 1. Here, we have considered a multi-operator cellular network comprising a set of BSs, denoted by N.

Each BS is equipped with a solar energy harvesting module and energy storage. Thus, they are capable of solar energy harvesting and storing the energy for later use. The BSs are classified into two types, i.e., on-grid BS and off-grid BS. The off-grid BSs are only bounded to operate and serve their users using the energy they acquire from energy harvesting or they can use the derived energy from other BSs in case of energy shortage. And in the case of surplus energy, they can be energy sources and transfer energy to other consumers' BSs. On the other hand, the on-grid BSs can purchase energy from the grid in case of energy shortage along with using their own harvested energy. In case of energy surplus, the on-grid BSs will serve the consumer BSs by transferring energy, and the residual energy after the energy sharing with the consumer BSs will be sold back to the grid. The off-grid BSs, denoted by $N_{off} \subset N$, can act both as a source BS and a consumer BS whereas the on-grid BSs, denoted by $N_{on} \subset N$, can only act as a source BS when it has an energy surplus.

A packet power grid (PPG) is implemented in this network for energy sharing among the BSs. In PPG, an entity called

TABLE 1. Notation table.

Notation	Description
N	The set of all BSs in the network
N_{off}, N_{on}	The set of off-grid BSs and on-grid BSs
$\mathcal{X}^c, \mathcal{X}^s$	The set of source BSs and consumer BSs
$B_i(t)$	The energy state of BS $i \in N$ at the beginning of time slot t
$E_i^h(t)$	The harvested energy of BS $i \in N$ during time slot t
$E_i^c(t)$	The energy consumption of BS $i \in N$ during time slot t
$E_i^s(t)$	The amount of total energy transferred to or from BS $i \in N$ during time slot t
$G_i(t)$	The amount of total energy BS $i \in N$ receives from or transfers to the grid during time slot t
$E_i^{c'}(t+1)$	The energy needed for BS $i \in N$ to serve the predicted load at time slot $t+1$
d_i	The energy demand of consumer BS $i \in \mathcal{X}^c$
$\mathcal{E}_i$	Total shareable surplus energy of source BS $i \in \mathcal{X}^s$
x_{ij}	The fraction of shareable energy of source BS $j \in \mathcal{X}^s$ transferred to consumer BS $i \in \mathcal{X}^c$
$\mathcal{E}_{ij}^s$	The amount of energy transferred to consumer BS $i \in \mathcal{X}^c$ from source BS $j \in \mathcal{X}^s$
ϕ_{ij}	The energy loss factor transferring energy from source BS $j \in \mathcal{X}^s$ to consumer BS $i \in \mathcal{X}^c$
ζ_{ij}	The normalized cost consumer BS $i \in \mathcal{X}^c$ pays to source BS $j \in \mathcal{X}^s$ for transferred energy.
δ_m	Maximum allowable relative standard deviation of the remaining energy of the source BSs
ρ_i	Minimum portion of energy demand of consumer BS $i \in \mathcal{X}^c$ that must be fulfilled

energy router is responsible for taking the decision of energy routing and finding suitable source BSs for the consumer BSs under it from where the consumer BSs can seek energy [13]. An energy router is responsible for maintaining the energy cooperation among the BSs of a certain area. The energy routers can collaborate and manage source BSs for that energy router has a shortage of source BSs to fulfill the energy demand of the consumer BSs under its domain. According to [15], the BSs are connected through a sequence of DC power lines and the energy transfer between them occurs in time division multiplexing (TDM) fashion. That means, when a particular link is busy transmitting the energy of a particular energy trade, it can not be accessed for any other energy transmissions. A single path comprising a sequence of one or more power links is considered from one BS to another in this network. While transferring energy from one BS to another, a certain amount of energy loss is induced that is proportional to the resistivity as well as the length of the power link between the source and consumer BS. We assume that the time is slotted and the length of the one-time slot is denoted by τ. The major notations used to design MCEL and their descriptions are presented in table 1.

IV. DESIGN OF MCEL
In this section, we unfold the operational details of various components of the proposed energy cooperation system, MCEL. First, we determine the source and consumer BSs based on their predicted energy harvesting and consumption profiles. Then, we present an optimization framework for optimal energy cooperation among the BSs. An energy router

manages the allocation of optimal energy sources for the consumer BSs under its domain.

A. DETERMINING SOURCE AND CONSUMER BSs
The energy state of an off-grid BS $i \in N_{off}$ at the beginning of a time slot $t+1$ is represented by,

$$B_i(t+1) = B_i(t) + E_i^h(t) - E_i^c(t) \pm E_i^s(t), \quad (1)$$

where, $B_i(t)$ is the stored energy of BS $i \in N_{off}$ at the beginning of time slot t, $E_i^h(t)$ is the harvested energy of BS $i \in N_{off}$ during time slot t, $E_i^c(t)$ is the energy consumption of BS $i \in N_{off}$ during time slot t and $E_i^s(t)$ is the amount of total energy transferred during time slot t, which will be added if BS $i \in N_{off}$ was a consumer or subtracted if BS $i \in N_{off}$ was a source at time slot t.

On the other hand, the energy state of an on-grid BS $i \in N_{on}$ at the beginning of a time slot $t+1$ is represented by,

$$B_i(t+1) = B_i(t) + E_i^h(t) - E_i^c(t) - E_i^s(t) \pm G_i(t), \quad (2)$$

where, $B_i(t)$ is the stored energy of BS $i \in N_{on}$ at the beginning of time slot t, $E_i^h(t)$ is the harvested energy of BS $i \in N_{on}$ during time slot t, $E_i^c(t)$ is the energy consumption of BS $i \in N_{on}$ during time slot t, $E_i^s(t)$ is the amount of total energy transferred from BS $i \in N_{on}$ during time slot t and $G_i(t)$ is the amount of energy shared with the grid at time slot t. It will be added if energy is purchased from the grid. Otherwise, it will be subtracted if energy is transferred to the grid at time slot t.

The set of consumer BSs under the router where the system is running is denoted by $\mathcal{X}^c$ and the set of source BSs is denoted by $\mathcal{X}^s$. The traffic load of a BS $i \in N$ at the next time slot is denoted by $L_i^p(t+1)$ and we calculate it using a hybrid traffic prediction model according to [31] which is a combination of double seasonal ARIMA (DSARIMA) and long-short term memory (LSTM) based networks. The energy needed for BS $i \in N$ to serve the predicted load at time slot $t+1$ is represented by $E_i^{c'}(t+1)$. The predicted harvested energy of a BS $i \in N$ in the next time slot is denoted by $E_i^{h'}(t+1)$ and we have proposed a deep Q-learning based approach to derive this.

For a particular BS (off-grid or on-grid) $i \in N$, if $E_i^{c'}(t+1) - \{B_i(t+1) + E_i^{h'}(t+1)\} \geq \omega_i$, where ω_i denotes a threshold energy level value that BS $i \in N$ wants to reserve as backup, it acts as an energy source at time slot $t+1$ and the amount of energy it is ready to transfer to other consumer BSs is presented by,

$$\mathcal{E}_i = B_i(t+1) + E_i^{h'}(t+1) - E_i^{c'}(t+1) - \omega_i. \quad (3)$$

Otherwise, if the BS is off-grid, i.e., BS $i \in N_{off}$, it behaves as a consumer BS and the amount of its energy demand is presented by,

$$d_i = E_i^{c'}(t+1) - B_i(t+1) - E_i^{h'}(t+1). \quad (4)$$

And if the BS is on-grid, i.e., BS $i \in N_{on}$, it purchases energy from the electrical grid and the amount of its purchased

energy is denoted by,

$$G_i(t+1) = E_i^{c'}(t+1) - B_i(t+1) - E_i^{h'}(t+1). \quad (5)$$

B. ENERGY HARVESTING PREDICTION MODEL

In this section, we present a Deep-Q-Learning (DQL) based approach for the prediction of the harvested energy of a particular BS $i \in N$. Exponentially-weighted moving average (EWMA) is a well-known and most used energy prediction algorithm which assumes that the harvested energy of a certain time slot shows a similarity with the exact time slot of the previous days [32]. In EWMA, the harvested energy for a particular time slot is predicted as a weighted average of the historical data of harvesting energy and prediction during the same time slot from the previous days. The weights are exponential which results in gradually decreasing weights for the older data. The predicted harvested energy for BS $i \in N$ at time slot t of the day d according to EWMA is represented by,

$$\overline{E_i}(d, t) = \beta \times \overline{E_i}(d-1, t) + (1 - \beta) \times E_i^h(d-1, t), \quad (6)$$

where, $\overline{E_i}(d-1, t)$ is the predicted energy at time slot t of the previous day according to EWMA, $E_i^h(d-1, t)$ is the actual harvested energy at time slot t of the previous day, β is the weight factor and $\beta \in [0, 1]$. EWMA exhibits a good accuracy of prediction in the areas where the weather remains stable. However, it will produce more prediction errors in case of frequently changing weather. The concept of EWMA combined with the consideration of current weather conditions, i.e., how the weather is changing in the most recent time slots is essential for frequent weather changing scenarios [33]. For ensuring this, we consider the energy generation information of n recent previous time slots. The modified equation of predicted energy of BS $i \in N$ for time slot t is presented by,

$$E_i^{h'}(t) = \overline{E_i}(d, t) \times (1 + \eta), \quad (7)$$

where, $\overline{E_i}(d, t)$ is the amount of predicted energy at time slot t of the present day according to EWMA and η is the weather changing ratio which represents the average harvested energy increasing/decreasing ratio of the previous slots. The value of η can be positive if there occurs an increase of harvested energy in the previous time slots on average, otherwise, negative. Now, we will describe the attributes of our deep Q-network (DQN) and its training process in the subsections.

1) SYSTEM STATE SPACE

As in our system model, each day is divided into slots and the length of each time slot is τ. There is a total of T repetitive time slots in a day. Each state in DQN is determined by the corresponding time slot and the actual harvested energy of BS $i \in N$ during n previous time slots. System state at time slot t is represented by $s_t = \{t, E_i^h(t-n), E_i^h(t-n+1), \ldots, E_i^h(t-1)\}$. Therefore, the system state space of our DQN system can be represented by,

$$\mathcal{S} = \{s_1, s_2, s_3, \ldots, s_T\}. \quad (8)$$

2) ACTION SPACE

The harvested energy is estimated according to Eq. (7) in our prediction algorithm. We have considered three different ways to calculate the weather changing ratio, η, and these three different ways represent our possible actions from a state in the DQN system. Therefore, the action space can be presented by,

$$\mathcal{A} = \{a_1, a_2, a_3\}, \quad (9)$$

where, action a_1 is calculating η considering all of the n previous time slots. In action a_2, we choose those time slots from n previous slots whose energy harvesting value falls in a threshold range. And, in action a_3, we choose those time slots whose energy harvesting value deviates by a certain allowable amount from the mean value of the harvested energy of n previous time slots.

According to action a_1, the weather-changing ratio is calculated by the following equation,

$$\eta = \frac{\sum_{t'=t-n}^{t-1} P^e(t') \times R(t') \times t'}{\sum_{t'=t-n}^{t-1} t'}, \quad (10)$$

where, $P^e(t')$ is the prediction error ratio at time slot t', $R(t')$ is the prediction reliability level of time slot t' which we determine using DQL algorithm. Here, we calculate the average weather change ratio in terms of the fluctuation of actual harvested energy from the predicted energy in the recent time slots. For giving more priority to the recent time slots, the slot number is multiplied as a weight. The prediction error ratio for any time slot t is presented by,

$$P^e(t) = \frac{E_i^h(t) - E_i^{h'}(t)}{E_i^{h'}(t)}, \quad (11)$$

where, $E_i^h(t)$ is the actual harvested energy and $E_i^{h'}(t)$ is the predicted harvested energy of BS $i \in N$ at time slot t.

According to action a_2, the weather-changing ratio is presented by the following equation,

$$\eta = \frac{\sum_{t'=t-n}^{t-1} P^e(t') \times R(t') \times t' \times I_{\{\nabla_{min}^t \leq E_i^h(t') \leq \nabla_{max}^t\}}}{\sum_{t'=t-n}^{t-1} t' \times I_{\{\nabla_{min}^t \leq E_i^h(t') \leq \nabla_{max}^t\}}}. \quad (12)$$

Here, $I_{\{.\}}$ is an indicator function whose value is 1 when the event {.} is true, otherwise, 0. Here, we are considering those previous time slots whose actual harvested energy falls in a threshold range $[\nabla_{min}^t, \nabla_{max}^t]$. In this paper, we have calculated these threshold values by taking two standard deviation ranges from the mean value of harvested energy of the corresponding BS during the particular time slot of the previous 30 days.

Finally, according to action a_3, the weather-changing ratio can be calculated by the following equation,

$$\eta = \frac{\sum_{t'=t-n}^{t-1} P^e(t') \times R(t') \times t' \times I_{\{|E_i^h(t')-\mu_h| \leq \lambda\}}}{\sum_{t'=t-n}^{t-1} t' \times I_{\{|E_i^h(t')-\mu_h| \leq \lambda\}}}, \quad (13)$$

where, μ_h is the statistical mean of the harvested energy of previous n time slots of BS $i \in N$. λ is the allowable difference of harvested energy of a particular time slot from the mean value. The statistical mean of the harvested energy of the recent previous n time slots is presented by,

$$\mu_h = \frac{1}{n} \sum_{t'=t-n}^{t-1} E_i^h(t'). \quad (14)$$

Here, we take the sum of harvested energy of n previous time slots and divide it by n to derive the mean value.

3) REWARD
For energy prediction at each time slot, there will be a reward based on the prediction accuracy. This will help the agent learn and adjust its action accordingly in order to derive more precision in predicting the harvested energy. The reward for taking an action $a \in \mathcal{A}$ from state s_t can be represented by,

$$r(s_t, a) = \begin{cases} 1, & \text{if } P^e(t) < P^e_{max} \\ -1, & \text{otherwise} \end{cases}, \quad (15)$$

where, P_{max} is the maximum allowable prediction error ratio in the system. If the prediction error ratio for a certain time slot becomes less than or equal to the maximum allowable prediction error ratio, then the agent will obtain a reward of 1. Otherwise, it will obtain a reward of -1.

4) TRAINING OF THE DQN
For accelerating the learning process of the traditional Q-learning method, we propose a deep Q-learning-based approach which is also known as a deep Q-network (DQN). In DQN, a deep neural network (DNN) is used to estimate the Q-function values for different state-action pairs. As DNN-based Q-Learning faces the problem of instability, we integrate the experience replay technique with Deep Q-Learning [34]. If the DNN learns from the sequential samples, the learning process often becomes inefficient as there is a chance of high correlation among the samples. And in the experience replay technique, a particular experience is potentially used for many weight updates that promote greater data efficiency.

Q-function value in reinforcement learning represents the quality of choosing an action from a particular state and helps in policy making. In our Deep Q-Learning method, there is a DNN called training network with weight vector, θ, and the system learns by updating the weight of this network. We represent the reliability level of a certain time slot as the predicted maximum Q-value by passing the corresponding

---

Algorithm 1 Training of the Deep Q-Network at Each BS

Input: State space $\mathcal{S}$, action space $\mathcal{A}$.
Output: Trained DQN

1: Initialize replay memory $\mathcal{D}$ with capacity N
2: Initialize the training network with state-action function Q and random weights θ
3: Initialize the target network with the state-action function $\hat{Q}$ and weights $\hat{\theta} = \theta$
4: $steps = 0$
5: **for** episode=1,M **do**
6:   **for** t = 1, T **do**
7:     **if** probability $\geq \epsilon$ **then**
8:       a_t = Random $(\mathcal{A})$
9:     **else**
10:       $a_t = \arg\max_{a \in \mathcal{A}} Q(s_t, a; \theta)$
11:     **end if**
12:     $\varrho_t = r(s_t, a_t)$
13:     $\mathcal{D} \leftarrow (s_t, a_t, \varrho_t, s_{t+1})$
14:     $e' \leftarrow$ Random-minibatch$(\mathcal{D})$
15:     **for all** $e_i \in e'$ **do**
16:       $\mathcal{Z}_i = \varrho_i + \gamma \times \max_{a' \in \mathcal{A}} \hat{Q}(s_{i+1}, a'; \hat{\theta})$
17:       $L(\theta) = \{\mathcal{Z}_i - Q(s_t, a_t; \theta)\}^2$
18:       Perform Gradient Descent on $L(\theta)$ with respect to θ
19:       $steps = steps + 1$
20:       **if** $steps == c$ **then**
21:         set $\hat{\theta} = \theta$
22:         $steps = 0$
23:       **end if**
24:     **end for**
25:   **end for**
26: **end for**
27: **return** Trained DQN

---

state to the training network. The prediction reliability at time slot t, denoted by $R(t)$, can be presented by,

$$R(t) = \max_{a \in \mathcal{A}} Q(s_t, a; \theta), \quad (16)$$

where, the reliability level of a time slot t is equal to the maximum Q-value we obtain from the training network for the corresponding state s_t.

In our deep Q-learning model, we use a separate neural network with weight vector $\hat{\theta}$ to find out the target Q-values. The target Q-value for taking an action $a_t \in \mathcal{A}$ from state s_t in Deep Q-Learning is presented by the following equation,

$$\mathcal{Z} = r(s_t, a_t) + \gamma \times \max_{a' \in \mathcal{A}} \hat{Q}(s_{t+1}, a'; \hat{\theta}). \quad (17)$$

Here, γ is the discount factor. Here, $r(s_t, a_t)$ the immediate reward of choosing action a_t from state s_t and s_{t+1} is the next state. The maximum Q-value from the next state is multiplied by the discount factor and this value is added with the immediate reward. The Q-value for the next state s_{t+1} is calculated by passing the state to the target network with weight vector $\hat{\theta}$. The actual predicted Q-value, denoted by

$Q(s_t, a_t; \theta)$, is obtained by passing the state s_t to training network with weight vector θ. In DQN, we try to minimize the loss function in terms of the difference between the predicted Q-value and the target Q-value. The loss function can be presented as,

$$L(\theta) = \{\mathcal{Z} - Q(s_t, a_t; \theta)\}^2, \tag{18}$$

where, the loss function is calculated as the square value of the difference between the target Q-value and the predicted Q-value. For updating the weight parameter of the training network, we perform a gradient descent on the loss function. To obtain the gradient, we differentiate the loss function with respect to weight parameter θ and it can be represented by,

$$\frac{\partial L(\theta)}{\partial \theta} = \{r(s_t, a_t) + \gamma \times \max_{a' \in \mathcal{A}} \hat{Q}(s_{t+1}, a'; \hat{\theta}) \\ - Q(s_t, a_t; \theta)\} \frac{\partial Q(s_t, a_t; \theta)}{\partial \theta}. \tag{19}$$

Gradient Descent [34] is then performed to update the weights of the training network to minimize the loss function. The training process of the DQN is presented by algorithm 1. At first, we initialize the replay memory $\mathcal{D}$, the training network with random weights θ and we initialize the target network with weight parameter, $\hat{\theta} = \theta$. For each state in each episode, with probability ϵ, we choose a random action, a_t. Otherwise, we select the action that gives us the maximum Q-value for the state in the training network. Then, we perform the action from the state, observe the reward and store the transition s_t, a_t, the reward we get, and the next state into the replay memory. After this, we sample a random minibatch of transition from the replay memory. For each transition, we calculate the target Q-value, and the loss function and then perform gradient descent on the weight parameter of the training network θ. And after each c update, we clone the training network into the target network by setting $\hat{\theta} = \theta$.

5) COMPLEXITY ANALYSIS
In this section, we have analyzed the time complexity of Algorithm 1. Let, $|s_t|$ denote the size of the state space at time slot τ. Therefore, the size of the state space of our developed system is $|s_t| \times T$. Let, $|\mathcal{A}|$ denote the size of the action space. Therefore, the required time complexity of line 2 is $O(|s_t| \times T \times |\mathcal{A}|)$. In line 10, we have selected an action from the action set $\mathcal{A}$, and the required time complexity of line 10 is $O(|\mathcal{A}|)$. In line 14, we have selected mini-batch e' and the size of the mini-batch is $|e'|$. Lines 15 to 24 are enclosed in a loop, which iterates at most $|e'|$ times. Therefore, the worst case time complexity from lines 5 to 26 is $O(M \times T \times (|\mathcal{A}| + |e'|))$. As a result, the overall time complexity of Algorithm 1 is $O(|s_t| \times T \times |\mathcal{A}| + M \times T \times (|\mathcal{A}| + |e'|))$.

C. OPTIMAL ENERGY COOPERATION
In this section, we first formulate the optimal energy cooperation framework. The related calculations and operations are described in the subsequent sub-sections.

1) PROBLEM FORMULATION
Let, ζ_{ij} denote the normalized energy buying cost of consumer BS $i \in \mathcal{X}^c$ from source BS $j \in \mathcal{X}^s$ for the purchased energy. ϕ_{ij} is the energy loss factor due to the transfer of energy from source BS $j \in \mathcal{X}^s$ to consumer BS $i \in \mathcal{X}^c$ through the DC power line between them. In our objective function, we bring a trade-off between minimizing the energy buying cost of the BSs and minimizing the loss of energy due to energy transfer among the BSs introducing a priority factor, α. We impose weight α to energy buying cost minimization of the consumer BSs and $(1 - \alpha)$ to the minimization of loss of energy due to energy transfer among the BSs. The objective function is presented as follows,
Minimize :

$$W = \sum_{i \in \mathcal{X}^c} \sum_{j \in \mathcal{X}^s} \{\alpha \times \zeta_{ij} + (1 - \alpha) \times \phi_{ij}\} \tag{20}$$

Here, the value of weight factor α is in between $(0, 1]$), determined by the system administrator. The objective function is subject to the following constraints:

$$0 \le x_{ij} \le 1, \quad \forall i \in \mathcal{X}^c, \quad \forall j \in \mathcal{X}^s \tag{21}$$

$$\sum_{i \in \mathcal{X}^c} x_{ij} \le 1, \quad \forall j \in \mathcal{X}^s \tag{22}$$

$$\sum_{j \in \mathcal{X}^s} \mathcal{E}_{ij}^s \le d_i, \quad \forall i \in \mathcal{X}^c \tag{23}$$

$$\sum_{j \in \mathcal{X}^s} \mathcal{E}_{ij}^s \ge \rho_i \times d_i, \quad \forall i \in \mathcal{X}^c \tag{24}$$

$$\sigma_s <= \delta_m \tag{25}$$

Here, in constraint (21), x_{ij} is a energy sharing factor. The value of x_{ij} is in between $[0, 1]$ and it means the portion of available energy of source BS $j \in \mathcal{X}^s$ that is transferred to consumer BS $i \in \mathcal{X}^c$. This constraint means that a consumer BS can receive any fraction of the available energy of any source BS. Constraint (22) is the availability constraint. It denotes that, the total amount of energy a certain source BS transfers to the consumer BSs cannot exceed the available energy of that source BS. Constraint (23) and (24) together represent the demand constraint. Constraint (23) means that the total amount of energy a certain consumer BS receives from the source BSs cannot exceed its energy demand. Constraint (24) means that the minimum portion of energy demand set by a consumer BS must be fulfilled by deriving energy from the source BSs. Constraint (25) means that the standard deviation of the remaining sharable energy of the source BSs after all the energy transfers cannot exceed the maximum allowable value, δ_m, basically for maintaining a balance between them and increasing the network lifetime.

2) CALCULATION OF COST
The normalized cost consumer BS $i \in \mathcal{X}^c$ pays to source BS $j \in \mathcal{X}^s$ for energy transfer is presented by,

$$\zeta_{ij} = \frac{C_{ij}^u \times \mathcal{E}_{ij}^s}{C_{max}^u \times \mathcal{E}_{max}^u}, \quad \zeta_{ij} \in [0, 1], \tag{26}$$

where, C_{ij}^u is the cost per unit energy source BS $j \in \mathcal{X}^s$ demands from consumer BS $i \in \mathcal{X}^c$, $x_{ij} \in [0, 1]$ means the fraction of energy that is transferred from source BS $j \in \mathcal{X}^s$ to consumer BS $i \in \mathcal{X}^c$, C_{max}^u is the maximum cost per unit energy fixed by any source BS in the network and $\mathcal{E}_{max}^u$ is the maximum possible energy that can be transferred from any source BS to consumer BS in the network. The amount of energy source BS $j \in \mathcal{X}^s$ transfers to consumer BS $i \in \mathcal{X}^c$ is denoted by,

$$\mathcal{E}_{ij}^s = x_{ij} \times \mathcal{E}_j, \quad (27)$$

where, $x_{ij} \in [0, 1]$ means the fraction of energy that is transferred from source BS $j \in \mathcal{X}^s$ to consumer BS $i \in \mathcal{X}^c$ and $\mathcal{E}_j$ is the total amount surplus energy of source BS $i \in \mathcal{X}^c$ that it is ready to transfer to the consumer BSs.

3) CALCULATION OF ENERGY LOSS FACTOR

The energy loss factor, in case of transferring energy from source BS $j \in \mathcal{X}^s$ under the domain of energy router $k \in R$ to consumer BS $i \in \mathcal{X}^c$ through the sequence of DC power lines, is represented by,

$$\phi_{ij} = \frac{\mathcal{E}_{ij}^s - \mathcal{E'}_{ij}}{\mathcal{E}_{max}^l}, \quad \phi_{ij} \in [0, 1], \quad (28)$$

where, $\mathcal{E'}_{ij}$ is the amount of energy received by the consumer BS $i \in \mathcal{X}^c$ after transferring through the power link. And $\mathcal{E}_{ij}^l$ is the maximum amount of energy loss observed in the network.

4) COMPUTING STANDARD DEVIATION OF REMAINING SHARABLE ENERGY

The amount of remaining energy of source BS $j \in \mathcal{X}^s$ after the energy cooperation is defined by,

$$r_j = \mathcal{E}_j - \sum_{i \in \mathcal{X}^c} \mathcal{E}_{ij}^s. \quad (29)$$

Here, we are subtracting the total amount of energy transferred to the consumers from the sharable energy of the source bs $j \in \mathcal{X}^s$. The standard deviation of the remaining energy of the source BSs after performing energy sharing is presented by,

$$\sigma_s = \sqrt{\frac{\sum_{j \in \mathcal{X}^s}(r_j - \mu)^2}{s}}, \quad (30)$$

where, μ is the mean remaining energy of the source BSs and s is the total number of source BSs.

V. PERFORMANCE EVALUATION

In this section, we implement our proposed MCEL system and we present the comparison of the performance of our proposed system with other state-of-the-art works: CONV [13], MinCost [14] and GPs+MPC+CONV [30]. We have implemented our proposed energy harvesting prediction model using MATLAB [35]. To solve our optimal energy cooperation problem, we use CPLEX solver at NEOS optimization server [36] (2x Intel Xeon E5-2698 @ 2.3-GHz 569 CPU and 92-GB RAM).

TABLE 2. Simulation parameters.

Parameter	Value
Simulation area	$2000 \times 2000 \ m^2$
Number of energy routers	5
Weigh factor, α	0.6
Maximum energy buffer capacity of a BS	360 KJ
Length of each power link	50-200 m
Resistivity of cable	0.023 $\Omega mm^2/m$
Cross sectional area of cable	10 mm^2
Learning rate of the DQN	0.001
Discount factor, γ	0.9

A. SIMULATION ENVIRONMENT

For the experimental environment setup, we have considered a simulation area of $2000 \times 2000 \ m^2$, 5 energy routers, and 5 to 25 consumer BSs under each energy router. The energy harvesting data and traffic load data were achieved randomly following a practical scenario according to [13]. We have considered 24 time slots in a day and the length of each time slot is one hour. For the deep Q-Network, we have considered three hidden layers each having 64 nodes. The ReLU activation function is used for the hidden layers and the Softmax activation function is used for the output layer. The learning rate of the deep Q-Network is set to 0.001. We have plotted the data points in each graph after taking the average result from 30 runs. The value and the range of performance parameters are presented in Table 2.

B. PERFORMANCE METRICS

For comparing the performance of our proposed system, the following performance metrics have been considered:

- The normalized cost defines the normalized value of total energy buying costs of the consumer BSs to fulfill the services of the users.
- The normalized loss of energy is defined by the normalized value of the total loss of energy occurring from the energy transfer among the BSs through the DC power lines.
- Integrated performance is defined by the integrated metrics of energy buying costs and loss of energy while transferring through the DC power lines. It can be presented by the following equation:

$$\text{Integrated Performance} = 1 - 0.5 \times (\vartheta + \kappa), \quad (31)$$

where, ϑ is the normalized energy buying cost of the BSs and κ is the normalized loss of energy due to energy transfer among the BSs. The value of the integrated performance is between [0, 1]. The higher the value of integrated performance, the better the performance of the system.

- User service disruption is defined as the percentage of user service requests that can not be fulfilled by the BSs due to a shortage in the available energy to serve them. The lower the percentage of user service disruption, the higher the performance of the system.
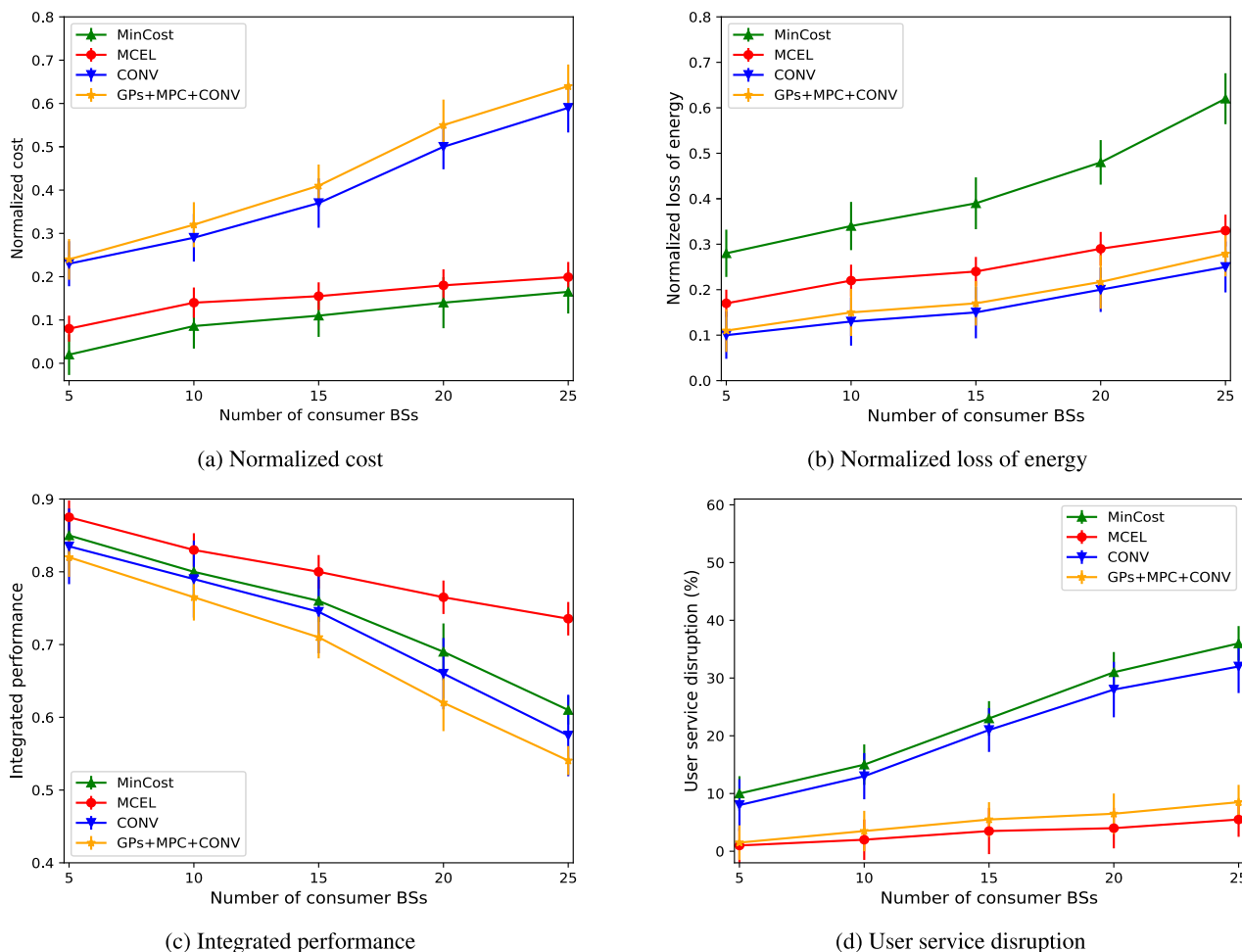
(a) Normalized cost

(b) Normalized loss of energy

(c) Integrated performance

(d) User service disruption

FIGURE 2. Impacts of varying the number of consumer BSs.

C. SIMULATION RESULT

In this section, we present the performance of our proposed system, MCEL, by varying the number of consumer BSs, the number of average users per consumer BSs and the number of source BSs.

1) IMPACT OF A VARYING NUMBER OF CONSUMER BSs

In this experiment, we have varied the number of consumer BSs keeping the number of source BSs at 15.

Fig. 2(a) represents the impact of varying the number of consumer BSs on the normalized energy buying cost of the consumer BSs. From the graph, we can see that, with the increase in the number of consumer BSs, the normalized energy buying cost increases. As MinCost only focuses on the minimization of energy buying cost of the consumer BSs, it is performing best in this case. However, it doesn't consider the loss of energy due to energy transfer among the BSs. On the other hand, CONV and GPs+MPC+CONV do not focus on cost minimization. That's why their performance is poor in this case. Here, our proposed system MCEL brings a trade-off between the minimization of the energy buying cost

and the loss of energy and it performs better than CONV and GPs+MPC+CONV.

Fig. 2(b) depicts the impact of varying the number of consumer BSs on the normalized loss of energy due to energy transfer. It can be realized from the graph that, with the increasing number of consumer BSs, the normalized loss of energy increases. As CONV and GPs+MPC+CONV only focus on minimizing the loss of energy, they are performing better in this case. However, they do not consider the minimization of the energy buying cost of the BSs. On the other hand, MinCost is performing worst in this case as it does not consider the loss of energy at all. Here, our proposed system MCEL is performing better than MinCost.

Fig. 2(c) depicts the impact of varying the number of consumer BSs on the integrated performance of the system in terms of normalized cost and loss of energy. It can be observed from the graph that, with the increase in the number of consumer BSs, the integrated performance decreases. In this case, our proposed system MCEL outperforms other solutions because the other three solutions only focus either on the minimization of energy buying cost or the loss of
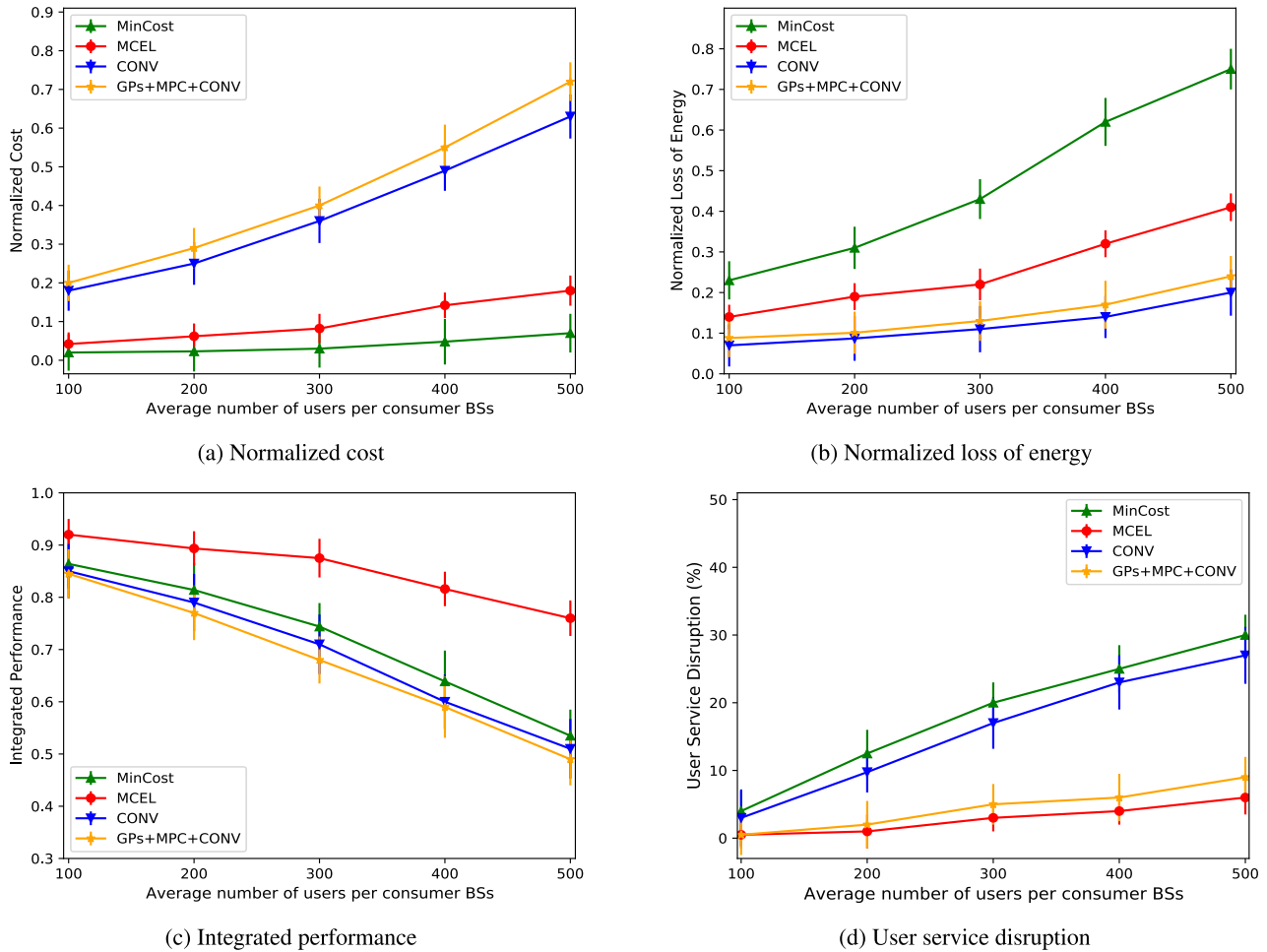
(a) Normalized cost

(b) Normalized loss of energy

(c) Integrated performance

(d) User service disruption

FIGURE 3. Impacts of varying the average number of users per consumer BSs.

energy due to transfer. When they try to minimize one, the other increases. As our proposed solution brings a trade-off between these two conflicting objectives, it is performing much better than the other three approaches.

Fig. 2(d) depicts the impact of varying the number of consumer BSs on user service disruption. As MCEL and GPs+MPC+CONV consider the prediction of harvested energy of the BSs and traffic load of the BSs, user service disruption is less for these two solutions than MinCost and CONV. MinCost and CONV do not consider the prediction of harvested energy and traffic load of the BSs. Rather they determine the energy demand of the BSs concerning certain threshold energy levels. That's why their performance is poor in this case and our proposed solution MCEL outperforms the other three approaches.

2) IMPACT OF A VARYING NUMBER OF AVERAGE USERS PER CONSUMER BSs

In this experiment, we have varied the number of average users per consumer BSs keeping the number of source BSs and consumer BSs at 15 and 20, respectively.

Fig. 3(a) presents the impact of the varying average numbers of users of the consumer BSs on the normalized energy buying cost of the consumer BSs. It can be observed from the graph that, with the increase in the number of average users, the energy buying cost of the consumer BSs increases as it creates increasing energy demands in the consumer BSs. As discussed earlier, MinCost only focuses on the minimization of the energy buying cost of the consumer BSs, it tends to achieve it by taking energy mostly from the source BSs of the same operators. Therefore, MinCost is performing better in this case. On the other hand, CONV and GPs+MPC+CONV do not consider the minimization of energy buying cost of the consumer BSs. For this reason, they are performing worse in this case. And our proposed system MCEL is performing better than CONV and GPs+MPC+CONV as it tries to bring a trade-off between the two conflicting objectives.

Fig. 3(b) depicts the impact of varying the average number of users of the consumer BSs on the normalized loss of energy. We can see from the graph that, with the increase in the number of average users, the loss of energy due to transfer increases since the energy demands of the consumer BSs
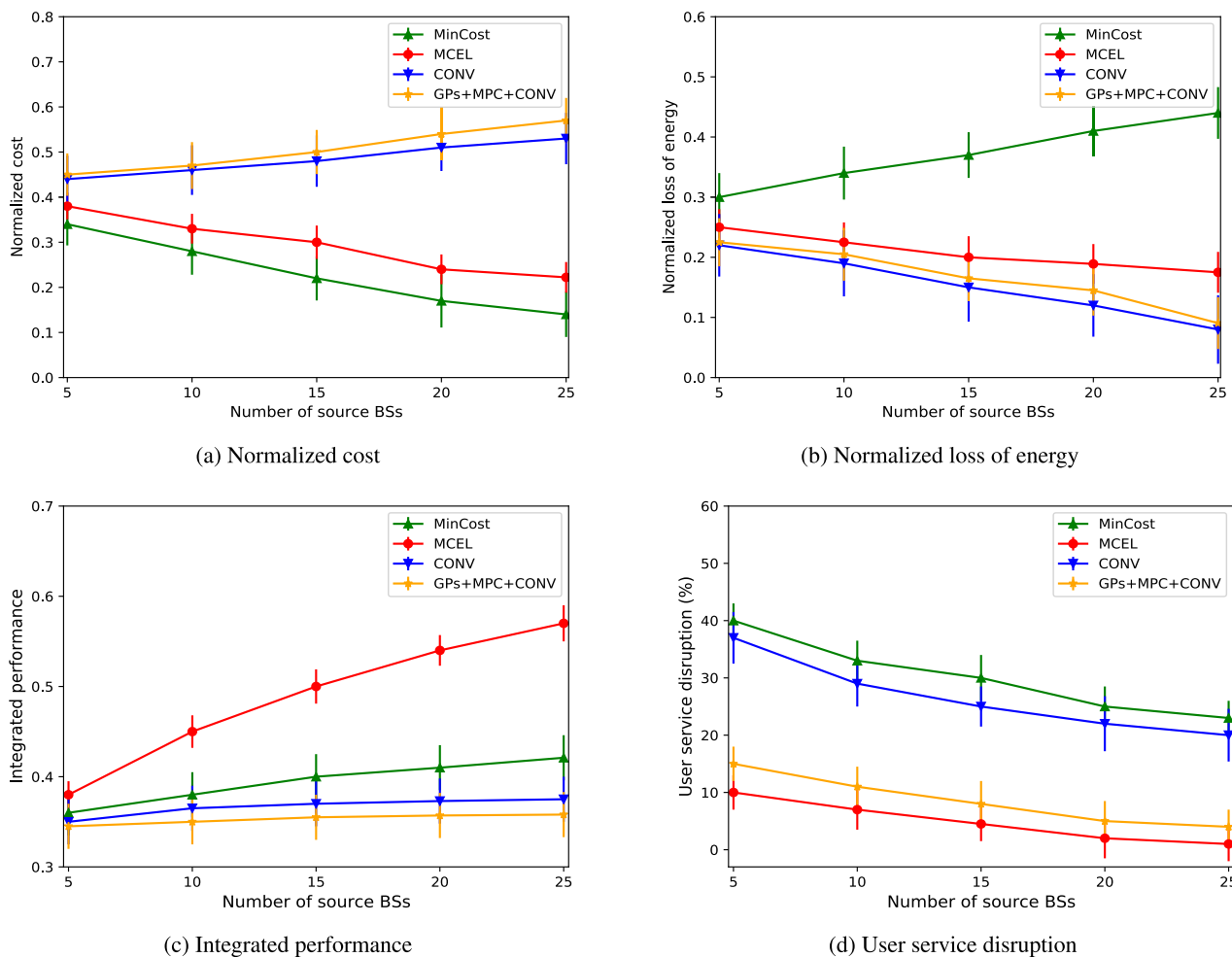
(a) Normalized cost

(b) Normalized loss of energy

(c) Integrated performance

(d) User service disruption

FIGURE 4. Impacts of varying the number of source BSs.

increase. As a result, the transfer of energy from the source of consumer BSs and the loss of energy due to energy transfer increase. As CONV and GPs+MPC+CONV only minimize the loss of energy due to energy transfer, they are performing better in this case. On the other hand, as MinCost does not consider minimizing the loss of energy due to energy transfer among the BSs, its performance is poor in this case. Here, our proposed solution MCEL is performing much better than MinCost.

Fig. 3(c) represents the impact of varying the average number of users of the consumer BSs on the integrated performance in terms of normalized cost and normalized loss of energy. It can be observed from the graph that, with the increasing number of users per consumer BSs, the integrated performance decreases for all of the four approaches as the normalized energy buying cost and loss of energy both increase. As our proposed solution brings a trade-off between these two conflicting objectives, i.e., minimization of the energy buying cost and the loss of energy, it outperforms the other three approaches.

Fig. 3(d) represents the impact of varying the average number of users of the consumer BSs on user service disruption. As MinCost and CONV do not perform the prediction of harvested energy and traffic load to determine the energy demand of consumer BSs for the next time slot, they are performing worse than MCEL and GPs+MPC+CONV. MinCost and CONV determine the energy demand of the consumer BSs based on pre-determined threshold energy levels and this is not a viable solution. When the user is less, they can be served mostly by the available energy of the consumer BSs. That's why the four approaches are performing near to each other for fewer users. But with the increasing number of users, MinCost and CONV perform worse than MCEL and GPs+MPC+CONV, which exploit the prediction of harvested energy and traffic load of the BSs. Here, our proposed system MCEL outperforms all the other three approaches.

3) IMPACT OF VARYING NUMBER OF SOURCE BSs
In this experiment, we have varied the number of source BSs keeping the number of consumer BSs at 10.

Fig. 4(a) represents the impact of a varying number of source BSs on the normalized energy buying cost of the consumer BSs. It can be observed from the graphs that, in the case of MinCost, with the increase in the number of source BSs, the energy buying cost of the consumer BSs decreases. That is because, with the increasing number of source BSs, the different options of selecting a source BS for a consumer BS increase, and MinCost further tries to minimize the energy buying cost of the consumer BSs by selecting the source BSs of the same operators. Therefore, MinCost is performing better in this case. On the other hand, with the increasing number of source BSs, CONV and GPs+MPC+CONV further try to minimize the loss of energy by selecting the nearby source BSs. However, usually nearby BSs appear to be of different operators and for this reason, the energy buying cost of the consumer BSs increases. As our proposed solution MCEL minimizes both the energy buying cost and loss of energy by bringing a trade-off, with the increased number of source BSs, it selects source BSs in such a way that, both the normalized cost and loss of energy decrease. Here, MCEL is performing better than CONV and GPs+MPC+CONV in this case.

Fig. 4(b) represents the impact of varying the number of source BSs on the normalized loss of energy due to energy transfer among the BSs. We can observe from the graph that, with the increased number of source BSs, CONV and GPs+MPC+CONV further try to minimize the loss of energy while transferring by choosing nearby BSs. CONV and GPs+MPC+CONV only focus on the minimization of loss of energy. Therefore, they are performing better than MCEL and MinCost in this case. On the other hand, with the increased number of source BSs, MinCost further tries to minimize energy buying costs by giving priority to choosing the source BSs of the same operators. Since BSs of the same operator are more distant than the BSs of different operators, the loss of energy due to energy transfer increases. That's why MinCost is performing worst in this case. Here, our proposed solution MCEL is performing much better than MinCost.

Fig. 4(c) represents the impact of varying the number of source BSs on the integrated performance in terms of the normalized energy buying cost and loss of energy due to energy transfer. It can be observed from the graph that, with the increasing number of source BSs, the integrated performance increases. As our proposed solution tries to bring a trade-off between minimizing energy buying costs and loss of energy, it outperforms the other three approaches in this case.

Fig. 4(d) represents the impact of varying the number of source BSs on user service disruption. We can observe from the graph that, with the increase in the number of source BSs, the user service disruption decreases. Here, MCEL and GPs+MPC+CONV are performing better than MinCost and CONV as they consider the prediction of harvested energy and load of the BSs for calculating the energy demand of the consumer BSs. Here, our proposed solution MCEL outperforms the other three approaches in this case.

VI. CONCLUSION

In summary, in this work, we proposed an optimal energy cooperation framework for energy sharing among the BSs in multi-operator cellular networks. We ensured a trade-off between two conflicting objectives- minimization of the energy buying cost and loss of energy. The inclusion of the prediction of future harvested energy and energy consumption profile of the BSs in the framework helped it to achieve significant performance impact. The simulation results showed that our proposed system MCEL achieved significant improvement in terms of service disruption reduction and integrated performance of energy loss reduction and cost as high as 15% and 20%, respectively. As a future work, the joint optimization of energy cooperation along with BS state scheduling in a multi-operator network environment will be an avenue for achieving better performance in 5G concerning ever-increasing traffic and data-intensive applications.

REFERENCES

[1] *ITU World Telecommunication/ICT Indicators Database 2019*, Int. Telecommun. Union, Geneva, Switzerland, Dec. 2019.

[2] P. Roy, S. Sarker, M. A. Razzaque, M. Mamun-or-Rashid, M. M. Hassan, and G. Fortino, "Distributed task allocation in mobile device cloud exploiting federated learning and subjective logic," *J. Syst. Archit.*, vol. 113, Feb. 2021, Art. no. 101972.

[3] P. Roy, A. Tahsin, S. Sarker, T. Adhikary, M. A. Razzaque, and M. M. Hassand, "User mobility and quality-of-experience aware placement of virtual network functions in 5G," *Comput. Commun.*, vol. 150, pp. 367–377, Jan. 2020.

[4] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, Nov. 2011.

[5] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.

[6] A. M. Aris and B. Shabani, "Sustainable power supply solutions for off-grid base stations," *Energies*, vol. 8, no. 10, pp. 10904–10941, Sep. 2015.

[7] U. K. Dutta, M. A. Razzaque, M. Abdullah Al-Wadud, M. S. Islam, M. S. Hossain, and B. B. Gupta, "Self-adaptive scheduling of base transceiver stations in green 5G networks," *IEEE Access*, vol. 6, pp. 7958–7969, 2018.

[8] J. Ferdous, M. P. Mollah, M. A. Razzaque, M. M. Hassan, A. Alamri, G. Fortino, and M. Zhou, "Optimal dynamic pricing for trading-off user utility and operator profit in smart grid," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 2, pp. 455–467, Feb. 2017.

[9] A. J. Williams, M. F. Torquato, I. M. Cameron, A. A. Fahmy, and J. Sienz, "Survey of energy harvesting technologies for wireless sensor networks," *IEEE Access*, vol. 9, pp. 77493–77510, 2021.

[10] J. Zhang, B. Zhang, J. Liu, and Z. Han, "A joint offloading and energy cooperation scheme for edge computing networks," in *Proc. IEEE Int. Conf. Commun.*, Aug. 2022, pp. 5537–5542.

[11] A. Balakrishnan, S. De, and L.-C. Wang, "Energy sharing based cooperative dual-powered green cellular networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[12] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[13] A. Fernandez Gambin and M. Rossi, "Energy cooperation for sustainable base station deployments: Principles and algorithms," in *Proc. IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–7.

[14] M. J. Farooq, H. Ghazzai, A. Kadri, H. ElSawy, and M.-S. Alouini, "Energy sharing framework for microgrid-powered cellular base stations," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[15] J. Ma, L. Song, and Y. Li, "Optimal power dispatching for local area packetized power network," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4765–4776, Sep. 2018.

[16] D. Zeller, M. Olsson, O. Blume, A. Fehske, D. Ferling, W. Tomaselli, and I. Gódor, "Sustainable wireless broadband access to the future internet-the earth project," in *The Future Internet Assembly*. Cham, Switzerland: Springer, 2013, pp. 249–271.

[17] R. Bolla, R. Bruschi, F. Davoli, L. Di Gregorio, L. Giacomello, C. Lombardo, G. Parladori, N. Strugo, and A. Zafeiropoulos, "The low energy consumption networks (ECONET) project," in *Proc. Sustain. Internet ICT Sustainability (SustainIT)*, 2012, pp. 1–5.

[18] M. A. Marsan, S. Buzzi, L. Chiaraviglio, M. Meo, C. Guerrero, F. Idzikowski, Y. Ye, and J. L. Vizcaíno, "TREND: Toward real energy-efficient network design," in *Proc. Sustain. Internet ICT Sustainability (SustainIT)*, 2012, pp. 1–6.

[19] J. Leithon, S. Sun, and T. Joon Lim, "Energy management strategies for base stations powered by the smart grid," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 2635–2640.

[20] S. Bu, F. R. Yu, M. Y.-K. Chua, and X. P. Liu, "When the smart grid meets energy-efficient communications: Green wireless cellular networks powered by the smart grid," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 3014–3024, Aug. 2012.

[21] J. Leithon, T. J. Lim, and S. Sun, "Energy exchange among base stations in a cellular network through the smart grid," *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 4036–4041.

[22] A. Balakrishnan, S. De, and L.-C. Wang, "Networked energy cooperation in dual powered green cellular networks," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6977–6991, Oct. 2022.

[23] S. Euttamarajah, Y. H. Ng, and C. K. Tan, "Energy-efficient joint power allocation and energy cooperation for hybrid-powered comp-enabled Het-Net," *IEEE Access*, vol. 8, pp. 29169–29175, 2020.

[24] F. Yin, M. Zeng, Z. Zhang, and D. Liu, "Coded caching for smart grid enabled HetNets with resource allocation and energy cooperation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12058–12071, Oct. 2020.

[25] H.-S. Lee and J.-W. Lee, "Adaptive traffic management and energy cooperation in renewable-energy-powered cellular networks," *IEEE Syst. J.*, vol. 14, no. 1, pp. 132–143, Mar. 2019.

[26] Y.-K. Chia, S. Sun, and R. Zhang, "Energy cooperation in cellular networks with renewable powered base stations," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6996–7010, Dec. 2014.

[27] J. Xu and R. Zhang, "CoMP meets smart grid: A new communication and energy cooperation paradigm," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2476–2488, Jun. 2015.

[28] H. Sugiyama, "Packet switched power network with decentralized control based on synchronized QoS routing," in *ICT&Applications Collocated Events*. USA: International Institute of Informatics and Systemics, 2012, pp. 147–152.

[29] L. Bonati, A. F. Gambin, and M. Rossi, "Wireless power transfer under the spotlight: Charging terminals amid dense cellular networks," in *Proc. IEEE 18th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoW-MoM)*, Jun. 2017, pp. 1–9.

[30] A. Fernandez Gambin, M. Scalabrin, and M. Rossi, "Online power management strategies for energy harvesting mobile networks," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 721–738, Sep. 2019.

[31] B. S. Shawel, T. T. Debella, G. Tesfaye, Y. Y. Tefera, and D. H. Woldegebreal, "Hybrid prediction model for mobile data traffic: A cluster-level approach," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[32] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embedded Comput. Syst.*, vol. 6, no. 4, p. 32, Sep. 2007.

[33] S. Kosunalp, "A new energy prediction algorithm for energy-harvesting wireless sensor networks with Q-learning," *IEEE Access*, vol. 4, pp. 5755–5763, 2016.

[34] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, and M. G. Bellemare, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, 2015.

[35] *Matlab Version R2020A*, I Mathworks, Natick, MA, USA, 2020.

[36] *NEOS Server: State-of-the-Art Solvers for Numerical Optimization*. Accessed: Aug. 1, 2022. [Online]. Available: http://www.neos-server.org/neos/

• • •