MDPI

Article

Coordinated Multi-Agent Deep Reinforcement Learning for Energy-Aware UAV-Based Big-Data Platforms

Soyi Jung [1] , **Won Joon Yun** [1] , **Joongheon Kim** [1,*] and **Jae-Hyun Kim** [2,*]

1 School of Electrical Engineering, Korea University, Seoul 02841, Korea; jungsoyi@korea.ac.kr (S.J.); ywjoon95@korea.ac.kr (W.J.Y.)
2 Department of Electrical and Computer Engineering, Ajou University, Suwon 16499, Korea
* Correspondence: joongheon@korea.ac.kr (J.K.); jkim@ajou.ac.kr (J.-H.K.); Tel.: +82-2-3290-3223 (J.K.); +82-31-219-2474 (J.-H.K.)

Abstract: This paper proposes a novel coordinated multi-agent deep reinforcement learning (MADRL) algorithm for energy sharing among multiple unmanned aerial vehicles (UAVs) in order to conduct big-data processing in a distributed manner. For realizing UAV-assisted aerial surveillance or flexible mobile cellular services, robust wireless charging mechanisms are essential for delivering energy sources from charging towers (i.e., charging infrastructure) to their associated UAVs for seamless operations of autonomous UAVs in the sky. In order to actively and intelligently manage the energy resources in charging towers, a MADRL-based coordinated energy management system is desired and proposed for energy resource sharing among charging towers. When the required energy for charging UAVs is not enough in charging towers, the energy purchase from utility company (i.e., energy source provider in local energy market) is desired, which takes high costs. Therefore, the main objective of our proposed coordinated MADRL-based energy sharing learning algorithm is minimizing energy purchase from external utility companies to minimize system-operational costs. Finally, our performance evaluation results verify that the proposed coordinated MADRL-based algorithm achieves desired performance improvements.

Keywords: big-data processing; multi-agent deep reinforcement learning; deep learning; smart grid; unmanned aerial vehicle (UAV)

1. Introduction

Modern technical advances in next-generation network and communication infrastructure enable reliable management and organization by utilizing mobile computing platforms, e.g., autonomous unmanned aerial vehicles (UAVs) [1–13]. Even though autonomous UAVs are considered major components in next-generation network design and implementation, it has several research challenges [14,15]. Among the research challenges, one of major problems is energy efficiency in power-hungry UAV platforms. Therefore, energy-efficient algorithms are obviously and definitely desired in UAV-based mobile communications and networks. In order to realize energy-aware reliable and robust UAV-based network design and implementation, the active use of charging infrastructure, such as charging towers with wireless power transfer technologies [16,17], is widely considered and discussed [1,11]. According to the fact that the charging infrastructure (including charging towers) are ground-mounted and AC-powered, the infrastructure gathers energy/power sources without strict limitations. Furthermore, the charging towers can (i) share their own energy resources in order to provide reliable and efficient energy resources among them or (ii) purchase energy resources from their associated utility company (also known as *external local energy market*) [18–22]. The dynamic active energy sharing sequential decision control process is essentially required for this given problem because the energy/power prices are determined based on auction-based economic theory in the local energy market [23]. Lastly, it is obvious that the charging towers are not only for energy distributors but also

for intelligent dynamic energy sharing traders, using deep learning computation. Thus, it it essential to use high-performance computing resources [24,25].

The proposed coordinated multi-agent deep reinforcement learning (DRL) (MADRL)-based autonomous and intelligent energy sharing in order to minimize energy purchases from the local energy market for minimizing/optimizing system-wide operational costs works with two tasks, as follows:

- The proposed algorithm determines the amount of energy sources purchased, where the corresponding prices can be dynamically updated depending on the energy-consuming patterns and auction-based economy theory for a local energy market. Note that the main objective of the proposed algorithm is to minimize this purchase price, which is also called the system-wide operational cost.
- The proposed algorithm shares energy resources at charging towers via coordinated MADRL-based cooperation among the charging towers.

For design and implementation of the proposed coordinated MADRL-based energy resource sharing learning in this paper, the considered charging towers can be considered MADRL agents and the agents collaboratively and coordinately work for autonomous and intelligent energy resource sharing under situations of time-varying unexpected observations. Among various MADRL-based algorithms, the proposed MADRL algorithm is designed fundamentally based on *communications neural network (CommNet)*, which is one of the well-known MADRL-based algorithms that obtains performance improvements via multi-agent intelligence coordination [26]. Furthermore, this proposed coordinated MADRL/CommNet-based algorithm is beneficial especially for big-data processing applications due to the fact that the processing requires a lot of computation resources within power-and-computation limited UAV platforms [27,28]. Therefore, efficient, active, and autonomous energy sharing mechanisms are essentially required for charging multi-UAV platforms.

Therefore, the novelties and contributions of our proposed MADRL-based energy resource sharing learning can be summarized and itemized as follows.

- *Joint scheduling:* The proposed scheduling in this paper is not only for the matching between UAVs and charging towers but also for charging energy allocation decisions.
- *DRL-based intelligent and autonomous energy management:* The proposed algorithm can dynamically and autonomously control energy sharing among charging towers based on DRL-based algorithms.
- *Multi-agent DRL computation:* Lastly, the multi-agent nature in our proposed MADRL-based algorithm is beneficial in terms of efficient and effective multiple charging-tower energy-sharing coordination.

The reminder of this paper is organized as follows. Section 2 summarizes related and previous work. Section 3 proposes a coordinated MADRL/CommNet-based energy source sharing algorithm among charging towers to minimize operational costs via minimizing energy purchases from the local energy market. Section 4 intensively evaluates the performance of the proposed coordinated MADRL/CommNet-based energy resource sharing algorithm via data-intensive simulations. Section 6 concludes this paper and provides future research directions.

2. Related Work

Nowadays, many UAV energy-efficient algorithms have been proposed. Among them, charging UAV devices via charging infrastructure that can be realized via wireless power transfer technologies is of interest [1,13]. For charging, the proposed algorithm in [11] designs an optimization framework for joint scheduling/matching UAVs and charging towers (i.e., charging infrastructure) and charging allocations. However, the proposed algorithm in [11] is not associated with charging tower coordination that is essentially required for active energy management. In [26], the proposed algorithm considers intelligent charging infrastructure coordination; however, scheduling is not considered because scheduling is

not required for electric vehicle (EV) charging problems because every EV driver decides where to go and the decision is independent from scheduling decisions. Furthermore, in [29,30], novel optimization and control algorithms for microgrid systems are discussed. However, the algorithms focus on infrastructure-level control; thus, UAV- and EV-related discussions and algorithm designs are not studied. Moreover, artificial intelligence and deep learning-based algorithms are not actively discussed; thus, the proposed algorithms in [29,30] are not superior in terms of stochastic and autonomous decision making under uncertainty. Therefore, to the best of our knowledge, our proposed algorithm is the first attempt for joint design of scheduling and charging infrastructure coordination.

In reinforcement learning algorithms, the use of a Markov decision process (MDP) is the simplest approach. Furthermore, mathematical analysis is also available under the concepts of Markov chain and dynamic programming. However, its computational complexity is huge, i.e., pseudo-polynomial; thus, it takes a lot of time to compute optimal solutions if the sizes of states are huge in reinforcement learning formulation. Thus, deep neural network based function approximation is used for reinforcement learning computation, and this is called deep reinforcement learning (DRL). Among various DRL algorithms, deep Q-network (DQN) is one of the most successful early-stage initial frameworks [31–33]. The DRL algorithms are extended from single-agent to multi-agent for cooperative and coordinated computation, and this is called MADRL [34,35]. In MADRL, CommNet [13,26] and the abstraction mechanism based on two-stage attention network (G2ANet) [36] are famous. The CommNet trains the multi-agent behaviors in a single deep neural network, and it assumes that all agents are homogeneous. On the other hand, in G2ANet, the relationship among agents are represented as graphs when the edge costs stand for the weights of correlation. Thus, the agents do not need to be homogeneous because the relationship can be trained with this graph structure. Therefore, G2ANet is beneficial for representing the sophisticated agent relationship, whereas it is computationally expensive because the relation graph is trained using two-stage attention models (i.e., hard attention and scale-dot attention). In considering our charging infrastructure coordination, we do not need to consider computationally expensive G2ANet because it is trivial to assume that all charging towers are equivalent. Therefore, a CommNet-based MADRL algorithm is used for our intelligent and autonomous learning computation.

3. Coordinated MADRL/CommNet-Based Energy Resource Sharing Learning

Our considered reference system model is explained in Section 3.1, and then, our considered scheduling algorithm for matching between charging towers and UAVs is presented in Section 3.2. Lastly, the proposed coordinated MADRL/CommNet-based energy resource sharing algorithm is introduced in Section 3.3.

3.1. System Model

In order to optimize and compute CommNet/MADRL-based energy resource sharing learning for charging towers, centralized computing (i.e., a cloud computing platform) is required in this paper. In the cloud, a deep learning neural architecture exists that optimizes and computes our proposed coordinated CommNet/MADRL-based energy resource sharing learning. Our cloud autonomously manages its own charging towers, where each charging tower has an energy storage for storing energy resources. Furthermore, the energy resources can be shared among charging towers if needed via CommNet/MADRL-based energy resource sharing learning mechanisms. If the shared energy resources are not enough to support charging UAVs, energy sources should be purchased from the local energy market (i.e., a utility company). The local energy market trades the energy based on the requests of charging towers in real-time.

3.2. Scheduling

The motivation of the scheduler design in our given problem is for efficiently and effectively providing energy/power resources from charging towers to their associated

UAVs via wireless power transfer technologies. Therefore, the scheduler should be able to determine a match between charging towers and UAVs. After that, the scheduler determines how much energy should be delivered from each charging tower to its associated UAV.

Thus, we can easily observe that this given scheduling problem is for the joint optimization for both of scheduling and energy resource allocation. Thus, it introduces the cases where two decision variables are multiplied [11].

3.3. Coordinated CommNet/MADRL-Based Energy Resource Sharing Learning

In order to design and implement MADRL-based algorithms for our given problem, we first have to identify that the problem cannot be formulated with single-agent deep reinforcement learning algorithms such as deep Q-network, as shown in Section 3.3.1. After that, our considering MADRL-algorithm, i.e., CommNet, is introduced to be used in our proposed coordinated MADRL-based energy resource sharing algorithm in Section 3.3.2.

3.3.1. Deep Q-Network and Its Limitation

In general MADRL problem formulations, states are formulated as matrices where the sizes is *A*-by-*B*, where *A* and *B* are the number of agents and the number of state variables, respectively. Assume that the states of agents can be denoted by $\mathcal{S} \triangleq \{s^1, \cdots, s^Z\}$. The state $\mathcal{S}$ in the policy π_θ returns action-value functions. The actions of individual agents are stochastically determined by following action-value functions [13],

$$Q(\mathbf{s}, \mathbf{a}; \theta) = \left\{ \max_a Q(s^1, a^1; \theta), \cdots, \max_a Q(s^Z, a^Z; \theta) \right\}. \tag{1}$$

Because dense layer computation in deep learning training occurs for each row in the state matrix, the actions of individual agents occur independent from the states of the other agents. Thus, $Q(s^z, a^z; \theta)$ in (1) is associated with a policy π_θ, and it is independent from the states of the other agents. Therefore, cooperative and coordinated actions among the individual agents cannot be expected with this deep Q-network-based deep learning neural architecture [31–33].

3.3.2. Cooperative Policy (CommNet)

In order to overcome the given problem in previous Section 3.3.1, each agent in CommNet gathers the states of the other agents s^{-j} to realize coordinated MADRL mechanisms. Here, the other agents can be represented as follows:

$$s^{-j} \triangleq \left\{ s^1, \cdots, s^{j-1}, s^{j+1}, \cdots, s^J \right\}. \tag{2}$$

For $\forall i$ and $\forall j$, the hidden variable $h^{i,j}$, which is the parameter of the *i*th hidden layer, gathers other hidden variables $h^{i,-j}$ and, then, $h^{i,-j}$ takes the mean operation. The computational process for NADRL/CommNet-based agents is represented as follows:

$$h^{i+1,j} \triangleq g\left(h^{i,j}, c^{i,j} \right), \tag{3}$$

$$c^{i,j} \triangleq \frac{\mid h^{i,-j} \mid}{J-1}, \tag{4}$$

where $g(\cdot)$ and $c^{i,j}$ mean an activation function and the communication variable of *j*th agent, respectively. The considered individual agents can receive average messages among them via this communication neural architecture. Notice that

$$s^{-j} \triangleq \bigcup_{j' \neq j} \left\{ s^{j'} \right\}, \tag{5}$$

$$h^{i,-j} \triangleq \bigcup_{j' \neq j} \left\{ h^{i,j'} \right\}, \tag{6}$$

where *i* and *j* are the orders of the neural layer and the agent, respectively. Figure 1 presents the neural-architectural comparison between deep Q-network and CommNet. As shown in Figure 1, the actions from the deep Q-network-based policy are independent from other agents; therefore, coordinated actions cannot be realized. On the other hand, the actions from this MADRL/CommNet-based policy are dependent on them because they share a single deep-learning neural architecture, and thus, coordinated and cooperative MADRL actions can be realized and obtained. Therefore, this MADRL/CommNet-based policy has only one policy, but it is possible to create a system that coordinates and cooperates while sharing learning information among them. The input and output of the deep learning neural architecture for performing training optimization in MADRL/CommNet-based energy resource sharing learning computation are the states (charging tower energy status values) and actions (charging decision values), respectively [26].
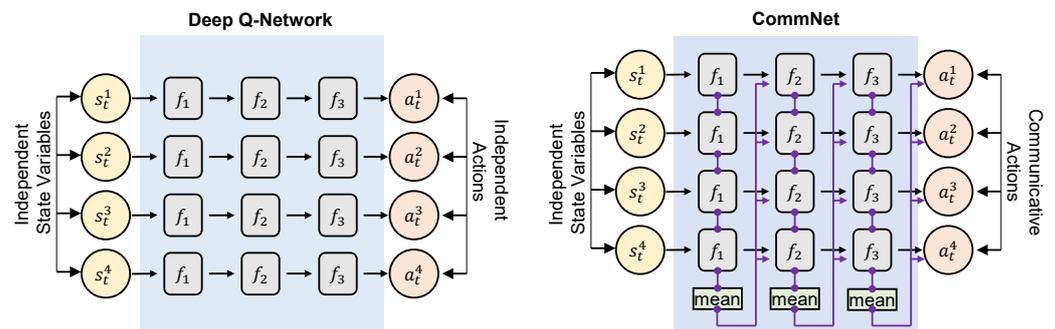
Figure 1. The comparison of deep Q-network and CommNet architectures in terms of the communications among individual agents.

4. Performance Evaluation

This section consists of the performance evaluation setting and setup (refer to Section 4.1) and the corresponding results (refer to Section 4.2).

4.1. Evaluation Setup

This section presents the basic setup for evaluation of the proposed coordinated MADRL/CommNet-based energy resource sharing learning in multi-UAV networking systems.

For the network simulation setup in performance evaluation, the movement coverage values of individual UAVs are set to 10×10 and the entire simulation topology is defined as an urban Manhattan grid 4390×2500. In addition, the number of UAVs and charging towers are $|\mathcal{U}| = 30$ and $|\mathcal{C}| = 4$, respectively, where $\mathcal{U}$ and $\mathcal{C}$ are defined as the sets of UAVs and charging towers. The other simulation-based performance evaluation parameters and settings are presented in Table 1.

Table 1. Simulation-based performance evaluation parameters.

Parameters	Value
The number of UAVs	30
The number of charging towers	4
Maximum energy generation of PV	17.7 W
Energy capacity of charging towers	500 Wh
State of charge ranges	Min: 25%, Max: 50%
Available energy of ESS	125 Wh
ϵ-greedy parameter, ϵ	1
ϵ decay, γ	10^{-4}
Wasted energy reward parameter, σ_w	200
Purchased energy reward parameter, σ_p	4000
Shared energy reward parameter, σ_s	27

This simulation-based performance evaluation is conducted while comparing the performances of following two methods with our proposed coordinated MADRL/CommNet-based energy resource sharing algorithm (denoted as Proposed in this paper).

- Our proposed coordinated MADRL/CommNet-based energy resource sharing without efficient and effective scheduling is considered one possible candidate for comparison. Our considered scheduling algorithm is introduced in Section 3.2, but this is excluded for performance comparison. Note that this algorithm is denoted as Random Scheduling in this paper.
- For the second algorithm, in order to conduct performance comparison, we consider the algorithm with efficient and effective scheduling in Section 3.2 but without coordinated MADRL/CommNet-based energy resource sharing. Note that this algorithm is denoted as Random Sharing in this paper.

As discussed in Section 2, the joint scheduling and DRL-based coordinated energy sharing in a charging infrastructure is not studied. Therefore, comparing our proposed algorithm with random scheduling and random sharing algorithms is considered in this performance evaluation.

Our simulation software is implemented with `Python 3.6.5` over the `Ubuntu 18.04 LTS` operating system machine. For scheduler implementation, well-known optimization tools, i.e., `CVXPY 1.1` and `MOSEK 9`, are used [37,38]. In addition, our proposed MADRL-based algorithm is implemented with `tensorflow-gpu 1.5.0`. For the MADRL/CommNet algorithm implementation, the two-layer neural network architecture of energy resource sharing is configured as follows. It includes 6 hidden layers, where the number of units in the first three layers (layer 1, layer 2, and layer 3) is 512 for each and the remainder (layer 4, layer 5, and layer 6) has 1024 units for each. The hyperbolic-tangent (denoted as tanh) and rectified linear unit (denoted as ReLU) functions are considered activation functions for the first three and reminder layers, respectively. Moreover, a Xavier initializer is used for weight initialization; andan Adam optimizer is used for parameter learning optimization. During the neural network training procedure, an ϵ-greedy method is used to make the charging tower agents explore a variety of actions.

Figure 2 presents the photovoltaic (PV) power generation distribution in each charging tower over time. The individual charging towers have their own PV power generation distribution because they have their own individual PV power generation capacities, locations, solar radiation quantities, and so forth. The loads of charging towers are defined as the numbers of UAVs determined to be charged by the towers (determined as explained in Section 3.2), and the numerical values and their fluctuations are illustrated in Figure 2. Lastly, the power/energy prices from the local energy market can be presented as a probabilistic distribution depending on time-of-use (ToU) at each unit time.
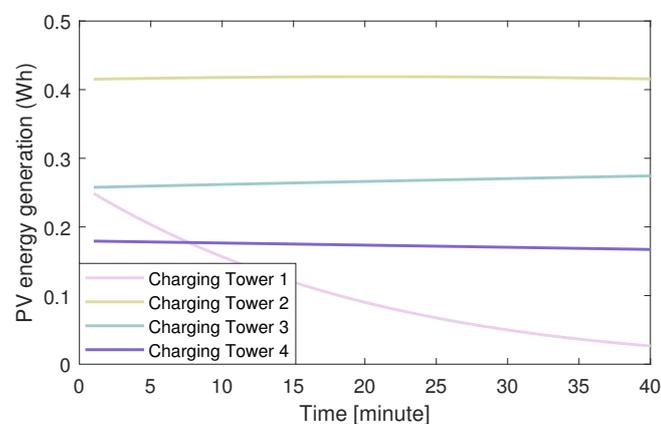

Figure 2. Photovoltaic power generation distributions in individual charging towers.

4.2. Evaluation Results

This section presents the simulation-based performance evaluation results for our proposed coordinated MADRL/CommNet-based algorithm (i.e., Proposed) compared with two algorithms, i.e., Random Scheduling and Random Sharing. This simulation-based evaluation is performed in terms of scheduling (refer to Section 4.2.1) and energy sharing (refer to Section 4.2.2). Lastly, the summary of this simulation-based performance evaluation is presented in Section 4.2.3.

4.2.1. Scheduling

Our proposed scheduling in Section 3.2 is designed for energy resource balancing among charging towers. Thus, the performance evaluation is conducted in this perspective. Figure 3a,b show the remaining battery/energy capacities distribution in UAVs. The initial batteries/energies of UAVs are uniformly randomly selected in [5283, 5870] mAh. As presented in Figure 3, the Proposed algorithm is superior to the Random Scheduling algorithm because Figure 3a shows better energy-aware behaviors. Moreover, as presented in Table 2, the average and variance of residual battery/energy amounts in UAVs are summarized for both Proposed and Random Scheduling. In Table 2, we can confirm that the Proposed algorithm takes higher average values of residual energies over the entire time period. The reason for this is that the number of charged UAVs with the Proposed algorithm is higher than the number of charged UAVs with the Random Scheduling algorithm. Furthermore, it can be also observed that the standard deviation of the Proposed algorithm is smaller. This means that the Proposed algorithm is able to provide charging services under consideration of energy charging load-balancing and fairness.

Table 2. Residual battery/energy amounts of unmanned aerial vehicles (UAVs) (unit: percentage) for both of the Proposed algorithm and the Random Scheduling algorithm, where μ and σ stand for the average and variance of UAV battery/energy remains.
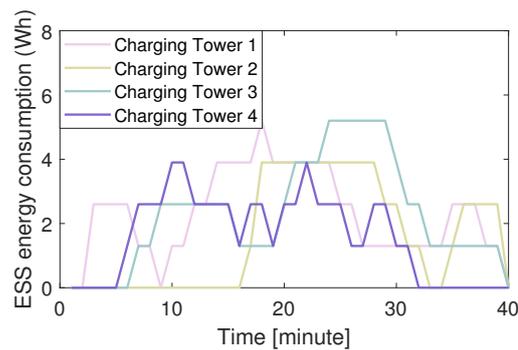
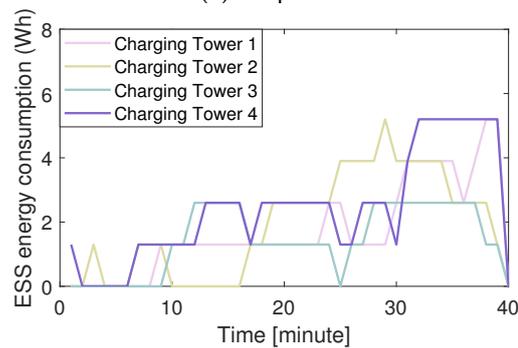|	Proposed		Random Scheduling	
t [min]	μ	σ	μ	σ
0 min–5 min	90.1%	0.2	89.9%	0.2
6 min–10 min	79.8%	0.3	78.2%	0.2
11 min–15 min	71.1%	0.9	68.1%	0.3
16 min–20 min	63.0%	1.5	58.9%	1.0
21 min–25 min	57.3%	1.9	50.4%	2.1
26 min–30 min	51.2%	2.7	42.2%	2.9
31 min–35 min	41.9%	3.2	36.5%	3.6
36 min–40 min	32.0%	2.7	31.1%	4.5

(**a**) Proposed.

(**b**) Random Scheduling.
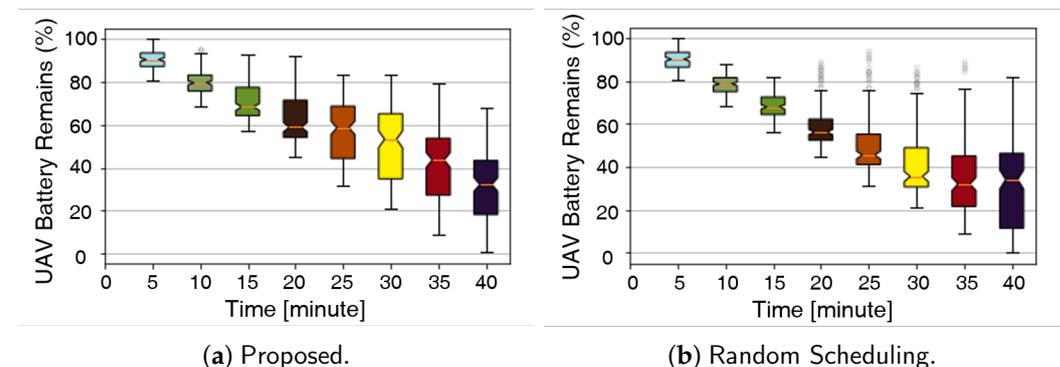
Figure 3. UAV residual battery/energy distribution comparison between (**a**) the Proposed algorithm and (**b**) Random Scheduling algorithm.

Figure 4a,b are the energy consumption (also called loads) in the charging towers when the Proposed algorithm and the Random Scheduling algorithm are utilized. In Figure 4c, the distributions of differences in terms of energy consumption (or loads) between the Proposed algorithm and the Random Scheduling algorithm are presented. As observed in Figure 4c, relatively fair energy consumption over time can be achieved with the Proposed algorithm compared to the energy consumption over time with the Random Scheduling algorithm.

As shown in Figure 5a,b, for the Proposed algorithm and the Random Scheduling algorithm, the purchased energy from local energy market in Figure 5a is obviously smaller than that of Figure 5b because of the novelty of the Proposed algorithm. This means that our proposed scheduling is efficient in terms of energy consumption load-balancing among charging towers.
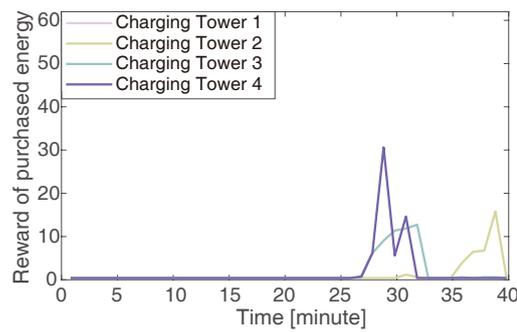

(**a**) Proposed.

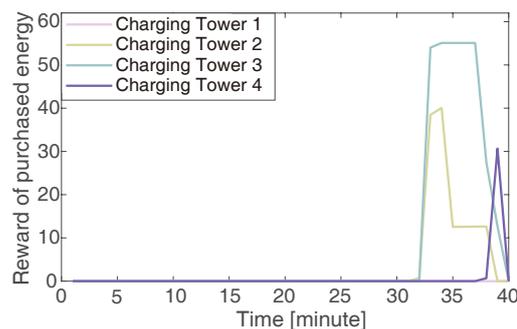(**b**) Random Scheduling.
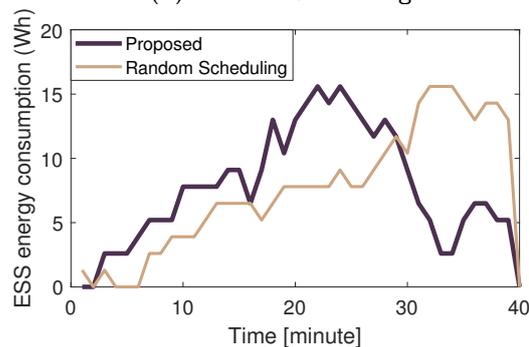
(**c**) Total load comparison.

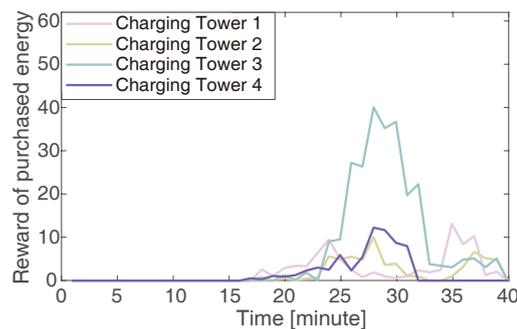**Figure 4.** Energy consumption (load) in each charging tower with (**a**) the Proposed algorithm, (**b**) the Random Scheduling algorithm, and (**c**) comparison of the total amount between the Proposed algorithm and the Random Scheduling algorithm.
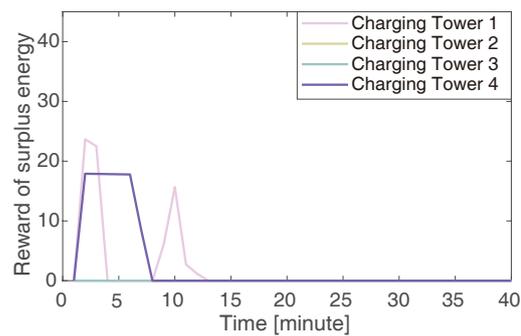
(**a**) Proposed.

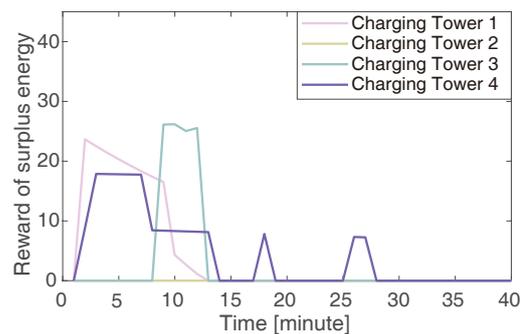(**b**) Random Scheduling.

(**c**) Random Sharing.

Figure 5. Purchased energy from a local energy market utility company with (**a**) the Proposed algorithm, (**b**) the Random Scheduling algorithm, and (**c**) the Random Sharing algorithm.

The surplus energy stands for the energy that overflowed due to unnecessarily energy purchases from the local energy market. As presented in Figure 6a,b, the amounts of surplus energies in the Proposed algorithm and the Random Scheduling algorithm are numerically simulated. The simulation results in terms of surplus energy show that the amount in Figure 6a is smaller than that of Figure 6b because our Proposed algorithm outperforms the other. The amounts of surplus energy in the Proposed algorithm is smaller because the corresponding loads in Figure 4 are bigger.
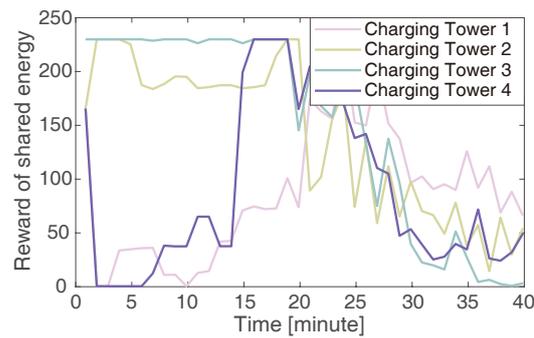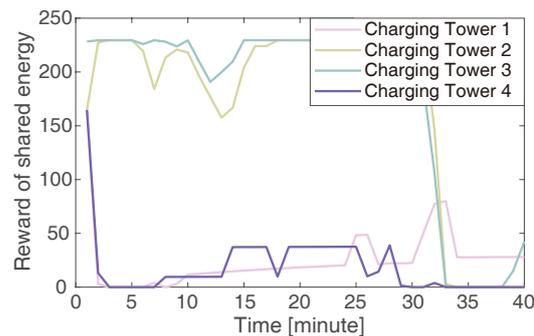
(**a**) Proposed.

(**b**) Random Scheduling.
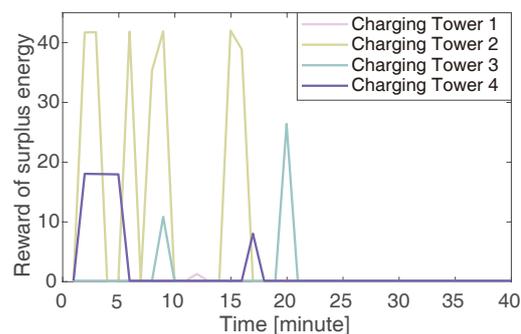
(**c**) Random Sharing.

Figure 6. Surplus energy with (**a**) the Proposed algorithm, (**b**) the Random Scheduling algorithm, and (**c**) the Random Sharing algorithm.

In our consideed charging systems for UAV networks, facilitating energy resource sharing among charging towers is obviously beneficial in terms of the minimization of energy purchase from the local energy market because sharing can increase the possibility of energy provisioning in charging towers that do not have sufficient energy resources. As shown in Figure 7a,b, the Proposed algorithm has relatively larger energy sharing among charging towers, whereas the Random Scheduling algorithm leads to dramatically less sharing during the last simulation runs. The reason for this is that the energy sharing with the Random Scheduling algorithm becomes exhausted due to the failure of energy consumption load-balancing.

(**a**) Proposed.

(**b**) Random Scheduling.
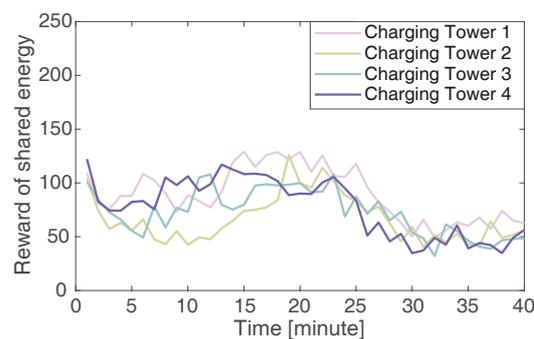
(**c**) Random Sharing.

Figure 7. Shared energy among charging towers with (**a**) the Proposed algorithm, (**b**) the Random Scheduling algorithm, and (**c**) the Random Sharing algorithm.

4.2.2. Learning-BASED Energy Sharing

The performance of coordinated MADRL/CommNet-based energy resource sharing learning was evaluated. As presented in Figure 5a,c, our Proposed algorithm has much less energy purchase from the local energy market because the reward of the MADRL/CommNet-based method in this paper is negative for energy purchase. Therefore, the Proposed algorithm minimizes energy purchase costs (which is strongly related to system-wide operational costs). Figure 6a,c show the distributions of surplus energies (set to negative reward in our MADRL/CommNet). As shown in Figure 7a compared to Figure 7c, the Proposed algorithm presents more frequent energy resource sharing because it maximizes positive reward in our proposed MADRL/CommNet. As shown in Figure 7c, the average amount of shared energy with the Proposed algorithm is larger than the amount with the Random Sharing algorithm.

4.2.3. Summary

As clearly stated in our simulation-based performance evaluation results, it has been verified that the Proposed algorithm is efficient in terms of energy consumption load-balancing among charging towers. As presented in Figure 8a, convergence of the total

reward of our proposed MADRL/CommNet verifies that the Proposed algorithm outper-
forms the other methods; thus, intelligent and efficient energy management and control
can be realized. Our Proposed algorithm eventually converges to positive optimal rewards,
whereas the other two comparing algorithms, i.e., Random Scheduling algorithm and Ran-
dom Sharing algorithm, converges to negative values, as shown in Figure 8a. Furthermore,
the values in Figure 8b,c of our Proposed algorithm are lower than the others because
they present negative reward values, i.e., purchased energy and surplus energy. Similarly,
values in Figure 8d of our Proposed algorithm is the highest in general, because it shows
positive reward (i.e., shared energy).

Finally, we can confirm that our proposed coordinated MADRL/CommNet-based en-
ergy resource sharing learning achieves desired performance improvements by optimizing
its own reward function that depends on purchased energy (negative reward), surplus
energy (negative reward), and shared energy (positive reward), as also verified based on
the performance evaluation data in Table 3.

(**a**) Total reward.

(**b**) Purchased energy.

(**c**) Surplus energy.
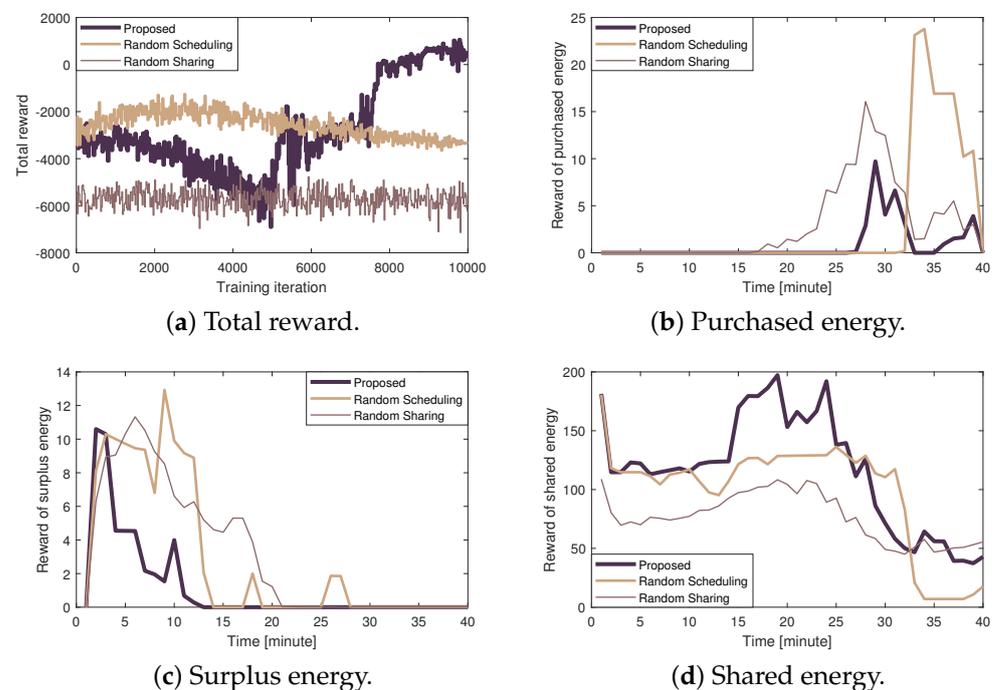
(**d**) Shared energy.

Figure 8. Various reward value distributions in terms of (**a**) total reward, (**b**) purchased energy,
(**c**) surplus energy, and (**d**) shared energy, respectively, while comparing the Proposed algorithm, the
Random Scheduling algorithm, and the Random Sharing algorithm.

Table 3. The list of each obtained reward value and load of charging towers.

Parameters	Proposed	Random Scheduling	Random Sharing
Load of charging tower (Wh)	360.8	360.8	360.8
Reward of purchased energy (negative)	137.9	475.4	472.8
Reward of surplus energy (negative)	180.5	449.9	498.9
Reward of shared energy (positive)	18,547.3	15,620.9	12,210.8

5. Applications in Big-Data Processing Platforms

Our considered multi-UAV networks can be widely used for many applications.
Furthermore, the proposed coordinated charging system and its related intelligent and
autonomous algorithms are also definitely useful.

Especially, multiple UAV devices are able to gather extremely large-scale surveillance and cellular network big-data [39–41]. For surveillance, multiple UAV devices can be utilized for monitoring extreme harsh areas and then for gathering security big-data from extreme areas such as dense forests and seaside coasts where network infrastructure cannot be established. Furthermore, the proposed coordinated algorithm can be also used for extending network coverage because individual UAVs are able to work as mobile base stations. Then, each UAV can gather big-data information such as massive user association and large-scale traffic patterns.

The mentioned surveillance and mobile cellular networks data are generated in real-time and the amounts are quite large. Thus, corresponding big-data processing algorithms are essentially required and it is obvious that the algorithms are generally computationally expensive and thus requires large amounts of energy resources. Therefore, design and implementation of energy-aware algorithms in UAVs as well as charging infrastructure such as charging towers are desired.

6. Concluding Remarks and Future Work

According to the autonomous and flexible characteristics of UAV networks, they are widely and actively used for next-generation mobile network design and implementation. The utilization of autonomous UAV systems can realize high-mobility aerial surveillance and mobile wireless cellular network base station deployment; therefore, large-scale flexible big-data processing where the data were gathered via multiple UAVs can be consequentially achieved. In order to facilitate the use of power-hungry UAVs for big-data computing applications, active and efficient energy-aware charging mechanisms for autonomous UAVs are required via wireless power transfer technologies. Therefore, the use of charging towers is required. In this system, we propose joint scheduling and coordinated energy sharing algorithm for energy-aware system management. For scheduling, the matching/scheduling between UAVs and charging towers is considered along with the optimal decision for energy/power source allocation amounts. In addition, fFor minimizing the operational costs in our considering systems, the energy stored in individual charging towers should be shared among charging towers in order to minimize energy purchase from the local energy market. Therefore, our proposed energy resource sharing learning algorithm minimizes operational costs by coordinating MADRL/CommNet-based intelligent cooperation among charging towers. This type of MADRL-based algorithm is beneficial because it realizes stochastic and autonomous decision making under uncertainty. Lastly, our simulation-based performance evaluation results verify that the proposed joint scheduling and coordinated MADRL/CommNet-based energy resource sharing algorithm achieves desired performance improvements.

As potential future work directions, we can consider safe deep reinforcement learning-related design and implementation, which is useful to consider safe, robust, and privacy-aware operations in UAV charging scheduling control and optimization. Furthermore, larges-scale data-intensive simulations are also valuable for more deep-dive discussions in terms of performance evaluation.

References

1. Shin, M.; Kim, J.; Levorato, M. Auction-Based Charging Scheduling With Deep Learning Framework for Multi-Drone Networks. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4235–4248. [CrossRef]
2. Geraldes, R.; Gonçalves, A.; Lai, T.; Villerabel, M.; Deng, W.; Salta, A.; Nakayama, K.; Matsuo, Y.; Prendinger, H. UAV-Based Situational Awareness System Using Deep Learning. *IEEE Access* **2019**, *7*, 122583–122594. [CrossRef]
3. Truong, N.Q.; Nguyen, P.H.; Nam, S.H.; Park, K.R. Deep Learning-Based Super-Resolution Reconstruction and Marker Detection for Drone Landing. *IEEE Access* **2019**, *7*, 61639–61655. [CrossRef]
4. Huang, H.; Yang, Y.; Wang, H.; Ding, Z.; Sari, H.; Adachi, F. Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1117–1121. [CrossRef]
5. Hu, J.; Zhang, H.; Song, L. Reinforcement Learning for Decentralized Trajectory Design in Cellular UAV Networks With Sense-and-Send Protocol. *IEEE Internet Things J.* **2019**, *6*, 6177–6189. [CrossRef]
6. Liu, X.; Liu, Y.; Chen, Y. Reinforcement Learning in Multiple-UAV Networks: Deployment and Movement Design. *IEEE Trans. Veh. Technol.* **2019**, *68*, 8036–8049. [CrossRef]
7. Wu, F.; Zhang, H.; Wu, J.; Song, L. Cellular UAV-to-Device Communications: Trajectory Design and Mode Selection by Multi-Agent Deep Reinforcement Learning. *IEEE Trans. Commun.* **2020**, *68*, 4175–4189. [CrossRef]
8. Yin, S.; Zhao, S.; Zhao, Y.; Yu, F.R. Intelligent Trajectory Design in UAV-Aided Communications With Reinforcement Learning. *IEEE Trans. Veh. Technol.* **2019**, *68*, 8227–8231. [CrossRef]
9. Cui, J.; Liu, Y.; Nallanathan, A. Multi-Agent Reinforcement Learning-Based Resource Allocation for UAV Networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 729–743. [CrossRef]
10. Jung, S.; Yang, P.; Quek, T.Q.S.; Kim, J.H. Belief Propagation based Scheduling for Energy Efficient Multi-drone Monitoring System. In Proceedings of the IEEE International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 21–23 October 2020; pp. 261–263. [CrossRef]
11. Jung, S.; Kim, J.; Kim, J.H. Joint Message-Passing and Convex Optimization Framework for Energy-Efficient Surveillance UAV Scheduling. *Electronics* **2020**, *9*, 1475. [CrossRef]
12. Kwon, D.; Kim, J. Optimal Trajectory Learning for UAV-BS Video Provisioning System: A Deep Reinforcement Learning Approach. In Proceedings of the IEEE International Conference on Information Networking (ICOIN), Kuala Lumpur, Malaysia, 9–11 January 2019; pp. 372–374.
13. Jung, S.; Yun, W.J.; Kim, J.; Kim, J.H. Infrastructure-Assisted Cooperative Multi-UAV Deep Reinforcement Energy Trading Learning for Big-Data Processing. In Proceedings of the IEEE International Conference on Information Networking (ICOIN), Jeju Island, Korea, 13–16 January 2021.
14. Zhang, S.; Zhang, H.; Song, L. Beyond D2D: Full Dimension UAV-to-Everything Communications in 6G. *IEEE Trans. Veh. Technol.* **2020**, *69*, 6592–6602. [CrossRef]
15. Shang, B.; Liu, L.; Ma, J.; Fan, P. Unmanned Aerial Vehicle Meets Vehicle-to-Everything in Secure Communications. *IEEE Commun. Mag.* **2019**, *57*, 98–103. [CrossRef]
16. Na, W.; Park, J.; Lee, C.; Park, K.; Kim, J.; Cho, S. Energy-Efficient Mobile Charging for Wireless Power Transfer in Internet of Things Networks. *IEEE Internet Things J.* **2018**, *5*, 79–92. [CrossRef]
17. Park, L.; Jeong, S.; Lakew, D.S.; Kim, J.; Cho, S. New Challenges of Wireless Power Transfer and Secured Billing for Internet of Electric Vehicles. *IEEE Commun. Mag.* **2019**, *57*, 118–124. [CrossRef]
18. Zhao, D.; Wang, H.; Huang, J.; Lin, X. Storage or No Storage: Duopoly Competition Between Renewable Energy Suppliers in a Local Energy Market. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 31–47. [CrossRef]
19. Correa-Florez, C.A.; Michiorri, A.; Kariniotakis, G. Optimal Participation of Residential Aggregators in Energy and Local Flexibility Markets. *IEEE Trans. Smart Grid* **2020**, *11*, 1644–1656. [CrossRef]
20. Ghorani, R.; Fotuhi-Firuzabad, M.; Moeini-Aghtaie, M. Optimal Bidding Strategy of Transactive Agents in Local Energy Markets. *IEEE Trans, Smart Grid* **2019**, *10*, 5152–5162. [CrossRef]
21. Siano, P.; De Marco, G.; Rolán, A.; Loia, V. A Survey and Evaluation of the Potentials of Distributed Ledger Technology for Peer-to-Peer Transactive Energy Exchanges in Local Energy Markets. *IEEE Syst. J.* **2019**, *13*, 3454–3466. [CrossRef]
22. Xiao, Y.; Wang, X.; Pinson, P.; Wang, X. A Local Energy Market for Electricity and Hydrogen. *IEEE Trans. Power Syst.* **2018**, *33*, 3898–3908. [CrossRef]
23. Park, L.; Jeong, S.; Kim, J.; Cho, S. Joint Geometric Unsupervised Learning and Truthful Auction for Local Energy Market. *IEEE Trans. Ind. Electron.* **2019**, *66*, 1499–1508. [CrossRef]
24. Mo, Y.J.; Kim, J.; Kim, J.; Mohaisen, A.; Lee, W. Performance of Deep Learning Computation with TensorFlow Software Library in GPU-Capable Multi-Core Computing Platforms. In Proceedings of the IEEE International Conference on Ubiquitous and Future Networks (ICUFN), Milan, Italy, 4–7 July 2017; pp. 240–242. [CrossRef]
25. Ahn, S.; Kim, J.; Lim, E.; Choi, W.; Mohaisen, A.; Kang, S. ShmCaffe: A Distributed Deep Learning Platform with Shared Memory Buffer for HPC Architecture. In Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS), Vienna, Austria, 2–6 July 2018; pp. 1118–1128.
26. Shin, M.; Choi, D.; Kim, J. Cooperative Management for PV/ESS-Enabled Electric Vehicle Charging Stations: A Multiagent Deep Reinforcement Learning Approach. *IEEE Trans. Ind. Inf.* **2020**, *16*, 3493–3503. [CrossRef]

27. Erdelj, M.; Natalizio, E.; Chowdhury, K.R.; Akyildiz, I.F. Help from the Sky: Leveraging UAVs for Disaster Management. *IEEE Pervasive Comput.* **2017**, *16*, 24–32. [CrossRef]

28. Chen, W.; Liu, B.; Huang, H.; Guo, S.; Zheng, Z. When UAV Swarm Meets Edge-Cloud Computing: The QoS Perspective. *IEEE Netw.* **2019**, *33*, 36–43. [CrossRef]

29. Zhou, Q.; Shahidehpour, M.; Paaso, A.; Bahramirad, S.; Alabdulwahab, A.; Abusorrah, A. Distributed Control and Communication Strategies in Networked Microgrids. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 2586–2633. [CrossRef]

30. Zhou, Q.; Tian, Z.; Shahidehpour, M.; Liu, X.; Alabdulwahab, A.; Abusorrah, A. Optimal Consensus-Based Distributed Control Strategy for Coordinated Operation of Networked Microgrids. *IEEE Trans. Power Syst.* **2020**, *35*, 2452–2462. [CrossRef]

31. Su, Y.; Fan, R.; Fu, X.; Jin, Z. DQELR: An Adaptive Deep Q-Network-Based Energy- and Latency-Aware Routing Protocol Design for Underwater Acoustic Sensor Networks. *IEEE Access* **2019**, *7*, 9091–9104. [CrossRef]

32. Luo, Y.; Yang, J.; Xu, W.; Wang, K.; Renzo, M.D. Power Consumption Optimization Using Gradient Boosting Aided Deep Q-Network in C-RANs. *IEEE Access* **2020**, *8*, 46811–46823. [CrossRef]

33. Xu, W.; Yu, J.; Miao, Z.; Wan, L.; Ji, Q. Spatio-Temporal Deep Q-Networks for Human Activity Localization. *IEEE Trans. Circ. Syst. Video Technol.* **2020**, *30*, 2984–2999. [CrossRef]

34. Kwon, D.; Kim, J. Multi-Agent Deep Reinforcement Learning for Cooperative Connected Vehicles. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019; pp. 1–6. [CrossRef]

35. Kwon, D.; Jeon, J.; Park, S.; Kim, J.; Cho, S. Multiagent DDPG-Based Deep Learning for Smart Ocean Federated Learning IoT Networks. *IEEE Internet Things J.* **2020**, *7*, 9895–9903. [CrossRef]

36. Liu, Y.; Wang, W.; Hu, Y.; Hao, J.; Chen, X.; Gao, Y. Multi-Agent Game Abstraction via Graph Attention Neural Network. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI), New York, NY, USA, 7–12 February 2020; pp. 7211–7218.

37. Diamond, S.; Boyd, S. CVXPY: A Python-embedded modeling language for convex optimization. *J. Mach. Learn. Res.* **2016**, *17*, 2909–2913.

38. Andersen, E.D.; Andersen, K.D. The MOSEK interior point optimizer for linear programming: An implementation of the homogeneous algorithm. *High Perform. Optim.* **2000**, *33*, 192–232.

39. Kim, J.; Lee, W. Stochastic Decision Making for Adaptive Crowdsourcing in Medical Big-Data Platforms. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 1471–1476. [CrossRef]

40. Jeon, J.; Kim, J.; Kim, J.; Kim, K.; Mohaisen, A.; Kim, J. Privacy-Preserving Deep Learning Computation for Geo-Distributed Medical Big-Data Platforms. In Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks–Supplemental Volume (DSN-S), Portland, OR, USA, 24–27 June 2019; pp. 3–4. [CrossRef]

41. Yoo, S.; Kim, H.; Kim, J. Secure Compute-VM: Secure Big Data Processing with SGX and Compute Accelerators. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS) Workshop on System Software for Trusted Execution*; ACM: New York, NY, USA, 2018; pp. 34–36.