

Global Optimization for Energy Efficient Resource Management by Game Based Distributed Learning in Internet of Things

ChunHua Ju¹, Qi Shao^{*1}

¹ College of Computer Science & Information Engineering, Zhejiang Gongshang University,
No.2 Shuguang Street, Hangzhou, China
[E-mail: juchunhua@hotmail.com, shaoqi2009@126.com]

*Corresponding author: Qi Shao

*Received March 22, 2015; revised June 25, 2015; revised July 20, 2015; accepted August 2, 2015;
published October 31, 2015*

Abstract

This paper studies the distributed energy efficient resource management in the Internet of Things (IoT). Wireless communication networks support the IoT without limitation of distance and location, which significantly impels its development. We study the communication channel and energy management in the wireless communication network supported IoT to improve the ability of connection, communication, share and collaboration, by using the game theory and distributed learning algorithm. First, we formulate an energy efficient neighbor collaborative game model and prove that the proposed game is an exact potential game. Second, we design a distributed energy efficient channel selection learning algorithm to obtain the global optimum in a distributed manner. We prove that the proposed algorithm will asymptotically converge to the global optimum with geometric speed. Finally, we make the simulations to verify the theoretic analysis and the performance of proposed algorithm.

Keywords: game theory, wireless networks, learning algorithm, Internet of Things

Research is supported by The National Key Technology R and D Program of China (Grant 2014BAH24F06), Natural Science Foundation of Zhejiang Province (Grants LY14F020002, LY15G010001), scientific innovation of university students of Zhejiang Province (2015R408079)

1. Introduction

The Internet of Things (IoT) has drawn much attention from academia as well as industry recently. In 1999, Kevin Ashton proposed the term IoT. With the development of internet and wireless communication, the IoT has been seen as an important technological revolution that brings us into a new ubiquitous connectivity, computing, and communication era [1]. The possible applications of IoT are significant, such as climate monitoring, transport safety, home automation, health care, supply chain, agriculture, rural development, border security and military application. In IoT, the objects (Obs, including people and things) are expected to have the capabilities of sensing the physical objects, communication, and operating smartly without human intervention. The major technologies that would dominate IoT applications include wireless sensor networks (WSN) and mobile communications [2].

Collaboration is an important feature of IoT. The IoT is widely regarded as global network which allows communication and collaboration between people and things, and between things themselves [3]. Wire connected Obs could share information via internet [4], however, much more Obs need wireless connection to overcome limitation of distance and location. The fast wireless communication could support the widely collaboration in the IoT through wireless networks, which significantly impels the development of IoT. Researches on the wireless communication based collaboration in the IoT should be paid much attention to.

In fact, collaboration in IoT needs huge information exchange, which would consume resources, such as communication channel, bandwidth and energy. In practical collaboration via wireless network, the channel resource is very important and limited, then the channel resource management is an important problem. In addition, energy is very sensitive to the IoT [3], especially for the sensors.

There are some researches on the technologies in IoT. Authors in [5-7] focused on the surroundings sensing in IoT based on the sensor network technologies. The identification in IoT was studied in [8-10]. Authors in [11-12] studied the communication protocols. The tracking effects and energy consumption optimization was studied in [13]. In all, to the best of our knowledge, there are only limited researches which focused on the energy optimization in IoT, and channel resource management in the IoT has not been paid attention to.

Considering the importance of energy sensitive channel resource management in the collaboration via wireless network, in this paper, we focus on the wireless channel selection for the Obs, towards the energy efficiency. Due to the dense and dynamically deployment of objects in the IoT [14], the centralized control in network management will be highly inefficient for the high complexity.

The game theory [15] is a powerful tool to study the interactions among multiple users and has been used in the distributed optimization in many researches about the distributed

network [16-17]. For some good properties to the distributed global optimization, the potential game was also applied in some researches. In [18], the authors proposed a multi-cell coordination approach to mitigate the mutual interference among base stations in the frequency slotted cellular networks. Authors in [19] proposed local altruistic game and local congestion game which belong to local interaction game in cognitive radio networks. In [20], the authors investigated the problem of joint base station selection and resource allocation in an orthogonal frequency division multiple access heterogeneous cellular network, analyzed this problem by using potential game theoretic approaches, and proposed two different variants of Max-logit learning algorithms which achieved outstanding performances.

For the IoT, there are only few game based researches to our best knowledge. Authors in [14] proposed an energy aware trust derivation scheme using game theoretic approach, which managed overhead while maintaining adequate security of WSNs. In [3], a service providing model was built by using a differential game model. The game solution was gotten in the condition of grand coalition, feedback Nash equilibrium and intermediate coalitions and an allocation policy was obtained by Shapley theory. Nevertheless, with the distributed global energy consumption optimization in IoT which we focus on, related studies are very limited.

To sum up, the multi-Obs energy consumption optimization is an important and challenging problem in IoT, and the distributed optimization approach is potential but not paid much attention to until now. In this paper, we solve the challenging multi-Obs energy efficient channel selection problem in IoT in the distributed optimization approach to achieve global optimum, by using the potential game theory and distributed learning algorithm. First, we formulate an energy efficient neighbor collaborative game model and prove that the proposed game is an exact potential game. Second, we design a distributed energy efficient channel selection learning algorithm to obtain the global optimization in a distributed manner. Thirdly, we analyze the converging speed and the computing complexity of the proposed algorithm. We prove that the proposed algorithm will asymptotically converge to the global optimum with geometric speed. Finally, we make the simulations to verify the theoretic analysis and the performance of proposed algorithm.

The rest of this paper is organized as follows: In Section II, we present the system model and problem formulation. In Section III, we formulate the proposed game model, and investigate its properties. In Section IV, we propose the distributed learning algorithm, analysis the convergence of the algorithm, the converging speed and the computing complexity. In Section V, simulation results and discussion are presented. Finally, we provide conclusions in Section VI.

2. System Model and Problem Formulation

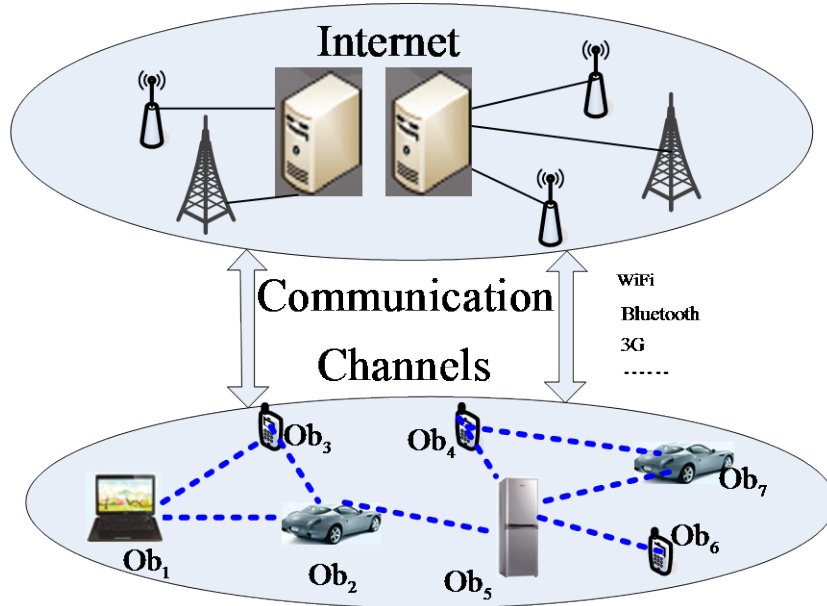


Fig. 1. System model

We consider a local interactive network [21] consisting of M objects (Obs), which are connected via various wireless communication technologies and the internet. Obs could transmit their data via 3G mobile communication channels, WiFi channels, Bluetooth and so on. Generally speaking, Obs use wireless frequency channels to communicate. Local interactive is a feature of the IoT that the Obs are always spatially distributed and the effect of each Ob's action is limited between the neighboring Obs [1]. For example, in Fig. 1, the data transmission of Ob_2 and Ob_3 would collide if they use the same frequency channel for their neighboring relationship. Ob_3 and Ob_6 could use the same frequency channel without collision because they are not neighbors.

Importantly, because the qualities of wireless channels are related to the location and channel frequency, choosing different frequency channel means different energy consumption. Denote $S_{CH} = \{1, 2, \dots, N\}$ as the set of channels, and denote $S_{Ob} = \{1, 2, \dots, M\}$ as the set of Obs. If Ob_m chooses channel n , according to the Shannon equation, the obtained capacity would be given by:

$$R_m = W_n \log \left(1 + \frac{P_m \eta_n}{W_n N_0} \right) = W_n \log \left(1 + \frac{P_m d_m^{-\gamma_n} \theta_{n,m}}{W_n N_0} \right) \quad (1)$$

where W_n is the bandwidth of channel n , N_0 is the noise power spectrum density, P_m is

the transmitting power of Ob_m , d_m is the distance between Ob_m and its destination of transmitting data. γ_n is the path loss exponent of channel n , and $\theta_{n,m}$ is the instantaneous random component of the path loss [22] on channel n for Ob_m .

Then the transmission power of Ob_m should be given by:

$$P_m = \frac{W_n N_0 \left(\exp^{(R_m/W_n)} - 1 \right)}{d_m^{-\gamma_n} \theta_{n,m}} \quad (2)$$

Due to the lack of central control, collisions may occur when there may be more than one Obs choose the same channel in local region. The typical slotted Aloha transmission mechanism is adopted to reduce the collision, where each Ob accesses the channel with probability α . The data transmission of Ob_m on channel n would be successful when no other neighboring Obs attend to access this channel. Denote Ω_m as the set of Ob_m 's neighbors (including Ob_m itself), and $|\Omega_m|$ is the number of Ob_m 's neighbors. The probability of successful transmission is given by:

$$\varepsilon_s^n = \alpha (1 - \alpha)^{|\Omega_m|-1} \quad (3)$$

For each Ob, the higher probability of successful transmission and also the lower transmission power are expected. We define the reward of Ob, e.g., Ob_m as follow:

$$\phi_m = \varepsilon_s^n / P_n = \frac{\alpha (1 - \alpha)^{|\Omega_m|-1} d_m^{-\gamma_n} \theta_{n,m}}{W_n N_0 \left(\exp^{(R_m/W_n)} - 1 \right)} \quad (4)$$

In the perspective of network, the optimization objective is to maximize sum of Ob's utilities:

$$\text{Problem: max } \phi_{net} = \sum_{m=1}^M \phi_m \quad (5)$$

Remark: From the aspect of each Ob, there is a tradeoff for between channel quality and collision probability when making the decision in the network consist of multi Obs. Good channels, i.e., channels which consume lower power, are more likely to be chosen by other Obs, which reduces to higher probability of collision. From the aspect of network, the optimization of network utility and stability are expected. The global optimization of network utility could be obtained in a centralized manner by the exhaustive search. However, the computation complexity is extremely high. For example, even in a relative small scenario where $N=5$ and $M=30$, the number of possible Obs' strategy profiles is $5^{30} = 9.3132 \times 10^{20}$. The desired method is the one which obtains the global optimization in a distributed manner to low the computation complexity. We solve this problem by using the game theory [15], which is one of the important mathematical tools that analyze interactions among multi objects.

3. The Energy Efficient Neighbor Collaborative Game

In this section, we study the distributed optimization of the formulated energy efficiency problem by using game theory. Every Ob is regarded as a player in the game, and we define the energy efficient neighbor collaborative game (EENCG) as:

$$G = \{S_{Ob}, S_{CH}, A_m, u_m\} \quad (6)$$

where S_{CH} is the set of channels, S_{Ob} is the set of Obs, A_m is the available strategy set of $Ob_m \in S_{Ob}$, where $a_m \in A_m$ is the chosen action of Ob_m . u_m is the utility function of player Ob_m , and $u_m(a_m, a_{-m})$ denotes the Ob_m 's utility where a_m is the action of player Ob_m and a_{-m} is the action profile of other players.

Based on the feature of collaboration in the IoT [3], motivated by local collaboration in biographical systems [30], [31] and the collaboration design in networks [18-20], [23-24], [27], we define the following utility function of player, e.g., Ob_m , to obtain the global optimization in a distributed decision manner.

$$u_m(a_m, a_{-m}) = \sum_{Ob_i \in \Omega_m} \phi_i \quad (7)$$

where ϕ_i is the reward of player Ob_i and Ω_m is the set of neighbors of Ob_m .

Definition 1: Denote an action profile of players as $\mathbf{a} = \{a_1, a_2, \dots, a_M\}$, and an action profile $\mathbf{a}^* = \{a_1^*, a_2^*, \dots, a_M^*\}$ is a pure strategy Nash equilibrium (NE) if and only if no player could improve its utility by deviating unilaterally, i.e.,

$$u_m(a_m^*, a_{-m}^*) \geq u_m(a_m, a_{-m}^*) \\ \forall Ob_m \in S_{Ob}, \forall a_m \in A_m, a_m \neq a_m^* \quad (8)$$

Theorem 1: the global optimal solution to the proposed energy efficient channel selection problem (5) is a pure strategy NE of the proposed EENCG game.

Proof: The following proof is based on the potential game theory [23]. We define the network utility as the potential function of the EENCG:

$$\Gamma(a_m, a_{-m}) = \phi_{net} = \sum_{Ob_i \in S_{Ob}} \phi_i = \sum_{Ob_i \in \Omega_m} \phi_i + \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i \quad (9)$$

When Ob_m unilaterally changes its action from a_m to a_m' and other Obs hold their original strategies, the value of potential function would be changed:

$$\begin{aligned}
\Delta\Gamma &= \Gamma(a_m, a_{-m}) - \Gamma(a'_m, a_{-m}) \\
&= \sum_{Ob_i \in \Omega_m} \phi_i(a_m) + \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a_m) - \left(\sum_{Ob_i \in \Omega_m} \phi_i(a'_m) + \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a'_m) \right) \quad (10) \\
&= \sum_{Ob_i \in \Omega_m} \phi_i(a_m) - \sum_{Ob_i \in \Omega_m} \phi_i(a'_m) + \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a_m) - \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a'_m)
\end{aligned}$$

Note that the mutual influence would not occur when the Obs are not neighbors (Fig. 1). Then for $Ob_i \in \{S_{Ob} \setminus \Omega_m\}$, which is not the neighbor of Ob_m , there will be no influence from the Ob_m 's action change, i.e.,

$$\sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a_m) - \sum_{Ob_i \in \{S_{Ob} \setminus \Omega_m\}} \phi_i(a'_m) = 0 \quad (11)$$

Then we have

$$\begin{aligned}
\Delta\Gamma &= \Gamma(a_m, a_{-m}) - \Gamma(a'_m, a_{-m}) = \sum_{Ob_i \in \Omega_m} \phi_i(a_m) - \sum_{Ob_i \in \Omega_m} \phi_i(a'_m) \\
&= u_m(a_m, a_{-m}) - u_m(a'_m, a_{-m}) = \Delta u \quad (12)
\end{aligned}$$

The result shows that the change in the potential function equals to the change in individual utility function. Thus, according to the theory in [18], the proposed EENCG game is an exact potential game which has at least one pure strategy NE point, and the potential function is $\Gamma(a_m, a_{-m}) = \phi_{net}$.

Furthermore, according to [23], any global or local maxima of the potential function should be a pure strategy NE in exact potential game. Note that, the proposed potential function equals to the network utility, then the global maxima of network utility, i.e., the globally optimal solution to the proposed problem (5) is a pure strategy NE of the proposed EENCG game. Hence, the Theorem is proved.

4. Decentralized Energy Efficient Channel Selection Learning Algorithm

According to above theoretic analysis, a method which obtains the best NE of the game is also the method to achieve the global optimal solution of our problem. In this section, we design a distributed energy efficient channel selection learning (DEECSL) algorithm to obtain the global optimization in a distributed manner.

For potential game, several learning algorithms could achieve the NE points, such as best response dynamic [28], no-regret learning [24]. However, these learning algorithms may converge to some suboptimal NE points, not the optimal NE point.

In order to escape from suboptimal NE points and finally converge to the optimal NE, the proposed DEECSL algorithm introduces a probabilistic decision mechanism inspired by Boltzmann exploration strategy [29]. The DEECSL algorithm is shown in Table 1.

Table 1. The proposed DEECSL Algorithm

Initialization: Set the iteration $t = 0$, each Ob randomly chooses an action a_m from its available action set A_m ;

Loop:

Step 1: All the Obs access its chosen channel, then calculate their reward (4) in this iteration.

Step 2: All the Obs exchange the channel choosing strategies and rewards with neighbors through the common control channel (we assume that there is a common control channel available and an 802.11 DCF like access mechanism can be applied).

Step 3: Randomly generate a strategy change player, e.g., Ob_m , which could change its channel selection strategy in the next iteration. Other players hold their strategies in the next iteration. (the strategy change player could be selected randomly by the random counting timer method like, where the player whose timer counts to zero first would be the selected player.)

Step 4: The generated strategy change player, e.g., Ob_m , exploring all its available strategy space based on all the neighbors' information, calculate the expected utilities (7) of candidate strategies, and randomly choose an action a_m^{\wedge} according to the mixed strategy where the probability is given by:

$$\Pr(a_m^{\wedge}) = \frac{\exp\{\beta u_m(a_m^{\wedge}, a_{-m})\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \quad (13)$$

where β is the learning parameter, $\beta > 0$.

Other players, e.g., Ob_i repeats their strategies as last iteration: $a_i(t+1) = a_i(t)$.

End loop

In step 4, the learning parameter β shows the tradeoff between exploration and exploitation. Smaller β implies that players tend to choose a suboptimal action, whereas higher β implies that players are prone to choose the best response action. Note that the probabilistic decision in step 4 is designed to avoid the learning algorithm converges to some suboptimal NE points.

Theorem 2: The proposed DEECSL algorithm will converges to a unique stationary distribution of players' strategy profile, and the global optimum would be achieved with an arbitrarily high probability with a sufficiently large β . The unique stationary distribution

of players' strategy profile $\bar{\mathbf{a}} = \{\bar{a}_1, \bar{a}_2, \dots, \bar{a}_M\}$ will be given by:

$$\pi(\bar{\mathbf{a}}) = \frac{\exp\{\beta\Gamma(\bar{\mathbf{a}})\}}{\sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\}} \quad (14)$$

where $\Gamma(\mathbf{a})$ is the potential function. $A = A_1 \otimes A_2 \otimes \dots \otimes A_M$ is the set of strategy profiles of all the players, where \otimes is the Cartesian product.

Proof: the following proof is based on the theory in [29] and Theorem 6.2 in [30], which is also used in [18], [19], [23]. Denote the network state at the t -th iteration by $\mathbf{S}(t) = (S_1(t), S_2(t), \dots, S_m(t), \dots, S_M(t))$, where $S_m(t)$ means Ob_m 's working channel in the t -th iteration. According to the proposed algorithm, $\mathbf{S}(t+1)$ is only determined by $\mathbf{S}(t)$, which means that $\mathbf{S}(t)$ is a discrete time Markov process with a unique stationary distribution [29].

Denote $\mathbf{S}(t+1) = \mathbf{a}_2$, $\mathbf{S}(t) = \mathbf{a}_1$, and denote the transition probability from state \mathbf{a}_1 to \mathbf{a}_2 as $P_{\mathbf{a}_1, \mathbf{a}_2}$, the transition probability from state \mathbf{a}_2 to \mathbf{a}_1 as $P_{\mathbf{a}_2, \mathbf{a}_1}$. There is one or zero element that may be changed between \mathbf{a}_2 and \mathbf{a}_1 in the proposed algorithm (step 3).

When $\mathbf{a}_2 = \mathbf{a}_1$, we have:

$$\pi(\mathbf{a}_1)P_{\mathbf{a}_1, \mathbf{a}_2} = \pi(\mathbf{a}_2) \quad (15)$$

When $\mathbf{a}_2 \neq \mathbf{a}_1$, one player changes its working channel, which reduces that one element of the network state has been changed, e.g., Ob_m 's working channel from S_m to S'_m by adopting action a'_m . In step 3, the probability for Ob_m be has being chosen to update its action is $1/M$. Then we have:

$$\begin{aligned}
\pi(\mathbf{a}_1)P_{\mathbf{a}_1, \mathbf{a}_2} &= \left[\frac{\exp\{\beta\Gamma(\mathbf{a}_1)\}}{\sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\}} \right] \times \left[\left(\frac{1}{M} \right) \left(\frac{\exp\{\beta u'_m(a'_m, a_{-m})\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \right) \right] \\
&= \frac{\exp\{\beta\Gamma(\mathbf{a}_1)\} \times \exp\{\beta u'_m(a'_m, a_{-m})\}}{M \times \sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\} \times \sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \\
&= \frac{\exp\{\beta(\Gamma(\mathbf{a}_1) + u'_m(a'_m, a_{-m}))\}}{M \times \sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\} \times \sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}}
\end{aligned} \tag{16}$$

Similarly, suppose Ob_m 's working channel from S_m to S'_m by adopting action a'_m , we have:

$$\begin{aligned}
\pi(\mathbf{a}_2)P_{\mathbf{a}_2, \mathbf{a}_1} &= \left[\frac{\exp\{\beta\Gamma(\mathbf{a}_2)\}}{\sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\}} \right] \times \left[\left(\frac{1}{M} \right) \left(\frac{\exp\{\beta u_m(a_m, a_{-m})\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \right) \right] \\
&= \frac{\exp\{\beta\Gamma(\mathbf{a}_2)\} \times \exp\{\beta u_m(a_m, a_{-m})\}}{M \times \sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\} \times \sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \\
&= \frac{\exp\{\beta(\Gamma(\mathbf{a}_2) + u_m(a_m, a_{-m}))\}}{M \times \sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\} \times \sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}}
\end{aligned} \tag{17}$$

Because the game has been proved be an exactly potential game, we have:

$$\Gamma(\mathbf{a}_1) - \Gamma(\mathbf{a}_2) = u_m(a_m, a_{-m}) - u'_m(a'_m, a_{-m}) \tag{18}$$

Based on (18), we have:

$$\exp\{\beta(\Gamma(\mathbf{a}_1) + u'_m(a'_m, a_{-m}))\} = \exp\{\beta(\Gamma(\mathbf{a}_2) + u_m(a_m, a_{-m}))\} \tag{19}$$

Thus,

$$\pi(\mathbf{a}_1)P_{\mathbf{a}_1, \mathbf{a}_2} = \pi(\mathbf{a}_2)P_{\mathbf{a}_2, \mathbf{a}_1} \tag{20}$$

Considering all the states of the network, we have:

$$\sum_{\mathbf{a}_1 \in A} \pi(\mathbf{a}_1)P_{\mathbf{a}_1, \mathbf{a}_2} = \sum_{\mathbf{a}_1 \in A} \pi(\mathbf{a}_2)P_{\mathbf{a}_2, \mathbf{a}_1} = \pi(\mathbf{a}_2) \sum_{\mathbf{a}_1 \in A} P_{\mathbf{a}_2, \mathbf{a}_1} = \pi(\mathbf{a}_2) \tag{21}$$

According to the discrete time Markov process theory [29], the proposed algorithm has the unique stationary distribution (14) which satisfies the balanced equation of Markov process.

Furthermore, according to **Theorem 1**, the global optimal solution to network utility is exactly the best pure strategy NE of the game. Suppose that \mathbf{a}^{opt} is the globally optimal action profile of players, according to the design of the potential function, we have

$$\mathbf{a}^{opt} = \arg \max_{\mathbf{a} \in A} \phi_{net} = \arg \max_{\mathbf{a} \in A} \Gamma(\mathbf{a}) \quad (22)$$

According to **Theorem 1**, the algorithm converges to a unique stationary distribution $\pi(\bar{\mathbf{a}}) = \exp\{\beta\Gamma(\bar{\mathbf{a}})\} / \sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\}$. When the $\beta \rightarrow \infty$, $\exp\{\beta\Gamma(\mathbf{a}^{opt})\} \gg \exp\{\beta\Gamma(\mathbf{a})\}, \forall \mathbf{a} \in \{A \setminus \mathbf{a}^{opt}\}$.

Then the probability of globally optimal solution \mathbf{a}^{opt} will be

$$\lim_{\beta \rightarrow \infty} \pi(\mathbf{a}^{opt}) = \frac{\exp\{\beta\Gamma(\mathbf{a}^{opt})\}}{\sum_{\mathbf{a} \in A} \exp\{\beta\Gamma(\mathbf{a})\}} = 1 \quad (23)$$

This result means that the proposed learning algorithm converges to the global optimum with an arbitrarily high probability. Thus, the proof is completed.

Notably, the information exchange in the proposed algorithm is strictly constrained between neighboring players, which makes the cost controlled. The global information is not needed for each player. In other words, the proposed algorithm achieves global optimum with only neighboring information exchange.

Furthermore, we analysis the converging speed and the computing complexity of the proposed algorithm, which are two important factors in the practical application.

Theorem 3: The proposed DEECSL algorithm will asymptotically converge to the global optimum with geometric speed, and the computing complexity would be $O(N)$.

Proof: the following proof is based on the theory and approach in [32], also used in [33].

According to the above **Theorem 2**, the network state $\mathbf{S}(t)$ is a discrete time Markov process with a unique stationary distribution, here denoted by \mathbf{S}_π . Denote all the possible states of this Markov process as set Ξ_s .

Denote the initial state as \mathbf{S}_I in the Initialization step of the proposed algorithm. Then the network state would transmit from \mathbf{S}_I to \mathbf{S}_π according to process of learning and updating in the algorithm. Denote the transition matrix of the Markov process as \square , which depict the transition probabilities between the possible network states.

According to [32], define the distance between two states $\mathbf{S}_1, \mathbf{S}_2$ as $d(\mathbf{S}_1, \mathbf{S}_2)$, which measures the variation between \mathbf{S}_1 and \mathbf{S}_2 . Then the distance between \mathbf{S}_I and \mathbf{S}_π would be $d(\mathbf{S}_I, \mathbf{S}_\pi)$. Construct a matrix $\aleph = (\mathbf{S}_I, \mathbf{S}_\pi)^T$, and define $\aleph = (\mathbf{S}_I^T \square, \mathbf{S}_\pi^T \square)^T$.

The ergodic coefficient [32] is defined as follow:

$$\delta(\mathfrak{S}) = \frac{1}{2} \sup_{i,j \in \Xi_S} \sum_{k \in \Xi_S} |\mathfrak{S}_{ik} - \mathfrak{S}_{jk}| = \frac{1}{2} \sup_{i,j \in \Xi_S} d(\mathfrak{S}_i - \mathfrak{S}_j) \quad (24)$$

$$\delta(\mathfrak{T}) = \frac{1}{2} \sup_{i,j \in \Xi_S} \sum_{k \in \Xi_S} |\mathfrak{T}_{ik} - \mathfrak{T}_{jk}| = \frac{1}{2} \sup_{i,j \in \Xi_S} d(\mathfrak{T}_i - \mathfrak{T}_j) \quad (25)$$

$$\delta(\square) = \frac{1}{2} \sup_{i,j \in \Xi_S} d(\square_i - \square_j) \quad (26)$$

According to the property of ergodic coefficient [32], we have:

$$d(\mathbf{S}_I \square - \mathbf{S}_\pi) = \delta(\mathfrak{T}) \leq \delta(\mathfrak{S}) \delta(\square) = \frac{1}{2} d(\mathbf{S}_I - \mathbf{S}_\pi) \delta(\square) \quad (27)$$

Each updating step in the learning algorithm would cause a state transition according to the transition matrix. Suppose the network state is $S_I \square^k$ after the k -th state transition after the k -th learning step. Then the distance between the k -th state to the stationary state would be:

$$d(S_I \square^k - S_\pi) \leq \delta(\mathfrak{S}) \delta(\square)^k = \frac{1}{2} d(S_I - S_\pi) \delta(\square)^k \quad (28)$$

Furthermore, with the $\delta(\square)^k$, according to the definition of ergodic coefficient [27], we have:

$$\delta(\square) \leq 1 - \inf_{i,j \in \Xi_S} |\square_{ij}| \quad (29)$$

where \square_{ij} is the probability from state \mathbf{S}_i to \mathbf{S}_j .

With the state transition in the proposed algorithm, according to the step 4, the network state transition happens when the selected updating user changes its strategy based on equation (13). Suppose the updating player is Ob_m , its current strategy is a_m and the corresponding current network state is \mathbf{S}_i . Denote the possible next strategy is \hat{a}_m and the corresponding current network state is \mathbf{S}_j . According to the mixed strategy (13), the network state transition probability from state \mathbf{S}_i to \mathbf{S}_j would be:

$$\begin{aligned} \square_{ij} &= \Pr(\hat{a}_m) = \frac{\exp\{\beta u_m(\hat{a}_m, a_{-m})\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \\ &= \frac{\exp\{\beta u_m(\mathbf{S}_j)\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(\mathbf{S}_{a_m})\}} \geq \frac{\exp\{\beta u_m(\mathbf{S}_j)\}}{|A_m| \exp\{\beta u_m(\mathbf{S}_{\max})\}} \geq \frac{\exp\{\beta u_m(\mathbf{S}_j)\}}{|\Xi_S| \exp\{\beta u_m(\mathbf{S}_{\max})\}} \\ &= \frac{1}{|\Xi_S|} \exp\{\beta (u_m(\mathbf{S}_j) - u_m(\mathbf{S}_{\max}))\} \end{aligned} \quad (30)$$

where $|A_m|$ is the number of possible strategies for Ob_m .

Define $\bar{Y} = \max_{\substack{m \in S_{Ob} \\ \mathbf{S}_j, \mathbf{S}_{\max} \in \Xi_S}} (u_m(\mathbf{S}_{\max}) - u_m(\mathbf{S}_j))$, which is biggest gap of utility between all

possible two states in the Markov process Ξ_S . According to the utility function, \bar{Y} is a finited constant value and $\bar{Y} \geq 0$. Then we have:

$$p_{ij} \geq \frac{1}{|\Xi_S|} \exp\{\beta(u_m(\mathbf{S}_j) - u_m(\mathbf{S}_{\max}))\} \geq \frac{1}{|\Xi_S|} \exp\{-\bar{Y}\beta\} \quad (31)$$

As a result, we have:

$$\delta(p) \leq 1 - |\Xi_S| \inf_{i,j \in \Xi_S} (p_{ij}) \leq 1 - \exp\{-\bar{Y}\beta\} \quad (32)$$

Further more:

$$d(S_I p^k - S_\pi) \leq \frac{1}{2} d(S_I - S_\pi) \delta(p)^k \leq \frac{1}{2} d(S_I - S_\pi) (1 - \exp\{-\bar{Y}\beta\})^k \quad (33)$$

This means that the algorithm converges from the initial state to the network stationary state geometric speed.

With the computing complexity, in the proposed algorithm, the computing is mainly on the strategy selection, i.e., for the selected strategy updating player, Ob_m , the probabilities for possible strategies need to be computed by:

$$\Pr(a_m^\wedge) = \frac{\exp\{\beta u_m(a_m^\wedge, a_{-m})\}}{\sum_{a_m \in A_m} \exp\{\beta u_m(a_m, a_{-m})\}} \quad (34)$$

There are $|A_m| \leq |S_{CH}| = N$ possible strategies for Ob_m to be computed. Then the computing complexity would be $O(N)$. Thus, the proof is completed.

It should be noted that geometric converging speed should be acceptable for most applications, and the computing complexity is much smaller than the central control based exhaustive searching approach, whose computing complexity would be $O(N^M)$.

5. Numeric Results and Discussion

In this section, we evaluate the performance of the proposed algorithm by Matlab simulations. Without loss of generality, the following parameters are set: the number of channels is 5. Similar with [31], we set the path loss exponent as $\gamma = 2$, and the noise power as $N_0 = -130$ dB. The transmission distance for Obs is the unit distance for simplicity. The transmission data rate is 1 Mbit/s. The instantaneous random components are assumed to be unit-constant and the channels are assumed to undergo Rayleigh fading with unit mean.

The neighboring relationship of Ob s is randomly generated. The bandwidth of channels are set as [1, 1.5, 2, 2.5, 3] MHz to simulate the different channels. The results are obtained by simulating 3000 independent experiments and then taking the average value.

Table 2. Simulation Parameters

| Parameters | Value |
|---------------------------------------|--------------------------------|
| The number of channels | 5 |
| The path loss exponent | 2 |
| The noise power | -130 dB |
| The transmission data rate | 1 Mbit/s |
| The instantaneous random component | unit-constant |
| Channel fading | Rayleigh fading with unit mean |
| The neighboring relationship | randomly |
| bandwidth of channels | [1, 1.5, 2, 2.5, 3] MHz |
| The number of independent experiments | 3000 |
| The number of players | 15, 50 |

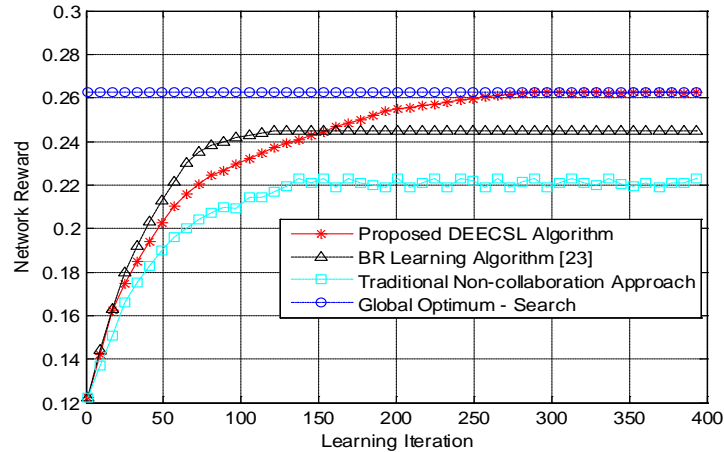


Fig. 2. The convergence and optimality of the proposed DEECSL algorithm, $M=15$.

The simulation is to show the performance of the proposed DEECSL algorithm compared with other approach, and verify the theoretic analysis of proposed algorithm, i.e., the convergence and the optimality. We first verify the convergence and the optimality of the proposed algorithm in a relative small scenario. Then we show the performance comparison furthermore in a relative large scenario to verify the adaption of the proposed algorithm.

To show the convergence and the optimality of the proposed algorithm, we compare the performance of the proposed algorithm with the globally optimal solution obtained by exhaustive search. We simulate 15 players scenario to avoid the limitation of computing capacity with the exhaustive search method. Fig. 2 shows that the proposed DEECSL algorithm converges to the globally optimal network utility after some learning iterations.

This result verifies the theoretic analysis on the proposed approach.

To show the advantage of the collaboration design in the proposed DEECSL algorithm, we also show the performance of the traditional non-collaboration approach, which is widely adopted in game theoretic learning algorithms. In the non-collaboration approach, each player does not care other neighboring players' benefits and make the decision totally based on its own reward. The result shows that the proposed collaboration design DEECSL algorithm outperforms the non-collaboration approach obviously. The main reason is that players in the non-collaboration approach would not consider the global reward, that also results in the unstability of non-collaboration approach. This comparison result verifies necessity and the advantage of the collaboration based utility function design in the proposed game model. In other words, with out the utility function design, the system would not converge to stable state and the performance would be worse.

In order to show the advangate of the design of probabilistic decision in the proposed algorithm, we compare with the best response (BR) algorithm [28]. Fig. 2 also shows that the proposed DEECSL algorithm obtains better performance compared with the BR algorithm. Here the BR algorithm is also based on the collaboration approach. The BR algorithm is a typical learning algorithm for games to achieve the NE points. The differenc between the BR algorithm and the proposed DEECSL algorithm is that players in BR algorithm choose the best strategy from all the candidate strategies rather than the probabilistic decision according to equation (13). Due to the "choosing the best" strategy, the BR algorithm's performance increases faster in the beginning and converges earlier. However, because players choose the current best strategies absolutely, they might loose some potential strategies in the long term. As a result, the BR learning algorithm may converges to some local optimum, which reduces the converged performance.

Compared with the BR algorithm, the proposed algorithm could escape from local optimum to achieve globally optimum due to the design of probabilistic decision. In other words, the probabilistic decision based on equation (13) rather than the "absolutely choosing the current best" would bring the chance for the player to make the right decision in the long term.

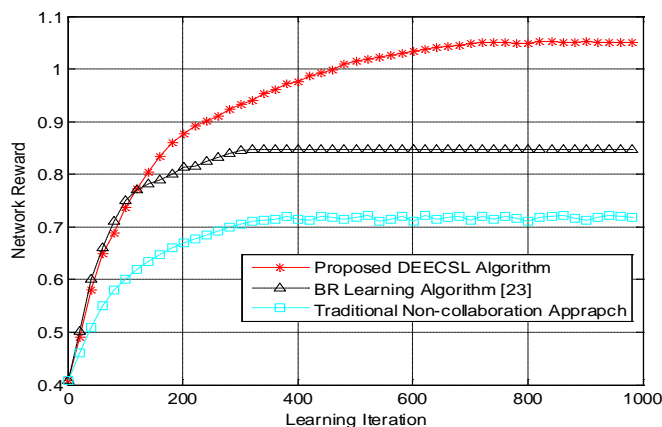


Fig. 3. The proposed DEECSL algorithm vs the BR learning algorithm, $M=50$.

To verify the proposed algorithm furthermore, we simulate a larger network scale where the number of players is 50. Due to the limitation of computation capability, the globally optimal solution by the exhaustive search is not given out. We compare the proposed algorithm with the BR learning algorithm and the non-collaboration approach. Fig. 3 shows the convergence of the proposed DEECSL algorithm again. Also, the proposed DEECSL algorithm achieves better performance than the BR algorithm. The non-collaboration approach achieves worst performance and instability for the lack of collaboration. This result verifies the theoretic analysis of the proposed algorithm and the performance advantage compared with other approaches again. That means the proposed approach is widely applicative in different scenarios.

6. Conclusion

In this paper, we studied the distributed energy efficient resource management in the Internet of Things. We proposed an energy efficient neighbor collaborative game model and proved that the proposed game was an exact potential game. Then, to achieve the global optimization to the proposed energy efficient channel selection problem in a distributed manner, we designed a distributed energy efficient channel selection learning algorithm to obtain the global optimization with only neighboring information exchange. We proved that the proposed DEECSL algorithm will asymptotically converge to the global optimum with geometric speed. Simulation results verified the theoretic analysis and the performance of proposed algorithm.

References

- [1] Q. Wu, G. Ding, Y. Xu, S. Feng, Z. Du, J. Wang, "Cognitive Internet of Things: A New Paradigm Beyond Connection," *IEEE Internet of Things Journal*, vol. 1, no. 2, Apr. 2014. [Article \(CrossRef Link\)](#)
- [2] S. Agrawal and M. L. Das, "Internet of Things – A Paradigm Shift of Future Internet Applications," *Institute of Technology, Nirma University, AHMEDABAD*, Dec. 2011. [Article \(CrossRef Link\)](#)
- [3] F. Lin, Q. Liu, X. Zhou, Y. Chen, D. Huang, "Cooperative Differential Game for Model Energy-Bandwidth Efficiency Tradeoff in the Internet of Things," *Network Technology and Application, China Communications*, Jan. 2014. [Article \(CrossRef Link\)](#)
- [4] X. Zheng, P. Martin, K. Brohman, and L. Xu, "Cloud Service Negotiation in Internet of Things Environment: A Mixed Approach," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, May. 2014. [Article \(CrossRef Link\)](#)
- [5] T. Serbulent, S. Murat, M. Wei, "Wi-Fi Enabled Sensors for Internet of Things: A Practical Approach," *IEEE Communications Magazine*, vol. 50, no. 6, pp. 134-143, 2012. [Article \(CrossRef Link\)](#)
- [6] L. T. Sanchez, D. C. Ranasinghe, M. Harrison, "Adding Sense to the Internet of Things," *Personal and Ubiquitous Computing*, vol. 16, no. 3, pp. 291-308, 2012. [Article \(CrossRef Link\)](#)

- [7] R. Roman, C. Alcaraz, J. Lopez, "Key Management Systems for Sensor Networks in the Context of the Internet of Things," *Computers and Electrical Engineering*, vol. 37, no. 2, pp. 147-159, 2011. [Article \(CrossRef Link\)](#)
- [8] R. Klauck, M. Kirsche, "A Case Study of a DNS-Based Discovery Service for the Internet of Things," *ADHOC-NOW*, pp. 316-329, 2012. [Article \(CrossRef Link\)](#)
- [9] Z. Chen, F. Xia, T. Huang, et al, "A Localization Method for the Internet of Things," *Journal of Supercomputing*, pp. 1-18, 2011. [Article \(CrossRef Link\)](#)
- [10] T. Y. Wu, G. H. Li, S. W. Huang, et al, "A GABased Mobile RFID Localization Scheme for Internet of Things," *Personal and Ubiquitous Computing*, vol. 16, no. 3, pp. 245-258, 2012. [Article \(CrossRef Link\)](#)
- [11] Z. Chen, H. Wang, Y. Liu, et al, "A Context-Aware Routing Protocol on Internet of Things Based on Sea Computing Model," *Journal of Computers*, vol. 7, no. 1, pp. 96-105, 2012. [Article \(CrossRef Link\)](#)
- [12] M. Danieleto, N. Bui and M. Zorzi, "Improving Internet of Things Communications Through Compression and Classification," in *Proc. of the 2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, 2012. [Article \(CrossRef Link\)](#)
- [13] S. Li and C. Fan, "Joint Optimization on Energy and Delay for Target Tracking in Internet of Things," *China Communications*, vol. 8, no. 1, pp. 20-27, 2011. [Article \(CrossRef Link\)](#)
- [14] J. Duan, D. Gao, D. Yang, C. H. Foh, and H. Chen, "An Energy-Aware Trust Derivation Scheme With Game Theoretic Approach in Wireless Sensor Networks for IoT Applications," *IEEE Internet of Things Journal*, vol. 1 no. 1, Feb. 2014. [Article \(CrossRef Link\)](#)
- [15] D. Fudenberg and D. K. Levine, "The Theory of Learning in Games," *The MIT Press*, 1998. [Article \(CrossRef Link\)](#)
- [16] J. Suris, L. Dasilva, Z. Han, A. Mackenzie, and R.Komali, "Asymptotic optimality for distributed spectrum sharing using bargaining solutions," *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 5225-5237, Oct. 2009. [Article \(CrossRef Link\)](#)
- [17] Z. Zhang et al., "A cooperation strategy based on Nash bargaining solution in cooperative relay networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2570-2577, Jul. 2008. [Article \(CrossRef Link\)](#)
- [18] J. Zheng et al., "Optimal power allocation and user scheduling in multicell networks: Base station cooperation using a game-theoretic approach," *IEEE Trans. Wireless Commun*, vol. 13, no. 12, pp. 6928–6942, Dec. 2014. [Article \(CrossRef Link\)](#)
- [19] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Signal Process*, vol. 6, no. 2, pp.180–194, Apr. 2012. [Article \(CrossRef Link\)](#)
- [20] H. Dai, Y. Huang, and L. Yang, "Game theoretic max-logit learning approaches for Joint Base Station Selection and Resource Allocation in Heterogeneous Networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, Jun. 2015. [Article \(CrossRef Link\)](#)
- [21] A. Montanari and A. Saberi, "Convergence to equilibrium in local interaction games," in *Proc. 50th Annu. IEEE Symp. Foundat. Comput*, pp. 303–312, 2009. [Article \(CrossRef Link\)](#)
- [22] G. Stuber, *Principles of Mobile Communications*, 2nd ed. Norwell, MA, USA: Kluwer, 2001. [Article \(CrossRef Link\)](#)
- [23] J. Marden, G. Arslan, and J. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst., Man, Cybern. B*, vol. 39, no. 6, pp. 1393–1407, Dec. 2009. [Article \(CrossRef Link\)](#)
- [24] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized dynamic spectrum access for cognitive radios: cooperative design of a noncooperative game," *IEEE Trans. Commun*, vol. 57, no. 2, pp. 459–469, 2009. [Article \(CrossRef Link\)](#)

- [25] M. A. Nowak, “Five rules for the evolution of cooperation,” *Science*, vol. 314, pp. 1560–1563, 2006. [Article \(CrossRef Link\)](#)
- [26] F. C. Santos, M. D. Santos, and J. M. Pacheco, “Social diversity promotes the emergence of cooperation in public goods games,” *Nature*, vol. 454, pp. 213–249, 2008. [Article \(CrossRef Link\)](#)
- [27] E. Altman, T. Jimenez, N. Vicuna, and R. Marquez, “Coordination games over collision channels,” in *Proc. WiOPT*, pp. 523–527, 2008. [Article \(CrossRef Link\)](#)
- [28] W. Zhong, Y. Xu, and H. Tianfield, “Game-theoretic opportunistic spectrum sharing strategy selection for cognitive MIMO multiple access channels,” *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2745–2759, Jun. 2011. [Article \(CrossRef Link\)](#)
- [29] P. J. M. van Laarhoven and E. H. L. Aarts, *Simulated Annealing: Theory and Applications*. Holland: Reidel, 1987. [Article \(CrossRef Link\)](#)
- [30] H. P. Young, *Individual Strategy and Social Structure*, Princeton, NJ: Princeton Univ. Press, 1998. [Article \(CrossRef Link\)](#)
- [31] B. Babadi and V. Tarokh, “GADIA: A greedy asynchronous distributed interference avoidance algorithm,” *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6228–6252, Dec. 2010. [Article \(CrossRef Link\)](#)
- [32] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, Queues*. New York, NY, USA: Springer-Verlag, 1999. [Article \(CrossRef Link\)](#)
- [33] Y. Li, D. Jin, J. Yuan, and Z. Han, “Coalitional Games for Resource Allocation in the Device-to-Device Uplink Underlying Cellular Networks,” *IEEE Transactions on wireless communications*, vol. 13, no. 7, Jul. 2014. [Article \(CrossRef Link\)](#)



Chunhua Ju graduated from Xiamen University with PhD degree in 2002. Now he is a professor, doctoral supervisor and division chief of science and technology department in Zhejiang Gongshang University. He focuses on intelligent information processing, data mining and collaborative innovation. He won the award for “New Century Excellent Talents in University” of China.



Qi Shao received the B.E. degree in Computer Science and Technology from Zhejiang Gongshang University, China, in 2014. He is currently pursuing the M.A. degree in Computer Science and Technology in Gongshang University, China.