

Research Article

Energy Cooperation in Ultradense Network Powered by Renewable Energy Based on Cluster and Learning Strategy

Chunhong Duo, Baogang Li, Yongqian Li, and Yabo Lv

North China Electric Power University, Baoding, China

Correspondence should be addressed to Chunhong Duo; duochunhong@163.com

Received 24 May 2017; Revised 30 July 2017; Accepted 7 September 2017; Published 16 October 2017

Academic Editor: Patrick Seeling

Copyright © 2017 Chunhong Duo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new method about renewable energy cooperation among small base stations (SBSs) is proposed, which is for maximizing the energy efficiency in ultradense network (UDN). In UDN each SBS is equipped with energy harvesting (EH) unit, and the energy arrival times are modeled as a Poisson counting process. Firstly, SBSs of large traffic demands are selected as the clustering centers, and then all SBSs are clustered using dynamic k -means algorithm. Secondly, SBSs coordinate their renewable energy within each formed cluster. The process of energy cooperation among SBSs is considered as Markov decision process. Q-learning algorithm is utilized to optimize energy cooperation. In the algorithm there are four different actions and their corresponding reward functions. Q-learning explores the action as much as possible and predicts better action by calculating reward. In addition, ϵ greedy policy is used to ensure the algorithm convergence. Finally, simulation results show that the new method reduces data dimension and improves calculation speed, which furthermore improves the utilization of renewable energy and promotes the performance of UDN. Through online optimization, the proposed method can significantly improve the energy utilization rate and data transmission rate.

1. Introduction

With the rapid development of mobile network, wireless data traffic has increased exponentially. More and more small base stations (SBSs) are needed to satisfy the traffic demand, which results in ultradense network (UDN). Meanwhile energy harvesting (EH) has become a very promising technology because of its flexible deployment and renewable energy [1]. The combination between SBS and EH can make it possible to harvest energy to extend the network running time. However, there are some difficulties for EH technology [2, 3], that is, the randomness and instability of renewable energy, the limited energy storage of EH devices, and so on.

The energy cooperation in EH wireless networks has recently been studied extensively. There are lots of researches for maximizing network throughput, energy cooperation, traffic management, and so on. Energy cooperation schemes among different BSs are considered [4–6], and the energy efficiency (EE) is improved by adopting the evolutionary algorithms. The jointly optimal policy that maximizes sum-throughput is proposed [7–9]; both energy cooperation and

traffic management are considered. The capacity region coincides with a traditional K -user Gaussian MAC, and users can perform energy cooperation [10]. The cooperation between primary and secondary users at information and energy levels is considered [11]. These studies mainly focus on energy cooperation between the cells or users in cellular network or the hybrid power supply system, which aims at improving the EE and traffic management. With the increase of SBSs in the future communication network, it is necessary to divide SBSs into small groups for optimal energy cooperation.

UDN has a much higher deployment density than current mobile network, which greatly improves the network throughput. Many studies on spectrum sensing, spectrum efficiency, and intercell interference have been published [12–15]. EE and resource management are studied in UDN [16–18], and a cluster-based EE resource allocation scheme is proposed [19]. Taking advantage of the density feature of SBS, enhancing the energy cooperation will promote the network performance. With the increasing density of SBSs, it is more efficient and practical to utilize cluster and online learning method for greater network throughput.

There are also researches about energy supply especially EH technology. A system consisting of two cooperative microgrids is considered [20], and the two microgrids exchange renewable energy through the transmission line. Two energy sources for supplying the energy required for system operation are proposed, that is, an energy harvester and a constant energy source driven by a nonrenewable resource [21–23]. An energy harvesting circuit is equipped at the energy harvesting receiver and introduces various nonlinearities into the wireless power transfer [24]. A dense small cell network consisting of a set of small cells and a set of users is considered. In the network every user selects an SBS by itself, and multiple users can be served by a single SBS [25]. For simplicity, we only consider the renewable energy cooperation among clustered SBSs in UDN.

Considering that clustering technique obtains optimal energy cooperation and reinforcement learning gains greater EE, we combine EH technology with clustered energy cooperation in UDN to solve EE problems. A renewable energy cooperation management algorithm is proposed based on clustered SBSs in UDN. Firstly, according to the geographical location and traffic load, sampling technique is used to determine the centers of data division, and then all SBSs are clustered using dynamic k -means algorithm. Secondly, within each formed cluster, SBSs coordinate their renewable energy. The process of energy cooperation among SBSs is considered as MDP. Q-learning algorithm is adopted to optimize energy cooperation, in which ϵ greedy policy can ensure algorithm convergence. There are four different actions and their corresponding immediate reward functions for each SBS. The reward function represents the energy value after taking the corresponding action. Q-learning explores the action as much as possible, learns how to predict the relationship between the action and the reward, and furthermore predicts better actions by calculating rewards. At last, simulation results show that the new method reduces data dimension and improves calculation speed, which furthermore improves the utilization of renewable energy and promotes the performance of UDN. Through online optimization, the proposed method can significantly improve the energy utilization rate and data transmission rate.

This paper is organized as follows: Section 2 presents the system model according to MDP. In Section 3, we optimize the problem of energy cooperation among SBSs in UDN using clustering technology and reinforcement learning. Simulation results are given in Section 4. Section 5 concludes the paper.

2. System Model

In practice, the EH model depends on its specific implementation. There are solar panel and wind turbine-generator which can generate renewable energy, but the EH characteristics are different in both cases. The energy arrival times in the energy harvester can be modeled as a Poisson counting process [21]. Sinusoidal curve can also be selected [22]. In order to provide a general model for EH communication systems, we model it as a stochastic process in order to isolate the considered problem from specific implementation

assumptions. In this paper, the location of SBSs in UDN is randomly deployed [26, 27], and each SBS is equipped with EH unit and a limited capacity battery. Assume that a limited time-slot (TS) system, the renewable energy, and the required data are arrived at the beginning of each TS. The channel state information is $H(t)$, which is kept constant in the same TS. The processing of energy/data packet can be considered as one-order discrete Markov model [28]. In each TS, the amount of data SBS processed is D_n , and the minimum energy required for data transmission is $E_n^T(t)$. The energy arrival times in the EH unit are modeled as a Poisson counting process with λ_E . The harvested energy $E_n^H(t)$ in each TS is $E_n^H(t) = (\lambda_E^t / t!) * e^{-\lambda_E}$, $t = 0, 1, 2, \dots$

In UDN, for a certain SBS n , the system parameters are given in System Parameters in UDN section.

Energy charged to the battery is $c_n(t)$, $c_n(t) \geq 0$; and energy discharged from the battery is $d_n(t)$, $d_n(t) \geq 0$. At any time, SBS will charge/discharge energy to/from battery. There is at most one of $c_n(t)$ and $d_n(t)$ that is strictly positive, that is, $c_n(t) * d_n(t) = 0$.

The SBS operation in TS t satisfies the following constraints (α represents battery charging efficiency):

$$\begin{aligned} E_n^H(t) &\geq 0 \\ E_n^T(t) &\geq 0 \\ c_n(t) &\geq 0 \\ d_n(t) &\geq 0 \\ c_n(t) * d_n(t) &= 0 \\ 0 &\leq B_n(t) \leq B_{n\max} \\ 0 &\leq B_n(t) + \alpha c_n(t) - d_n(t) \leq B_{n\max}. \end{aligned} \quad (C1)$$

We consider one hot region in UDN, which is shown in Figure 1. The colored SBSs are the cluster centers, which form groups with the uncolored SBSs in their circles. Other SBSs join to the nearest group. The SBSs in one cluster achieve energy cooperation. For simplicity, we only consider the scenario of one user. For the request data rate of user, we ignore the fluctuation of data services and suppose that there is the full traffic data case. The user always has data to be received, and SBSs are in full load operation.

To simplify the system model, we give one SBS energy harvesting model in UDN, which is shown in Figure 2.

To find the required energy to reliably transmit a data packet over the channel we consider Shannon's capacity formula for Gaussian channels.

$D_n(t)$ can be approximately calculated by

$$D_n = W \Delta_{Tx} \log_2 \left(1 + \frac{H_n P}{W N_0} \right), \quad (1)$$

where P is the transmit power and N_0 is the noise power density.

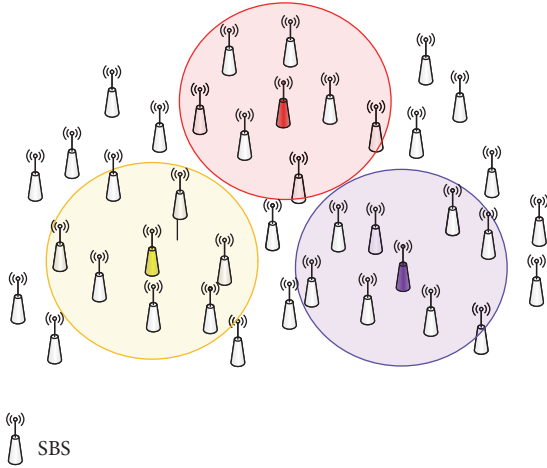


FIGURE 1: System model.

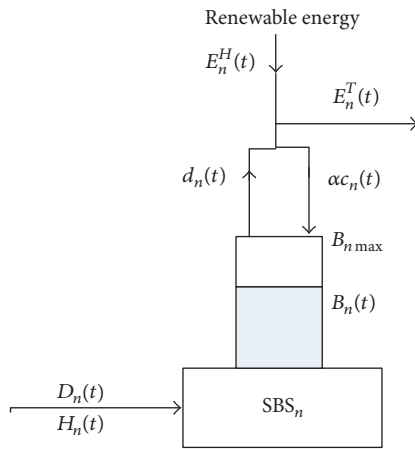


FIGURE 2: SBS energy harvesting model in UDN.

$E_n^T(t)$ is an integer multiple of the energy unit [22] and can be calculated by

$$E_n^T = f_e(D_n, H_n) = \frac{D_n \log_2 N_0}{H_n}. \quad (2)$$

In this paper, we define that EE in terms of bits/Hz/Joule is

$$EE = \frac{\sum_{k=1}^K \log_2(1 + p_k * g_k)}{P_f + P_c + \varphi \sum_{k=1}^K P_k}, \quad (3)$$

where $\sum_{k=1}^K p_k$ is the sum power consumption of K SBSs in one cluster, g_k denotes the channel-gain-to-noise ratio, P_f denotes the addition power dissipation due to SBS's circuitry, P_c is the static power dissipation, and φ is the power amplifier efficiency.

A multiuser scenario can be derived similarly. At the beginning of each TS, each SBS schedules its users, such that a single user is allocated to each subcarrier. Let $\mu_{u,m} = 1$

denote that subcarrier m is assigned to user u , and $\mu_{u,m} = 0$, otherwise. Accordingly, EE can be expressed as

$$EE = \frac{\sum_{u=1}^U \sum_{m=1}^M (u_{u,m}/M) \log_2(1 + p_{u,m} * g_{u,m})}{P_f + P_c + \varphi \sum_{u=1}^U \sum_{m=1}^M P_{u,m}}, \quad (4)$$

where U is the user number assigned for one SBS, M is the subcarrier number, and $p_{u,m}$ denotes power allocation.

$$g_{u,m} = \frac{h_{m,u}^2}{N_0 (W/M)}, \quad (5)$$

where $h_{m,u}$ is the channel gain, N_0 is the noise power spectral density of additive white Gaussian noise, and W is the bandwidth.

3. Energy Cooperation Optimization Algorithm Based on SBS Clustering and Learning Strategy

This section optimizes the problem of energy cooperation among SBSs in UDN using clustering technology and reinforcement learning.

3.1. Dynamic k-Means Clustering Method for SBSs in UDN. In this paper, a cluster-based approach for maximizing EE in UDN is proposed. According to the different realization process, the common clustering methods are divided into hierarchical clustering and partition clustering. There are also clustering methods for large scale data sets. One of the most classic clustering methods is k -means, which uses cluster centers to represent the whole group. The cluster centers need to be updated repeatedly in the process of determining the final result [29, 30]. All data should be redivided before updating cluster centers, which makes k -means lose the ability to deal with very large scale data, and the execution time of these algorithms grows with the data number increases.

Since there are more SBSs than the traditional network in UDN, it is very time-consuming to run the clustering algorithm on the whole data set. To remedy this problem, sampling strategy is selected, which can greatly save storage space and reduce computation amount. Sampling technology is used to select some samples from the original data. According to the cluster result, the distribution of the original data set is estimated. In this paper, the sampling technique is used to determine the centers of data division. If the distance between SBSs is too long, energy cooperation may cause power loss and transmission delay. We cluster SBSs in UDN according to distance and traffic. The sampled SBS is selected according to SBS traffic from large one to small, and the SBS traffic refers to the average data packets sent in the past week or month. Each SBS is viewed as a data point in a two-dimensional space, and the distances between sampling SBSs are saved [31, 32].

Assume that there are m SBSs in UDN. The distances between all SBSs are saved in the matrix M_{m*m} . The dynamic k -means algorithm is described as follows:

- (1) Select r sampling SBSs (according to SBS traffic from large to small) and save their distances in the matrix $D_{r \times r}$; each column represents the distances between one SBS and other sampling SBSs, $D_{r \times r} = d(x_i, x_j)$ ($i, j = 1, 2, \dots, r$), where $d(x_i, x_j)$ represents the distances between sampling SBSs.
- (2) Find the minimum distance between sampling SBSs, $\min d(x_i, x_j)$ ($i \neq j$).
- (3) Calculate the average value of each column $R_i = \sum_{j=1}^r d(x_i, x_j)/(T-1)$ ($i = 1, 2, \dots, r$), where $r-1$ represents each column vector in the distance matrix $D_{r \times r}$ which has a "0" element, which is the distance between each point and its own.
- (4) Calculate the average value of all R_i , $\bar{R} = \sum_{i=1}^r R_i/r$.
- (5) Calculate high density radius $R = \bar{R} + \min d(x_i, x_j)$; $\min d(x_i, x_j)$ is added to make R big enough so as to ensure that most high density points are correctly labeled.
- (6) Calculate cluster radius r according to $r = R + \min d(x_i, x_j)$ and select the two furthest points from sampling SBSs as the initial cluster centers and mark $k = 2$ to ensure that the centers come from different clusters.
- (7) Divide the data near the centers into two clusters according to r , find the next farthest point according to the centers, mark $k = k + 1$, and divide the data again until all the data are clustered completely.
- (8) Cluster the remaining SBSs into the nearest center point.
- (9) Calculate $m = r + r$, and $h = \min d(x_i, x_j) + \min d(x_i, x_j)$. If the distance between the cluster centers is smaller than m and the distance between the boundary points is smaller than h , then combine the two clusters and label $k = k - 1$.
- (10) Give the final clustering results and the k value.

How to determine k value is a very difficult problem. Once the k value is not reasonable, it is likely to lead to great errors in clustering. For this case, according to the data distribution properties and their distances, after a series of transformations, the final clustering number is obtained. Obtaining k value is a changing, dynamic process, and there is no need to know initial experience value. The dynamical division is closer to actual demand. It solves the problem that the algorithm needs manual input and improves the automatic clustering ability.

EE is considered to be one of the main benefits from clustering architecture. When the network is dense, the benefit of clustering is the improved stability of cooperative relationships. Clusters are formed in order to maximize the time availability of clusters, hence in this paper maximizing the availability of energy cooperative. The clustering centers are all SBSs of large traffics, which is benefit for energy cooperation among SBSs of small traffics. It also ensures the supply of renewable energy and improves the data transmission rate of the network.

TABLE I: Actions for SBS processing each data packet.

Actions	Transmitted/dropped and achieve/no energy cooperation
a_1	Transmit the data packet and achieve energy cooperation
a_2	Transmit the data packet and have no energy cooperation
a_3	Drop the data packet and achieve energy cooperation
a_4	Drop the data packet and have no energy cooperation

3.2. Q-Learning Approach for Energy Cooperation. Due to the instability of the renewable energy and the arbitrary distribution of SBSs, it is necessary to improve EE through energy cooperation. In this paper, we propose a renewable energy cooperation scheme among different SBSs, in which one SBS can collect/share energy from/to another SBS. The energy cooperation efficiency will be improved when the number of the cooperation SBSs is large.

Consider energy cooperation among SBSs in UDN as the finite states and discrete time MDP. $(S, A, P_{a_t}(s_t, s_{t+1}), R_{a_t}(s_t, s_{t+1}))$, where S is the finite environment state space; A is the finite system action space; $P_{a_t}(s_t, s_{t+1}) \in [0, 1]$ and $R_{a_t}(s_t, s_{t+1})$, respectively, represent the state transition probability and the immediate reward of transferring the state from s_t to s_{t+1} by taking action a_t . The probability and the immediate reward depend only on the current state s_t and the selected action a_t and are irrelevant to the past states and actions.

In the proposed model, the system state of n th SBS in TS t is $S_n = (E_n^H(t), E_n^T(t), H_n(t), B_n(t))$, and action set is $A = \{a_1, a_2, a_3, a_4\}$. At the beginning of each TS, SBS has four actions for each data packet, which is shown in Table 1.

Assume that the optimal state value function and the optimal action value function are V^* and Q^* , which satisfy the Behrman optimal equation:

$$\begin{aligned}
 V^*(s) &= \max_{\pi} V^{\pi}(s) \\
 &= \max_{a \in A} \sum_{s' \in S} P_a(s, s') (R_a(s, s') + \gamma V^*(s')). \\
 Q^*(s, a) &= \max_{\pi} Q^{\pi}(s, a) \\
 &= \sum_{s' \in S} P_a(s, s') \left(R_a(s, s') + \gamma \max_{a' \in A} Q^*(s', a') \right)
 \end{aligned} \tag{6}$$

As a result, the optimal policy can be obtained:

$$\begin{aligned}
 \pi^*(s) &= \arg \max_{a \in A} \sum_{s' \in S} P_a(s, s') (R_a(s, s') + \gamma V^*(s')) \\
 \pi^*(s) &= \arg \max_{a \in A} Q^*(s, a).
 \end{aligned} \tag{7}$$

The goal of MDP is to find the system's optimal policy π^* , which can be obtained by the optimal value function [33].

TABLE 2: Four choices of a_t and the corresponding r_t .

a_t	Corresponding r_t
a_1	Transmit the data packet and achieve energy cooperation $r_t = E_n^T(t) + \beta \left(E_n^H(t) - E_n^T(t) - \frac{B_n(t) - B_n(t-1)}{\alpha} \right)$
a_2	Transmit the data packet and have no energy cooperation $r_t = E_n^T(t)$
a_3	Drop the data packet and achieve energy cooperation $r_t = \beta \left(E_n^H(t) - \frac{B_n(t) - B_n(t-1)}{\alpha} \right)$
a_4	Drop the data packet and have no energy cooperation $r_t = 0$

Q-learning algorithm is adopted, and its iterative formula is as follows:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \delta \left[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right], \quad (8)$$

where (s_t, a_t) is a state-action pair in TS t , δ ($0 < \delta < 1$) is the learning factor, and γ ($0 < \gamma < 1$) is the discount factor.

Q iterative learning uses the reward of state-action pair as the evaluation function. First initialize Q value, then determine the action a_t in the state s_t according to ε greedy policy, get the knowledge and experience of training samples (s_t, a_t, s_{t+1}, R_t) , and then modify Q value. When agent gets the target state, the algorithm terminates one iteration loop. The algorithm starts from the initial state again until the end of learning.

Q-learning algorithm is applied in renewable energy cooperation within one SBS cluster based on UDN, and its process is as follows:

- (1) Initialize: $Q(s, a) \leftarrow$ any value, δ and γ : given values.
- (2) Repeat.

Given initial state S: $(E_n^H(0), E_n^T(0), H_n(0), B_n(0))$.

Repeat.

- ① Choose a_t according to ε greedy policy, then obtain R_t and s_{t+1} .
- ② $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \delta [r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$.
- ③ $s_t \leftarrow s_{t+1}$.

Until s_t is termination state.

Until all $Q(s, a)$ are converged.

- (3) Output the final policy: $\pi(s) = \arg \max_{a \in A} Q(s, a)$.

In ①, a_t has four choices and the corresponding r_t shown in Table 2 (α represents storage efficiency of the battery, and β represents resistance loss).

Actually, four actives can be combined together:

$$r_t = A \times E_n^T(t) + B \times \beta \left(E_n^H(t) - E_n^T(t) - \frac{B_n(t) - B_n(t-1)}{\alpha} \right). \quad (9)$$

TABLE 3: Values of A and B according to different actives.

Action	A	B	r_t
a_1	1	1	$r_t = E_n^T(t) + \beta \left(E_n^H(t) - E_n^T(t) - \frac{B_n(t) - B_n(t-1)}{\alpha} \right)$
a_2	1	0	$r_t = E_n^T(t)$
a_3	0	1	$r_t = \beta \left(E_n^H(t) - \frac{B_n(t) - B_n(t-1)}{\alpha} \right)$
a_4	0	0	$r_t = 0$

Their values are listed in Table 3.

The reward function is actually the sum of the energy for sending data package and the energy for cooperation. In each TS, it encourages sending data and carrying out energy cooperation.

In Q-learning algorithm, actions with the highest Q values at a particular state should be taken at each step. The agent who rigidly follows this rule might underperform since the same decision will be investigated over and over again. In order to be exploited, the state-action pair needs to be explored firstly [34]. In this paper, we utilize ε greedy policy ($0 < \varepsilon < 1$), which is commonly used during the process of state-action space exploration. It enforces sporadic jumps to suboptimal states for the exploration purposes, but also to detect changes of the environmental conditions. Whenever a decision is to be made, the one will be picked at random with the $(1 - \varepsilon)$ probability, which is given to the action with the highest Q value.

Q-learning converges to the optimal Q function. While $t \rightarrow \infty$, $Q(s, a)$ converges to $Q^*(s, a)$ with probability 1 [35], and the convergence rate is related to many factors. The convergence rate increases with the value of δ and the number of learning iterations N_L and decreases with the number of a , s , and γ [36]. Action selection follows the ε greedy policy with probability $(1 - \varepsilon)$ at each TS. The exploration probability is ε , and the exploitation probability is $(1 - \varepsilon)$.

4. Numerical Simulations

According to the traffic amount of SBS from large to small, their relative positions in UDN are marked as Table 4. All SBSs satisfy the constraints (C1).

20 SBSs are, respectively, $[x_i, y_i]$, $i = 1, 2, \dots, 20$, which are denoted by i th data point. 20 SBSs are clustered according to dynamic k -means clustering method mentioned in Section 3.1. The sampling rate is 50%; that is, the (1-10) data points are selected. As shown in Figure 3, the cluster radius is $r = 50.7556$, the number of cluster is $k = 3$, and the cluster centers are $[x_3, y_3]$, $[x_5, y_5]$, and $[x_6, y_6]$, which are denoted as \oplus . The three final clusters are $\{3^\oplus, 1, 4, 7, 8, 10, 17\}$, $\{5^\oplus, 2, 9, 12, 13, 14, 15, 16, 19\}$, and $\{6^\oplus, 11, 18, 20\}$.

Through dynamic k -means clustering, without knowing the k value before, SBSs that have large business amount can be selected as the cluster centers. There are relatively larger energy demands for cluster centers, which can effectively improve the utilization rate of energy. In addition, Q-learning

TABLE 4: The positions of SBS.

x_i	58	42	52	33	43	22	58	76	53	64	21	38	79	68	46	57	80	5	60	5
y_i	69	21	84	63	13	20	61	63	37	58	45	4	2	31	1	38	68	9	3	61

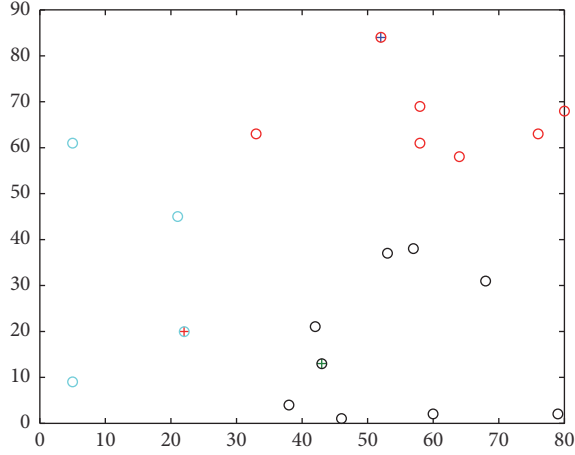
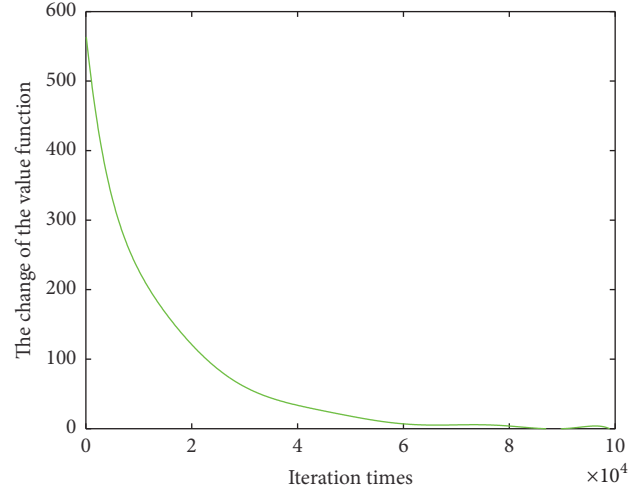
FIGURE 3: The dynamic k -means clustering result for SBSs in UDN.

FIGURE 4: Algorithm convergence process.

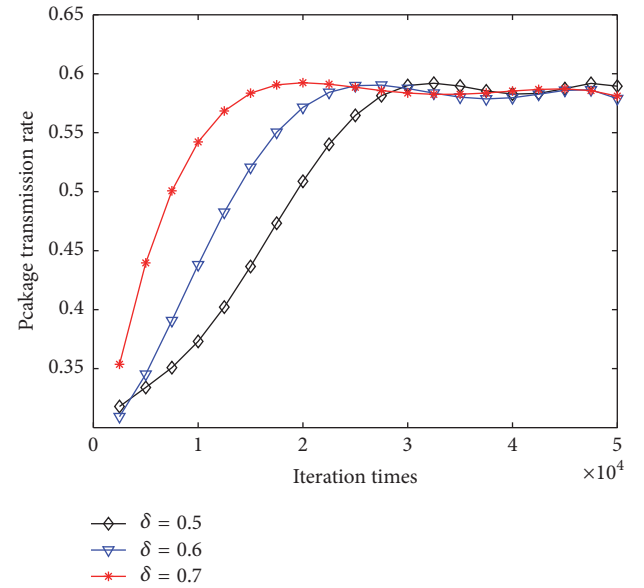
TABLE 5: System parameters.

Parameters	Values
W	2 MHz
Δ_{TS}	10 ms
N_0	$10^{-20.4}$ w/Hz
h_1	$1.655 * 10^{-13}$
h_2	$3.311 * 10^{-13}$
P_{H_n}	0.9
B_{max}	5
γ	0.9

in this paper only considers the energy cooperation between the same clusters, which can reduce the dimension disaster problem caused by too many states.

In the numerical analysis, we take one class $\{\sigma^{\oplus}, 11, 18, 20\}$ as an example. All parameters are based on an IEEE802.15.4e [36] communication system. The system parameters are listed in Table 5. Each time-slot Δ_{TS} is 10 ms, in which 5 ms is used to send data [30] and 5 ms to zero signal level. The channel state at TS n is H_n , $H_n \in \{h_1, h_2\}$. The channel state transition probability function is characterized by $P_{H_n}(h_1, h_1) = P_{H_n}(h_2, h_2) = 0.9$. The transmit power is P . We consider Shannon's capacity formula for Gaussian channels. The battery capacity is 5, that is, $B_n = \{0, 1, 2, 3, 4, 5\}$. Each basic energy unit is $2.5 \mu j$ [37, 38]. The possible data packet sizes are 300 bits or 600 bits.

As shown in Figure 4, Q-learning approach for energy cooperation in one class is convergent. The horizontal axis is the iteration times (the time of one iteration is 10 ms, and it is the same as other figures), and the vertical axis is the difference between the adjacent two sampling value functions. The line uses 5-degree polynomial fitting, and the

FIGURE 5: Package transmission rate for $\delta = \{0.5, 0.6, 0.7\}$.

sampling interval is $100 \mu s$ (it is the same as in Figures 4–10). The change of the value function proves the algorithm's convergence. When the iteration number (N_L) reaches about 6×10^4 , the function value is basically unchanged.

In Figure 5 it shows that the learning factor δ in Q-learning algorithm can influence the iteration times when the packet transmission rate reaches stable. When $\delta = 0.5$ and $N_L \approx 3.0 \times 10^4$, the black diamond line is basically unchanged; when $\delta = 0.6$ and $N_L \approx 2.5 \times 10^4$, the triangle blue line reaches stable; when $\delta = 0.7$ and $N_L \approx 1.5 \times 10^4$, the red star line keeps

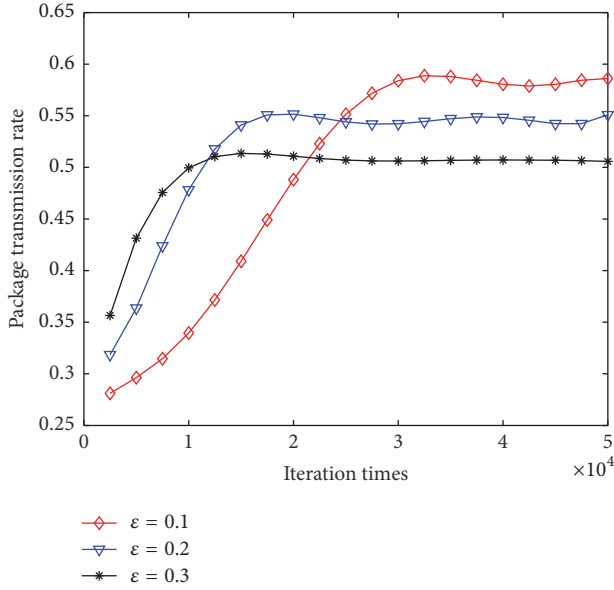


FIGURE 6: Package transmission rate for $\epsilon = \{0.1, 0.2, 0.3\}$.

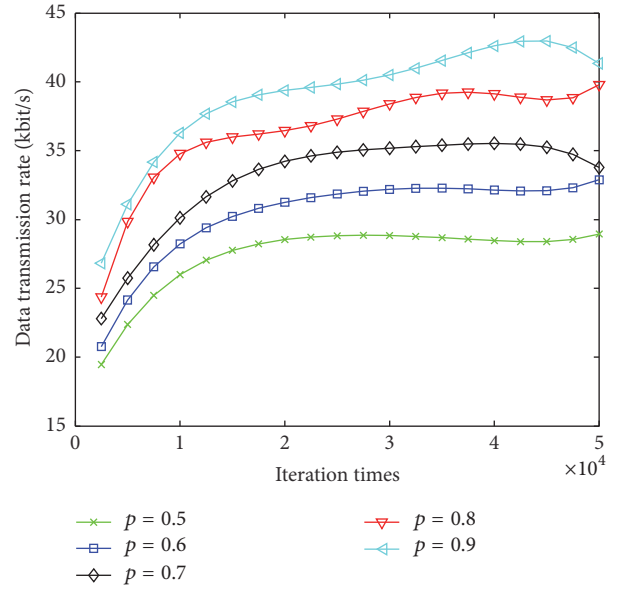


FIGURE 8: Data transmission rate for $P_{H_n} = \{0.5, 0.6, 0.7, 0.8, 0.9\}$ and $B_{\max} = 5$.

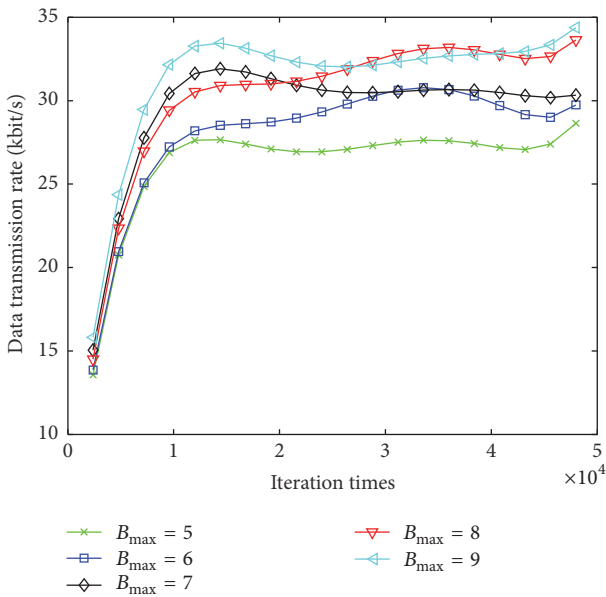


FIGURE 7: Data transmission rate for $B_{\max} = \{5, 6, 7, 8, 9\}$ and $P_{H_n} = 0.9$.

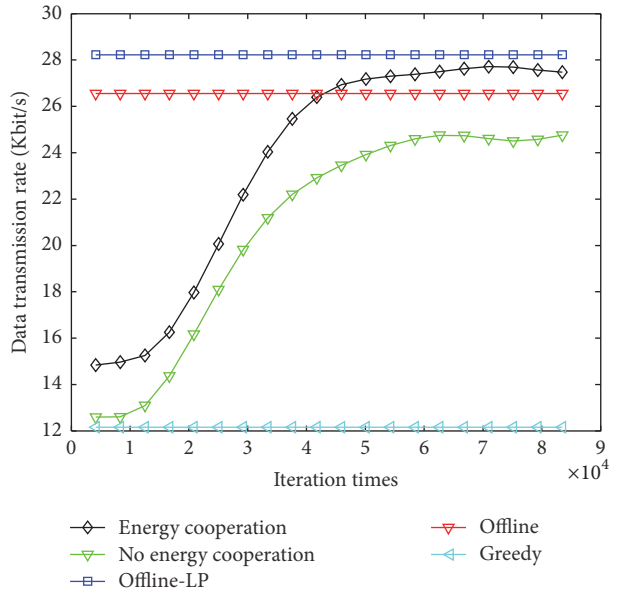


FIGURE 9: Data transmission rate with and without energy cooperation.

stable. We can conclude that Q-learning algorithm can keep stable with smaller iteration times as δ increases.

As shown in Figure 6, ϵ greedy policy can ensure the convergence of Q-learning algorithm. When ϵ uses different values, it can influence the final package transmission rate. When $\epsilon = 0.1$, the package transmission rate reaches about 60% as N_L increases; when $\epsilon = 0.2$, it reaches about 55%; when $\epsilon = 0.3$, it reaches about 50%. We can conclude that Q-learning algorithm is able to learn the optimal policy with increasing accuracy as ϵ decreases.

In Figure 7 we show the effect of the max battery size B_{\max} on the expected data transmission rate. We can conclude that

the expected data transmission rate increases with B_{\max} for the proposed algorithm. In our model, the system state of n th SBS in TS t is $S_n = (E_n^H(t), E_n^T(t), H_n(t), B_n(t))$, in which all the parameter values are finite and discrete. If B_{\max} has a relatively larger value, then $B_n(t) \in \{0, 1, \dots, B_{\max}\}$ has more choices, which leads to increased computation and dimension disaster in Q learning. In this paper we make $B_{\max} = 5$ to simplify the algorithm, but we have executed exhaustive numerical simulations with different parameter settings and observed similar results.

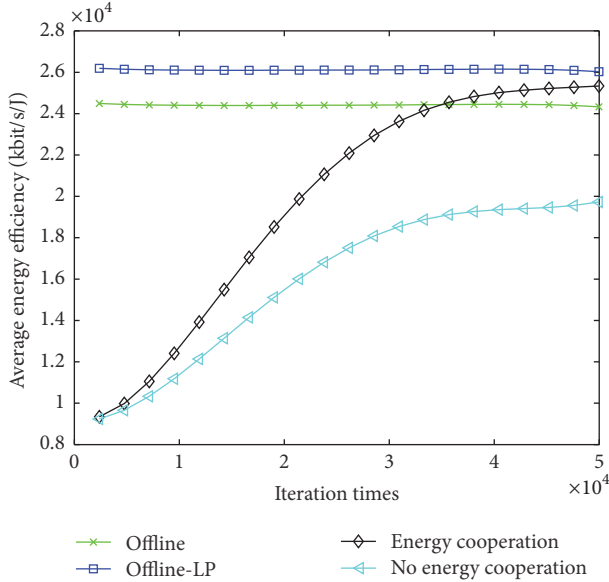


FIGURE 10: Average energy efficiency with and without energy cooperation.

Figure 8 displays the data transmission rate for different P_{H_n} values. We see that the expected data transmission rate increases with P_{H_n} . It means that the more stable the channel state is, the higher the data transmission rate is. In class $\{6^{\oplus}, 11, 18, 20\}$, when $P_{H_n} = 0.5$, the data transmission rate is about 28 kbit/s; when $P_{H_n} = 0.9$, the data transmission rate is about 40 kbit/s. As P_{H_n} increases, EH process becomes less random, and the proposed algorithm can better estimate its future states and adapt to it.

In Figure 9 it shows the relationship between the data transmission rate and iteration times. Energy cooperation can improve the energy utilization; that is, the data transmission rate can be higher through energy cooperation. The black diamond line which represents the rate of energy cooperation reaches about 28 kbit/s, and the green triangular line which represents the rate of no energy cooperation is below 24 kbit/s. The transmission rate through energy cooperation is about 15% higher than no energy cooperation. It is the same with other clusters in UDN. The proposed method in this paper can significantly improve the data transmission rate.

The ratio of the network throughput to the power consumption per unit area is defined as EE (the energy efficiency). The energy efficiency metric is a performance indicator that measures the benefit-cost ratio by comparing the achievable rate to the energy costs. In Figure 10 we illustrate, together with the performance of the other approaches, the expected average energy efficiency by the proposed approach against the number of learning iterations times. It can be observed that the average EE of energy cooperation is higher than that of no energy cooperation. The black curve is more closer to the Offline-LP algorithm. The proposed method can significantly improve energy utilization rate.

5. Conclusion

This paper presents a renewable energy cooperation management algorithm based on cluster and learning strategy in UDN. Firstly, according to the geographical location and traffic load, SBSs are clustered using dynamic k -means algorithm, in which sampling technology is utilized to improve computation speed and clustering effect. Secondly, within each formed cluster, SBSs coordinate their renewable energy. The process of energy cooperation is considered as MDP. Q-learning algorithm is adopted to optimize energy cooperation, which considers four immediate reward functions, and the convergence of the algorithm is realized by ϵ greedy policy. Thirdly, simulation results show that the new method can improve the utilization of renewable energy and promote the data transmission rate. At last, conclusion and future research directions are presented, which include energy cooperation between clusters, combination of renewable energy, smart grid, and so on.

System Parameters in UDN

- TS: Time-slot
- $H(t)$: Channel station information
- $E_n^H(t)$: Energy harvested
- $E_n^T(t)$: Energy for transmitting data
- $D_n(t)$: Data amount to be transmitted
- $B_{n,max}$: The maximum capacity of the battery.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work in this paper is supported by “Beijing Natural Science Foundation (Grant no. 4164101),” “National Natural Science Foundation of China (Grants no. 61501185, no. 61377088),” “Hebei Province Natural Science Foundation (Grant no. F2016502062),” and “the Fundamental Research Funds for the Central Universities (2015MS125, 2016MS97).”

References

- [1] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, “User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions,” *IEEE Wireless Communications*, vol. 23, no. 2, pp. 78–85, 2016.
- [2] S. Tang and L. Tan, “Reward Rate Maximization and Optimal Transmission Policy of EH Device with Temporal Death in EH-WSNs,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1157–1167, 2017.
- [3] M. Virili, A. Georgiadis, F. Mira et al., “EH performance of an hybrid energy harvester for autonomous nodes,” in *Proceedings of the IEEE Topical Conference on Wireless Sensors and Sensor Networks, WiSNet 2016*, pp. 71–74, usa, January 2016.
- [4] R. Ramamonjison and V. K. Bhargava, “Energy allocation and cooperation for energy-efficient wireless two-tier networks,”

- IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6434–6448, 2016.
- [5] N. Reyhanian, B. Maham, V. Shah-Mansouri, W. Tushar, and C. Yuen, “Game-Theoretic Approaches for Energy Cooperation in Energy Harvesting Small Cell Networks,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7178–7194, 2017.
 - [6] Q. Wang, H.-L. Liu, and Y.-M. Cheung, “A renewable energy cooperation scheme for OFDM systems using evolutionary many-objective optimization algorithm,” in *Proceedings of the 12th International Conference on Computational Intelligence and Security, CIS 2016*, pp. 194–197, chn, December 2016.
 - [7] H.-S. Lee and J.-W. Lee, “Energy cooperation and traffic management in cellular networks with renewable energy,” in *Proceedings of the 59th IEEE Global Communications Conference, GLOBECOM 2016*, usa, December 2016.
 - [8] A. Jahid, A. B. Shams, and M. F. Hossain, “Energy cooperation among BS with hybrid power supply for DPS CoMP based cellular networks,” in *Proceedings of the 2nd International Conference on Electrical, Computer and Telecommunication Engineering, ICECTE 2016*, bgd, December 2016.
 - [9] Y. Dong, Z. Chen, and P. Fan, “Capacity region of gaussian multiple-access channels with energy harvesting and energy cooperation,” *IEEE Access*, vol. 5, pp. 1570–1578, 2017.
 - [10] W. Ni and X. Dong, “Energy harvesting wireless communications with energy cooperation between transmitter and receiver,” *IEEE Transactions on Communications*, vol. 63, no. 4, pp. 1457–1469, 2015.
 - [11] J. J. Pradha, S. S. Kalamkar, and A. Banerjee, “On information and energy cooperation in energy harvesting cognitive radio,” in *Proceedings of the 26th IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, PIMRC 2015*, pp. 943–948, chn, September 2015.
 - [12] Z. Xiaorong and Z. Weiran, “Interference coordination-based cell clustering and power allocation algorithm in dense small cell networks,” *Journal of Electronics & Information Technology*, vol. 38, no. 5, pp. 1173–1178, 2016.
 - [13] Z. Fang, Y. Li, H.-T. Li, and Y. Li, “Spatial coordination beamforming for ultra-dense wireless networks,” *Dianzi Keji Daxue Xuebao/Journal of the University of Electronic Science and Technology of China*, vol. 45, no. 2, pp. 185–190, 2016.
 - [14] B. Lu, L. Tingting, and Y. Chenyang, “Interference coordination method and performance analysis in Ultra-dense Network (UDN),” *Journal of Signal Processing*, vol. 31, no. 10, pp. 1263–1271, 2015.
 - [15] C. Guo, C. Feng-en, C. Lei, and G. Yuehong, “Study on the interference coordination technology based on the traffic characteristics in ultra dense network,” *Telecom Engineering Technics and Standardization*, vol. 3, pp. 75–78, 2016.
 - [16] L. Liu, V. Garcia, L. Tian, Z. Pan, and J. Shi, “Joint clustering and inter-cell resource allocation for CoMP in ultra dense cellular networks,” in *Proceedings of the IEEE International Conference on Communications, ICC 2015*, pp. 2560–2564, gbr, June 2015.
 - [17] R. Wei, Y. Wang, and Y. Zhang, “A two-stage cluster-based resource management scheme in ultra-dense networks,” in *Proceedings of the 2014 IEEE/CIC International Conference on Communications in China, ICC 2014*, pp. 738–742, chn, October 2014.
 - [18] M. Kamel, W. Hamouda, and A. Youssef, “Ultra-dense networks: a survey,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522–2545, 2016.
 - [19] L. Liang, W. Wang, Y. Jia, and S. Fu, “A cluster-based energy-efficient resource management scheme for ultra-dense networks,” *IEEE Access*, vol. 4, pp. 6823–6832, 2016.
 - [20] K. Rahbar, C. C. Chai, and R. Zhang, “Real-time energy management for cooperative microgrids with renewable energy integration,” in *Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm '14)*, pp. 25–30, IEEE, November 2014.
 - [21] D. W. K. Ng, E. S. Lo, and R. Schober, “Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3412–3427, 2013.
 - [22] Y.-K. Chia, S. Sun, and R. Zhang, “Energy cooperation in cellular networks with renewable powered base stations,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6996–7010, 2014.
 - [23] Z. Ren, S. Chen, B. Hu, and W. Ma, “Energy-efficient resource allocation in downlink OFDM wireless systems with proportional rate constraints,” *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2139–2152, 2014.
 - [24] E. Boshkovska, A. Koelpin, D. W. K. Ng, N. Zlatanov, and R. Schober, “Robust beamforming for SWIPT systems with non-linear energy harvesting model,” in *Proceedings of the 17th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2016*, gbr, July 2016.
 - [25] S. Maghsudi and E. Hossain, “Distributed user association in energy harvesting small cell networks: a probabilistic bandit model,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1549–1563, 2017.
 - [26] S. Zhao, T. Zhang, and X. Zhu, “A handover algorithm based on prediction of adjustable threshold hysteresis margin in ultra dense network,” *Dianzi Yu Xinxu Xuebao/Journal of Electronics and Information Technology*, vol. 38, no. 3, pp. 649–654, 2016.
 - [27] W. Junxuan, T. Shiyang, and S. Changyin, “Resource allocation based on user clustering in ultra-dense small cell networks,” *Journal of XI'AN University of Posts and Telecommunications*, vol. 21, no. 1, pp. 16–20, 2016.
 - [28] P. Blasco, D. Gunduz, and M. Dohler, “A learning theoretic approach to energy harvesting communication system optimization,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1872–1882, 2013.
 - [29] F. Jinmei, “Research on large scale data clustering analysis methods,” *Harbin: Harbin Engineering University*, pp. 56–73, 2015.
 - [30] T. A. Hearn and L. Reichel, “Fast computation of convolution operations via low-rank approximation,” *Applied Numerical Mathematics*, vol. 75, pp. 136–153, 2014.
 - [31] T. Sen-ping and W. Wen-liang, “Algorithm of automatic gained parameter value k based on dynamic k-means,” *Computer Engineering and Design*, vol. 32, no. 1, pp. 274–276, 2011.
 - [32] B. Zhen-shan, X. Bo, and Z. Wen-bo, “Improved LEACH algorithm based on k-means in clustering,” *Computer Science*, vol. 40, no. 10, pp. 215–218, 2013.
 - [33] S. Ruoying and Z. Gang, “Reinforcement learning and coordination in multiagent systems,” *Tsinghua University Press*, pp. 1–19, 2014.
 - [34] J. N. Tsitsiklis, “Asynchronous Stochastic Approximation and Q-Learning,” *Machine Learning*, vol. 16, no. 3, pp. 185–202, 1994.
 - [35] C. Watkins, “Learning from delayed rewards,” *Royal Holloway, University of London*, pp. 220–228, 1989.

- [36] E. Even-Dar and Y. Mansour, "Learning rates for Q-learning," *Machine Learning Research*, no. 5, pp. 1–25, 2003.
- [37] IEEE 802.15.4e Draft Standard: Wireless Medium Access Control(MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs), IEEE Std., Mar. 2010.
- [38] S. Chalasani and J. M. Conrad, "A survey of energy harvesting sources for embedded systems," in *Proceedings of the IEEE Southeastcon*, pp. 442–447, Huntsville, Ala, USA, April 2008.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

