Balancing Energy Consumption and Reputation Gain of UAV Scheduling in Edge Computing

Juan Zhang, Yulei Wu, *Senior Member, IEEE*, Geyong Min, *Member, IEEE*, Fei Hao, and Laizhong Cui, *Senior Member, IEEE*

***Abstract*—Due to the extensive use of unmanned aerial vehicles (UAVs) in civil and military environment, effective deployment and scheduling of a swarm of UAVs are rising to be a challenging issue in edge computing. This is especially apparent in the area of Internet of Things (IoT) where massive UAVs are connected for communications. One of the characteristics of IoT is that an operator can interact with more than one UAVs for the effective scheduling under multi-task requests. Based on this scenario, we clarify the issue on how to maintain the energy efficiency of UAVs and guarantee the reputation gain during the scheduling deployment. In this paper, we first formulate the energy consumption and reputation into the decision model of UAVs scheduling. A game-theoretic scheme is then developed for the optimal decision searching. With the developed model, a range of important parameters of UAV scheduling are thoroughly investigated. Our numerical results show that the proposed scheduling strategy is able to increase the reputation and decrease the energy consumption of UAVs simultaneously. In addition, in the game process, the profit of an operator can be maximized and the network economy research can be explored.**

***Index Terms*—Energy efficiency, reputation, UAV scheduling, game theory, performance analysis.**

I. Introduction

UNMANNED aerial vehicles (UAVs), originally used in military tasks, including reconnaissance and monitoring from the sky, have been extended to civilian domains such as remote sensing of mapping, bridge inspection, and precision agriculture monitoring [1], [2]. With the support of 5G, UAV-related technologies play a critical role in the field of Internet of Things (IoT) for its important mobile edge services, such as aerial video surveillance, providing unprecedented aerial perspective for ground monitoring [3]. Specifically, UAVs can work at the places where people may not be able to reach and can be scheduled effectively to execute multiple tasks.

The flight time of UAVs has not kept pace with technology advancement due to the limitations of traditional battery technologies. Radical improvement in energy efficiency is always the essential requirement of UAVs scheduling. Existing studies on the restriction problem of battery capacity have been developed and performed. For example, Ghazzai *et al.* [4] proposed a generic scheduling framework to manage a fleet of micro UAVs, in which technical specifications of UAVs and the limitation of battery capacities are considered. With the aim of operators in commerce, the profits of operators can be improved by minimizing the energy cost.

In addition to the effect of energy consumption, the reputation of UAVs can be further affected for UAVs' performance in a mission by measuring their credibility. For example, low energy supply can decrease the reputation of a working UAV and reduce the probability of being scheduled. A series of reputation schemes have been designed for UAVs scheduling. Trestian *et al.* [5] proposed a framework which considers the reputation into a network selection problem to model a reputation-based system.

Most of existing research on UAV scheduling have energy consumption and reputation gain modeled separately [6], [7], [8], [9], [10], [11], [12]. Even though a few models integrate energy efficiency and reputation improvement into the modeling of UAV scheduling [13], [14], there has been no study constructing UAV scheduling models from the perspective of commercial profits. Different from existing research, this paper addresses the scheduling issue of UAVs from the integration of energy consumption and reputation gains of UAVs that affect the practical economic benefits.

In this study, we propose a scheduling strategy by considering characteristics of UAVs in terms of energy consumption and reputation in threefold. Firstly, we propose an energy efficient strategy and a reputation-based mechanism separately to construct the scheduling strategy rules of UAVs. Secondly, we build a game-theoretic model to study the scheduling problem of working UAVs when they receive new task requests. Thirdly, we derive the balance of energy minimization and reputation maximization of UAVs to achieve the Nash equilibrium efficiently for UAVs scheduling. The main contributions of this study can be summarized as follows:

- The energy consumption model and the reputation gain model are formulated separately with respect to game theoretic features. The energy consumption model is

Manuscript received February 15, 2020; revised May 13, 2020; accepted June 16, 2020. Date of publication June 24, 2020; date of current version December 9, 2020. This work was supported in part by the Engineering and Physical Sciences Research Council of United Kingdom under Grant No. EP/R030863/1 and the EU Horizon 2020 Programme Marie Sklodowska-Curie Individual Fellowship under Grant No. H2020-MSCA-IF-2018-840922. The associate editor coordinating the review of this article and approving it for publication was X. Chen. *(Corresponding authors: Yulei Wu; Geyong Min.)*

Juan Zhang, Yulei Wu, and Geyong Min are with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, U.K. (e-mail: y.l.wu@exeter.ac.uk; g.min@exeter.ac.uk).

Fei Hao is with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, U.K., and also with the School of Computer Science, Shaanxi Normal University, Xi'an 710119, China.

Laizhong Cui is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China.

Digital Object Identifier 10.1109/TCCN.2020.3004592

constructed according to the working state and future scheduling strategy of an UAV. The compensation mechanism of the reputation model is developed by considering the practical scenario in commercial competitions.

- The energy model and the reputation model are built into a game theoretic model to maximize the payoff of a network operator. The reached Nash equilibrium point by the developed game mechanism can help make the scheduling decisions of UAVs in multi-tasks requests from network operators.
- By virtue of the developed game model, we thoroughly investigate the effects of critical system parameters on the energy consumption, reputation gain and the payoff. The relationships among reputation, energy consumption and payoff are explored based on these parameters. In addition, fluctuations of the objectives in our proposed mechanism are further analyzed.

The rest of this paper is organized as follows. Section II introduces the related studies including state-of-the-art on energy efficiency and reputation of UAV scheduling. Section III provides preliminary work and important concepts that can facilitate the understanding of our proposed work. Section IV presents the system architecture, the energy consumption model, and the reputation gain model. Section V describes the game mechanism. The numerical results are provided in Section VI. Finally, Section VII concludes this paper.

II. Related Work

Sufficient energy supply is critical for the implementation of mobile edge tasks, but it generates cost to network operators. From the perspective of the lucrative purpose, sustainable task assignment to UAVs can promote the profit gain of network operators. However, it is a challenging problem of constructing the energy consumption and reputation models with the purpose of commercial profit maximization for UAV operators. Therefore, effective UAV scheduling strategies with the consideration of energy and reputation are urgent to be proposed. In what follows, we introduce the related work on this problem.

A. Energy Consumption

Features-oriented modelling can achieve UAV scheduling and maximize the profit of network operators effectively. In disaster areas, UAVs in the optimum deployment state can benefit search and rescue operations, which was studied in [7]. Li *et al.* [8] proposed the Dijkstra algorithm which is used to analyse the shortest path from the UAV site to the gathering site. The arrival time of a UAV can be computed according to the shortest path and permissible speed. Therefore, the UAVs that can satisfy the time constraint of rescue will be obtained, which can significantly improve the working efficiency when UAVs are called from other sites. The research reported in [9] described the scheduling framework of solving the problem of sequential search tasks. The authors used previous empirical data to formulate the problem of task allocation to network operators. The allocation strategy is worked as an optional non-preemptive scheduling system. In geometrically complex environments such as dense urban areas and mountainous terrains, quality-aware UAV coverage and path planing problems were researched in [10]. An occlusion-aware way-points generation algorithm was designed to find the best set of way-points for taking pictures in a target area to satisfy the spatial resolution requirement. These selected waypoints are assigned to multiple UAVs by solving a vehicle routing problem, so that all the way-points are visited within a global deadline to meet the temporal resolution requirement. An automated and decentralized surveillance system for the problem of detecting and tracking on UAV scheduling in a bounded area was proposed in [15]. In this work, a non-trivial robotic implementation of the distributed algorithm was provided for a set of instances and challenging problems. For the unique UAV characteristics of being replaced with grounded and fixed sensors, several researchers devised tailored algorithms based on the top-down methodology for the specific application scenarios [15], [16], [17]. The authors in [18] formulated an automated system to ease the operator's workload by breaking up the video streams into parts and scheduling a subset of them for the operator's inspection. In this work, problems are formulated as a time-indexed integer programming where time is discrete, and mixed-integer non-linear programming is combined with cutting framework to obtain the optimal solution.

Dynamic modelling and approaches in tasks allocation are key solutions to UAVs scheduling. In optimization, dynamic ant colony algorithm has been widely employed in areas including UAVs task allocation, power optimization, and modelling. Particularly, UAVs need long-term running and light weight power source for multi-tasks implementation. Therefore, tasks allocation mechanisms require high adversarial models to overcome the intrinsic complexity (such as the dynamic nature of UAVs) in different systems. Wu *et al.* [6] stated how to allocate dynamic tasks in UAV systems by constructing a dynamic ant colony's lobar division (DACLD) model which is based on the classic fixed response threshold. The DACLD model has the characteristics of a distributed framework which solve not only the dynamic problems in execution order, multi-state, adaptive response threshold and multi-individual response, but also the complex tasks by a swarm of agents. Rosalie *et al.* [19] proposed a mobility ant colony optimization (ACO)-based model for multilevel swarms of UAVs by generating unpredictable trajectories with a chaotic solution of dynamic systems. In the research, chaotic dynamics are explained as the typical solutions to deterministic systems with the specific properties, but these solutions are bounded, globally time invariant and sensitive to initial conditions, and consequently they are unpredictable in a long time. Integrating these solutions to the mobility model can obtain deterministic but unpredictable trajectories. Therefore, their system combines the Ordinary Differential Equations with the existing Ant Colony Algorithm, and considers the nonlinear chaotic system [20] to illustrate the transition from the random part of an ACO algorithm to a chaotic one.

Another emerging mobility model was designed for a fleet of UAVs based on a fuzzy logic inference system [21]. It develops an alpha-based mobility model which integrates energy level, coverage area and network connectivity attributes to make a mobility decision. This novel fuzzy inference system has been implemented to compute the values of a fellowship weighting parameter, named Alpha. Then the most suitable UAV can be selected by the system.

Energy modelling and optimization determine the level of tasks implementation in an UAV system. Phung and Morin [22] proposed a modeling of the energy consumption of a class of small convertible system called vertical Take-off and Landing Unmanned Aerial Vehicles (VTOL-UAVs). This method is based on a six-parameter-analytical model which lies on a set of coplanar propellers for propulsion and wings to improve the energy efficiency. In 5G based IoT and body sensor networks (BSN), the discovery of the energy over-consumption is a challenging problem in the removal, query and routing of network nodes. However, a high transmission capacity cannot be guaranteed in some approaches that demand for the energy conservation and fault tolerance. Therefore, the authors in [23] presented an architecture which utilizes XML charts to perform device discovery on the basis of networks state cost and available energy.

B. Reputation Gain

Reputation mechanisms in economics have drawn extensive attentions and been extended to engineering and industry for long-term development. A semi-distributed reputation mechanism based on a dynamic data-driven application system was proposed in [11], where the local reputation and global reputation are involved. Local reputation is dynamically and selectively injected to the central controller, where injected data are collected for global reputation computation. Subsequently, the central controller can detect malicious nodes in unknown networks by the information of dynamic change of trust and the balance of distributed nodes. Tian *et al.* [12] considered dynamic and diversity attacking strategies in the simulation of evaluating reputation management schemes. The evaluation method is proved to be able to depict a detailed evolution process. Tang *et al.* [24] discussed a new reputation-based mechanism for the proof of work (PoW) computation in the blockchain, in which miners are incentive to conduct honest mining. An algorithm is designed to encourage the honest mining of miners in this work. The evolution of cooperation in public goods games on complex networks is concentrated in [25], where individuals have various reputation tolerances. Reputation tolerance-based scheme is proved to prevent defectors' free-riding behaviour and enhances the formation of cooperative clusters. Reputation-based voting scheme was designed in [26] to ensure secure miner selection. This scheme evaluates candidates' reputation combining both past interactions and recommended opinions from other vehicles.

With the above metrics, this work aims to achieve the balance of energy consumption and reputation gain in order to perform the UAVs scheduling effectively.

III. Preliminaries

A. Dynamic Game of Incomplete Information

Game theory has been broadly applied in collisions of resource allocation and routing problems in communication networks. This tool provides an effective way to achieve an equilibrium point by a game approach so as to realize their maximum optimization to some extent. There are three critical elements in a game theory model: player, strategy and payoff. The *players* are the participants of a game event. The *strategy* is the rule that players need to follow. The *payoff*, also known as revenue, is the benefit or reward that a player/system gains in a game event under the strategy. Based on these factors, we consider the players' characteristics in our model as follows:

- Whether players embrace the complete information: complete information game or incomplete information game;
- Whether players are cooperative: cooperative game or non-cooperative game;
- Whether the game is dynamic or static over a number of time periods.

Our model is immersed in non-cooperative dynamic games with incomplete information according to the application scenario. To analyse dynamic games with incomplete information, we introduce the concept of Perfect Bayesian Equilibrium, which will be described as below.

In an incomplete information game, players are unsure about the payoffs or preferences of opponents. This game can be modeled as a *Bayesian Game* that consists of not only players, strategies and payoff, but the probability distribution $P(I_1, \ldots, I_N)$ of players over types, where $I_1, \ldots, I_N$ are players in the game. A Bayesian Nash equilibrium is a Nash equilibrium of the "expanded game" in which each pure strategy is the set of maps of player types Θ to the strategy space Ω.

Dynamic game of incomplete information implies that the strategy is sequentially rational, and no player can improve the payoffs individually at any stage of the game [27], [28]. The optimal method of a Bayesian game is to define a belief system, denoted by μ, which determines a posterior for each player over the set of scheduling in an information set. A belief system is consistent if it is derived from equilibrium strategies by *Bayes Rule*. Here, we have player *i* and the *i*'s strategy choice α^h, the posterior probability of *i* belonging to types θ^k can be presented with classical Bayes rules

$$\begin{aligned} Prob\left\{\theta^k|\alpha^h\right\} &\equiv \frac{p\left(\alpha^h|\theta^k\right)P\left(\theta^k\right)}{prob\left\{\alpha^h\right\}} \\ &\equiv \frac{p\left(\alpha^h|\theta^k\right)P\left(\theta^k\right)}{\sum_{j=1}^{K} p\left(\alpha^h|\theta^j\right)P\left(\theta^j\right)} \end{aligned}$$

For all $i \in I$ and all $\theta_i \in \Theta_i$, we have Bayesian Nash equilibrium as

$$s_i(\theta_i) \in arg\max_{s_i' \in S_i} \sum_{\theta_- i} P(\theta_- i|\theta_i)\mu_i\left(s_i', s_- i(\theta_- i), \theta_i, \theta_- i\right)$$

Perfect Bayesian Equilibrium in a dynamic game of incomplete information is about strategy profile s and belief

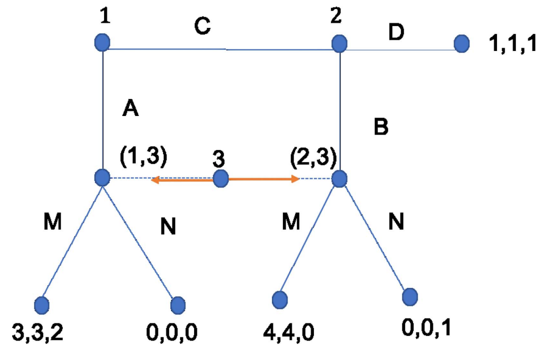


Fig. 1. An example of dynamic game of incomplete information.

system μ. The strategy profile s is sequentially rational given by μ, and the belief system μ is consistent given by s. In order to better demonstrate the knowledge, we give an example to represent the perfect Bayesian equilibrium in Fig. 1, where 1, 2, 3 are players, and the strategy profile $\{A, B, C, D, M, N\} \in \Omega_i$. We suppose that the strategy of player 1 is A with probability p and the strategy of player 2 is B with the probability q. Then, the Bayes rule implies the belief as

$$\mu_3(M) = \frac{p}{p + (1-p)q}$$

Based on the above strategy rule, the equilibrium we pursue is "Perfect Bayesian Equilibrium" – a function which relies on a strategy profile Ω_i and a belief system μ. It is noticeable that the strategy profile Ω_i is sequentially rational under the belief μ and the belief system μ is consistent if given the strategy profile Ω_i. Therefore, if player 1 chooses strategy A or C, player 2 and player 3 will have no idea about which exact strategy that player 1 has taken. The game goes into the information set of players 2 and 3, where strategy M is strictly preferred to N because it is not sequential and rational.

We suppose that p and $1-p$ are the predicted probabilities of A and C for players 2 and 3, respectively. Player 2 has the probability of q to choose B, so that the probability of choosing D is $1-q$, then the expectation payoff of player 3 choosing M can be expressed as $p \cdot 3 + (1-p) \cdot q \cdot 0 = 3p$. The probability of choosing N is $p \cdot 0 + (1-p) \cdot q \cdot 1 = q - pq$. For player 3, M is strictly preferred to N, so Nash Equilibrium can be concluded as (C, D, N) and (A, D, M).

IV. The System Architecture

The flexibility of UAVs is illustrated in different working scenarios such as agriculture monitoring, major competitions, and so on. The collected data can be either processed locally or delivered back to the control centre. The process firstly starts from customers' requests initiated in the data control centre, then the task signal is broadcast to UAVs by the control centre. Fig. 2 presents a system architecture of an UAVs scheduling network, where all UAVs can receive the broadcast of emergency requests from the control center. In this situation, one of the UAVs may decide to transfer for the execution of a new task when it receives the task request.

Let $i \in I = \{1, \ldots, N\}$, where I is the set of players and N is the total number of players. The strategy of player

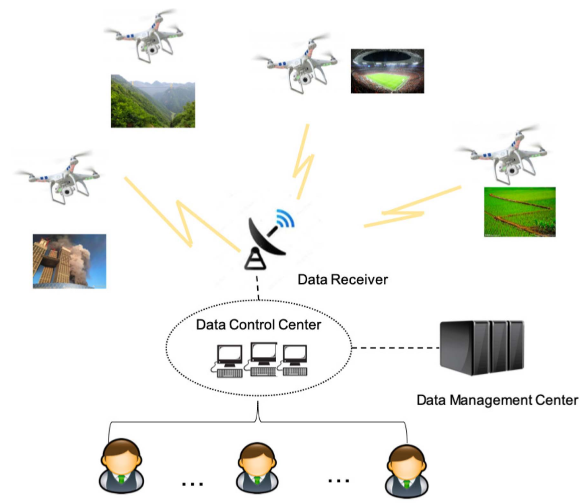


Fig. 2. System structure of UAV scheduling network.

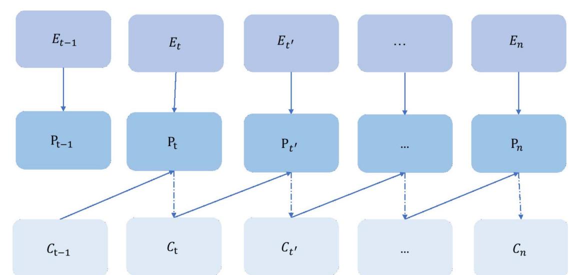


Fig. 3. An intuitive understanding of the proposed model.

i is denoted by $s_i \in \Omega$, where Ω is the strategy space. A strategy problem lies in a scenario where a player i works in the place $\overline{I}$ for customer A until time t during which it receives a new request of serving customer B in place $\overline{II}$. Because the bid of players are different, we suppose all payers are selfish and economically rational. These players have the purpose of 1) decreasing their energy consumption, while 2) guaranteeing the increase of their reputation, and 3) attempting to maximize the payoffs of each individual and the group. The parameters refereed in the proposed models are summarized in Table I.

Based on the above description, the strategy model, energy model and reputation mechanism are built in the following sections.

V. Strategy Model

A strategy reveals the action rule which is conducted by UAVs in a game. In our scheme, the strategy relies on predicted energy consumption and player's reputation. Our proposed model can be intuitively depicted in Fig. 3. As shown in this figure, at time t, the payoff of player i consists of two elements: consumed energy E_t and credits inherited from time $t-1$. For the new task request at time t', whether the player accepts the new task will be determined according to expected $P_{t'}$ based on demanded energy. The proposed model, named payoff-credit model, can be mathematically expressed in the following sections.

TABLE I
PARAMETER REPRESENTATIONS

Parameter	Explanation
j	The current task in site $\overline{I}$
k	A new task in $\overline{II}$
t	The working time of player i working in task j before flying away
t'	The time point of player i working in task k before next flying
$P_{i,j}$	Payoff of player i of task j
$P_{i,k}$	Payoff of player i of task k
Pi, j_t	Payoff of player i of task j at time t
$P_{i,j_t k}$	Payoff of player i flying for new task k after doing some time in task j
$c_{i,j}^t$	Cost of energy of i for task j at time t
$c_{i,k}$	Average cost of i doing task k
$M_{i,j}$	Task$j$ that player i need to working with
$M_{i,j}^t$	Task based on j that i finished at time t
$M_{i,j}^r$	Residual task based on j after t time
p_j	Average unit price of energy for i staying in the current task
p_k	Average unit price of energy for i leaving for a new task
E_i^o	Original energy of player i
$E_{i,j}$	The total energy consumption of finishing the task j in place $\overline{I}$
$E_{i,j}^r$	Residual energy of the player i after working some time on j
$E_{i,k}$	The total energy consumption of finishing the task k in place $\overline{II}$
br_i	The reputation compensation that the player i obtain from tasks
$cr_{i,j}$	The cost of the reputation compensation of player i working on task k
$c_{i,j}$	Cost of player i to do task $M_{i,j}$
r_i	The reputation of player i
r_i'	The reputation after the change of player i
$o_{i,j}$	The increment of i working in task j
$o_{i,j}'$	The fluctuating increment of i working in j
$f(r_{i,j})$	Reputation compensation
α, β	Weight parameters
ξ_1	Minimum coefficient energy required for task k
ξ_2	Maximum coefficient energy required for task k
μ, ψ	Adjust coefficient of reputation
W_i	Residual tasks waiting i to do
λ	Coefficient vector
ω_i	Weight of energy and reputation

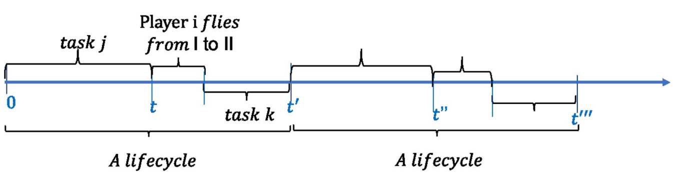


Fig. 4. The life-cycle of UAV scheduling.

More explicitly, all UAVs embrace the same options of stay or leave in a lifecycle $0 \longrightarrow t'$ (as depicted in Fig. 4). This study falls into a dilemma issue, where the decision on transferring or not will be studied when an UAV player i works on the task j until time t during which it receives the request of new task k. Similarly, in the next life-cycle, the player i has the same strategic circumstances.

Following the strategy model, the detailed energy model and reputation model are elaborated in the subsections to support the strategy.

A. Energy Model

Requests of tasks are broadcast to UAVs by a control centre or other relaying UAVs. These tasks are the drivers of energy consumption of UAVs. Our energy model can increase the overall payoff of a working UAV system by taking the energy consumption minimization as one of the important metrics. Under an emergency request, the player should consider energy consumption and decide whether to obey the instructions (stay or leave).

In our scenario, we assume that all energy capacities of players are fixed and equal. In place $\overline{I}$, the payoff of the player i working on task j until time t can be described as

$$P_{i,j,t} = M_{i,j} p_j - \left(E_i^o - E_{i,j}^r\right) c_{i,j,t} \tag{1}$$

where E_i^o and $E_{i,j}^r$ stand for the original energy and residual energy of a player i working on task j, respectively. $c_{i,j,t}$ is the cost of a unit energy. Therefore, $(E_i^o - E_{i,j}^r)c_{i,j,t}$ is the total cost of the player i working in the current place until time t. $M_{i,j}$ is the workload of a task j, and p_j represents the average price per workload unit charging from customers. Let us define $M_{i,j}p_j - (E_i^o - E_{i,j}^r)c_{i,j,t}$ as the payoff benefit gained from working on task j. This process is dynamic because of the change of payoff along with the growth of UAV's working time.

Once player i receives a new task request, the corresponding strategy in strategy space $\Omega = \{stay, leave\}$ will be taken. We assume that the game strategy is only determined by energy in this stage. In the following, we discuss the two strategies.

- *Player i stays in the current working place:* player i keeps working on the task j, and the residual workload is $M_{i,j}^r$. The total energy consumption of finishing the task in place $\overline{I}$ is expected as $E_{i,j}$. We have the payoff $P_{i,j}$ which is based on the energy cost during working time.

$$P_{i,j} = \left(M_{i,j,t} + M_{i,j}^r\right) p_j - E_{i,j} c_{i,j}$$
$$s.t. \quad E_{i,j} \leq E_i^o \tag{2}$$

where

$$M_{i,j,t} + M_{i,j}^r = M_{i,j}$$

Therefore, according to Equation (2), we have

$$P_{i,j} = M_{i,j} p_j - E_{i,j} c_{i,j}. \tag{3}$$

- *Player i leaves for a new task:* player i leaves the working place $\overline{I}$ to place $\overline{II}$. Let k represent the new task, $E_{j,k}^f$ denote the energy consumption of player i flying from place $\overline{I}$ to place $\overline{II}$, $E_{i,k}$ stand for the energy consumption of the new task k. Thus, $(E_{i,k} + E_{j,k}^f)c_{i,k}$ is the total cost for energy consumption. For player i, according energy consumption, the payoff becomes

$$P_{i,k} = M_{i,k} p_k - \left(E_{i,k} + E_{j,k}^f\right) c_{i,k} \tag{4}$$

where $M_{i,k}$ is the workload of the new task of player i, and p_k is the charging price of providing service. The expression of $M_{i,k}p_k - (E_{i,k}^n + E_{j,k}^f)c_{i,k}$ is the payoff considering the new task execution. Because the energy consumption is continuous, it is reasonable to integrate the payoff of a player working until time t in the place $\overline{I}$ into the payoff of the "*leave*" strategy. In this case, all payoff of "*leave*" will be greater than "*stay*", which is meaningless in some extent. Therefore, we concentrate on the payoff of energy cost if player i makes a "leave" strategy.

$$\begin{aligned} P_{i,j,t,k} &= P_{i,k} + P_{i,j,t} \\ &= M_{i,k} p_k - \left(E_{i,k} + E^f_{j,k}\right) c_{i,k} + P_{i,j,t} \end{aligned} \tag{5}$$

This model has linear features, we shall give the utility function based on the payoffs of player i in different strategies. The difference between the payoffs of "stay" and "leave" can be used to make decisions in terms of energy cost, which can be expressed with

$$W_i = P_{i,j} - P_{i,j,t,k} \tag{6}$$

We build a comparison in terms of energy consumption with respect to working states and future strategies.

$$\begin{cases} W_i > 0, & i \text{ stays in the task } j \text{ state;} \\ W_i < 0, & i \text{ leaves for a new task } k; \\ W_i = 0, & i \text{ has the same probability of stay or leave.} \end{cases}$$

$$\min_{s_i \in \Omega} \; U = \left| \sum_{i=1}^{N} W_i \right| = \left| \beta \sum_{i=1}^{N} P_{i,j} - \sum_{i=1}^{N} P_{i,j,t,k} \right| \tag{7}$$

where the parameter β is the weight of the preference of player i's action.

B. Reputation Mechanism

As part of the incentive mechanism, reputation promotes players to make a decision of stay or leave. The mechanism is divided into two categories: credit-exchange systems and reputation-based systems [29]. Generally, the reputation has two functions: 1) identifying the misbehaviour nodes by monitoring packet forwarding, and 2) helping the routing protocol avoid those nodes by informing the source node which has the selfish nodes on the path. In our study, players' behaviours are selfish and can be measured in a reputation-based system, by which the strategy behaviour of players will be suppressed for a reputation estimation.

Players usually have strategy tendency when they receive a task request. However, customers[1] intend to choose "suitable" players for their tasks such as the UAVs with high reputations and economic advantages. From the historic records, the fact that an UAV is rejected in the stage of being chosen implies the unsatisfactory working efficiency of the UAV in previous services. Our reputation mechanism can solve this problem when a customer needs to choose services in the future. Besides, the reputation mechanism facilitates the player's decision on whether it accepts the request, as an addition to the consideration of energy consumption. To be more specific, we consider a player's reputation at time t from two sides: customer's evaluation and system's compensation. The player i will receive an evaluation according to the performance, which can be defined by $V = \{0, 1, 2, 3, 4, 5\}$. The greater the value is, the more satisfactory the customer will gain. When the player switches to a new task without finishing the current task, its credits will be reduced. On the contrary, the player's credit at time t' will be added. This credit value will be considered into a new life-cycle of the game iteration. The goal of reputation models is to maximize the credit and integrate credit evaluation mechanism into the reputation model, meanwhile, maximize the profit of the operator. If player i has the current reputation r before joining the next task k, the new reputation will be

$$r_i' = r_i + \epsilon_{i,\cdot} \tag{8}$$

where r_i' is the reputation of the player i receiving the reputation increment $\epsilon_{i,j}$ or $\epsilon_{i,k}$. r_i is the initial reputation. We reference the modelling method of [30] and formulate the expected increment performance of i under constrained energy:

$$\begin{gathered} o_{i,j} = \mu \frac{E_i^o - E_i^r}{E_i^o} r_i \\ 0 \le E_i^r \le E_r^o \end{gathered} \tag{9}$$

where $o_{i,j}$ is the expected reputation increment of player i when it continues to complete task j with the remaining energy after time t. μ is the adjust coefficient of the reputation.

- If E_i^r is close to zero, $o_{i,j}$ is μr_i, and the reputation of player i will be $(1+\mu)r$. Player i will not leave for a new task.
- If E_i^r gets close to E_i^o, the reputation increment is close to zero.

Therefore,

$$\begin{cases} E_i^r \to E_i^o, & i \text{ reputation increment is } 0; \\ E_i^r \to 0, & i \text{ reputation increment is } \mu r_i. \end{cases}$$

The reputation increment of the player set can be concluded as

$$o_j = \sum_{i=1}^{N} \mu \frac{E_i^o - E_i^r}{E_i^o} r \tag{10}$$

Case 1: $E_i^r < E_{i,k} + E_i^f$, where E_k is the energy required for the task k. The residual energy E_i^r is not enough to finish the future task k in $\overline{II}$, the evaluation credit, from the customer side, for the task k is 0, and the reputation increment from system is

$$\begin{aligned} o_{i,j}' &= -\psi \frac{v_{i,j}}{sum} r_i \\ sum &= \sum_{v \in V} v \end{aligned} \tag{11}$$

where $v_{i,j}$ is the value that the player i receives from the customer. ψ is the adjustment coefficient of the evaluation value.

Case 2: $E_i^r > E_{i,k} + E_i^f$, where the player i can finish the task k within the energy E_i^r in place $\overline{II}$. In this case the reputation of player i will be updated to

$$o_{i,k}' = \psi \frac{v_{i,k}}{sum} r_i \tag{12}$$

where $o_{i,k}'$ is the reputation increment, and $v_{i,k} \in V$ is the evaluation value given by the customer.

[1]The party that is served by UAVs.

Therefore, the reputation of player *i* can be calculated as

$$\begin{aligned} r_i' &= r_i + \epsilon_i \\ &= \begin{cases} r_i + \eta o_{i,k} + \beta o_{i,k}', & E_i^r > \xi_1\left(E_{i,k} + E_i^f\right); \\ r_i + o_{i,k}, & \text{others} \end{cases} \end{aligned} \tag{13}$$

where η and β are the weight parameters, where $\eta + \beta = 1$, and $o_{i,k}'$ is used to update the reputation.

We consider that the reputation compensation brings about the incentives to players who accept tasks to obtain more reputation. The value of compensation is made up of players reputation and the emergency degree of tasks. This mechanism promotes the competition of players. In the two cases above, the reputation compensation model is given by

$$\begin{aligned} br_i &= f\left(r_{i,j}\right) = \lambda log_2\left(1 + \frac{r_i}{C}\right) v \\ v_k &= \delta E_i^r \end{aligned} \tag{14}$$

where br_i is the reputation compensation that the player *i* obtains. $f(r_{i,j})$ is the reputation compensation, and *C* is the threshold that the player *i* can trust. *v* is the value of task *k*, and δ is the coefficient related to the value.

In the next section, the game mechanism and the analysis process are introduced in detail.

VI. The Game Mechanism

Based on the modelling of energy consumption and reputation gain, the players strategy is constructed for the game mechanism. We first make a hypothesis and then discuss in two cases. We suppose

$$\begin{aligned} c_{i,k} &= \left(E_{i,k} + E_i^f\right)\bar{c}_i \\ E_{i,k} + E_i^f &= \bar{E}_k \end{aligned} \tag{15}$$

where $c_{i,k}$ is the average cost of the energy consumption that player *i* leaves for working on a new task *k*. $E_{i,k}$ is the energy consumption that player *i* spends on task *k*, and E_i^f is the consumed energy that player *i* transfers from previous task *j* to task *k* (*j* is the task that player *i* initially works on, and *k* is the potential task that player i transfers to work for. They are stated in the strategy model and have been described in the Fig. 4). Let $E_{i,k} + E_i^f = \bar{E}_k$ be the total energy cost that player *i* spends for its transferring. $\bar{c}_i$ is the average unit cost of various forms of energies. The cost of the reputation compensation of player *i* working on task *k* will be

$$\begin{aligned} cr_{i,k} &= c_{i,k} - br_{i,k} \\ &= \bar{E}_k\bar{c}_i - \lambda log_2\left(1 + \frac{r_i}{C}\right) v \end{aligned} \tag{16}$$

where $br_{i,k}$ is the reputation compensation of player *i* working on task *k*. Let $\xi_1 E_k$ and $\xi_2 E_k$ be the minimum and maximum energy that *i* requires to complete the task *k*. The expected energy is

$$E\left(\bar{E}_k\right) = \frac{\xi_1 E_k + \xi_2 E_k}{2} \tag{17}$$

Then, according to Equations (16) and (17), the expectation of reputation compensation is

$$\begin{aligned} E\left(cr_{i,k}\right) &= E\left(c_{i,k}\right) - E\left(br_{i,k}\right) \\ &= \left[\frac{\xi_1 E_k + \xi_2 E_k}{2}\right]\bar{c}_i - E\left(br_{i,k}\right) \end{aligned} \tag{18}$$

The probability of player *i* completing task *k* within its energy allowance will be

$$P\left(\bar{E}_k \le E_i^r\right) = \begin{cases} 0, & E_i^r < \xi_1 E_k; \\ \frac{E_i^r - \xi_1 E_k}{\xi_2 E_k - \xi_1 E_k}, & \xi_1 E_k \le E_i^r \le \xi_2 E_k; \\ 1, & E_i^r > \xi_2 E_k. \end{cases} \tag{19}$$

1) for *case 1*: when $\bar{E}_k < \xi_2 E_k$, the residual energy cannot support player *i* to complete the task *k*. The probability of player *i* performing the task *k* is

$$\begin{aligned} P_{\mathit{finish}}(i,k) &= P\left(\bar{E}_k \le E_k\right) \\ &= \frac{(1-\xi_1)E_k}{\xi_2 E_k - \xi_1 E_k} \\ &= \frac{1-\xi_1}{\xi_2 - \xi_1} \end{aligned} \tag{20}$$

Thus, the probability of not finishing the task *k* is

$$\begin{aligned} P_{\mathit{notfinish}} &= 1 - P_{\mathit{finish}}(i,k) \\ &= \frac{\xi_2 - 1}{\xi_2 - \xi_1} \end{aligned} \tag{21}$$

If player *i* can finish the task *k*, the expectation reputation increment will be

$$\begin{aligned} &E_{\mathit{finish}}\left(\epsilon_{i,k}\right) \\ &= P_{\mathit{finish}}(i,k)\left(\eta E\left(\mu\frac{E_k - \bar{E}_k}{E_k} r_i\right) + \beta E\left(o_{i,k}^{\prime}\right)\right) \\ &= \frac{1-\xi_1}{\xi_2 - \xi_1}\left(\eta\mu\frac{1-\xi_1}{2} r_i + \beta E\left(o_{i,k}^{\prime}\right)\right) \end{aligned} \tag{22}$$

Otherwise, the player *i* cannot finish the task *k* and the expectation reputation increment can be expressed as

$$\begin{aligned} &E_{\mathit{notfinish}}\left(\epsilon_{i,k}\right) \\ &= P_{\mathit{notfinish}}(i,k)\left(\eta E\left(-\mu\frac{\bar{E}_k - E_k}{E_k} r_i\right)\right) \\ &= \frac{\xi_2 - 1}{\xi_2 - \xi_1}\left(-\eta\mu\frac{\xi_2 - 1}{2} r_i\right) \end{aligned} \tag{23}$$

Therefore, the reputation increment in this case is

$$\begin{aligned} E_1\left(\epsilon_{i,k}^{\prime}\right) &= E_{\mathit{notfinish}}\left(\epsilon_{i,k}\right) + E_{\mathit{finish}}\left(\epsilon_{i,k}\right) \\ &= \frac{\xi_2 - 1}{\xi_2 - \xi_1}\left(-\eta\mu\frac{\xi_2 - 1}{2} r_i\right) \\ &\quad + \frac{1-\xi_1}{\xi_2 - \xi_1}\left(\eta\mu\frac{1-\xi_1}{2} r_i + \beta E\left(o_{i,k}^{\prime}\right)\right) \end{aligned} \tag{24}$$

2) for *case 2*: when $\bar{E}_k \ge \xi_2 E_k$, the residual energy can support the player *i* to finish task *k*. The reputation increment becomes

$$E\left(o_{i,k}\right) = E\left(\mu\frac{E_k - \bar{E}_k}{E_k} r_i\right)$$

$$= \mu \frac{E_k - E(\bar{E}_k)}{E_k} r_i$$
$$= \mu\left(1 - \frac{\xi_1 + \xi_2}{2}\right) \quad (25)$$

In this situation, the expected reputation increment is

$$E\left(o_{i,k}^{'}\right) = \sum_{v=1}^{5} \psi \frac{v_{i,k}}{sum} r_i \quad (26)$$

Then, the reputation increment in this case is

$$\begin{aligned} E_2\left(\epsilon_{i,k}^{'}\right) &= \eta E\left(o_{i,k}\right) + \beta E\left(o_{i,k}^{'}\right) \\ &= \eta\mu\left(1 - \frac{\xi_1 + \xi_2}{2}\right) + \beta \sum_{v=1}^{5} \psi \frac{v_{i,k}}{sum} r_i \quad (27) \end{aligned}$$

Therefore, the player i can receive the expected reputation in total with

$$\begin{aligned} r_i' &= r_i + E'\left(\epsilon_{i,k}\right) \\ &= r_i + \begin{cases} E_1\left(\epsilon_{i,k}^{'}\right), & E_k < \xi_2 E_k; \\ E_2\left(\epsilon_{i,k}^{'}\right), & E_k \geq \xi_2 E_k. \end{cases} \quad (28) \end{aligned}$$

For the next round of life-cycle of the game, the compensation based on tasks emergency degree and the benefit of the compensation $f_{i,k}$ is

$$f\left(r_i', k\right) = \lambda log_2\left(1 + \frac{r_i'}{C}\right) v \quad (29)$$

and

$$\begin{aligned} f_{i,k} &= f\left(r_i', k\right) - f\left(r_{i,j}\right) \\ &= \lambda log_2\left(1 + \frac{r_i'}{C}\right) v - \lambda log_2\left(1 + \frac{r_i}{C}\right) v \quad (30) \end{aligned}$$

We have the overall cost of reputation including the reputation increment as follows

$$\begin{aligned} c_{i,k}^{*} &= E\left(cr_{i,k}\right) - d_i f_{i,k} \\ &= \frac{\xi_1 E_k + \xi_2 E_k}{2} \bar{c}_i - \lambda log_2\left(1 + \frac{r_i}{C}\right) v_k \\ &\quad - d_i\left(\lambda log_2\left(1 + \frac{r_i'}{C}\right) v - \lambda log_2\left(1 + \frac{r_i}{C}\right) v\right) \end{aligned} \quad (31)$$

where $c_{i,k}^{*}$ is the cost of player i implementing task k, and d_i is the discounting coefficient of reputation compensation. When the expected payoff exceeds a payoff value, the player i considers to do task k.

$$\left\{c_{i,k}^{*}, +\infty\right\} \quad (32)$$

Since the payoff is related to the cost according to Equation (5), the expected payoff of player i doing task k will be

$$\begin{aligned} \Psi_{i,k} &= P_{i,j,t,k} - c_{i,k}^{*} \\ &= P_{i,k} + P_{i,j,t} - c_{i,k}^{*} \\ &= M_{i,k} p_k - \left(E_{i,k} + E_{j,k}^{f}\right) c_{i,k} + P_{i,j,t} - c_{i,k}^{*} \\ &= M_{i,k} p_k - \left(\frac{\xi_1 E_k + \xi_2 E_k}{2}\right) c_{i,k} + P_{i,j,t} - c_{i,k}^{*} \end{aligned} \quad (33)$$

Let Ψ be the payoff. The cost needs to be minimized, and the equilibrium is assumed to exist in the fixed workload

$$\begin{aligned} &\text{Min} \;\; c_i^{*} \\ &\text{Max} \;\; \Psi_i \\ &\qquad \forall i \in I \end{aligned} \quad (34)$$

The purpose of this part is to search the Pareto optimal for this optimization problem. There are several approaches which convert approximation problems into number of scalar optimization problems. These methods include weighted sum approach [31], Tchbycheff approach and boundary intersection approach [32], [33], [34], [35]. Let $\omega = (\omega_1, \omega_2)^T$ be a weight vector, and $\omega_i \geq 0$ for all $i = 1, 2$ and $\omega_1 + \omega_2 = 1$. Then, the optimal solution to the following scalar problem can be concluded

$$\begin{aligned} &\text{Max} \;\; \Phi(i|\omega) = \omega_1 \sum_{i=1}^{2} \omega_i \Psi(i) \\ &\text{subject to} \;\; s_i \in \Omega \end{aligned} \quad (35)$$

It exists a Pareto optimal point in the above objective function, where ω is characterised to be a coefficient vector and i is the variable to be optimized. The extrema $\Phi'(i|\omega)$ proves the equilibrium point. Our proposed mechanism is characterised in Algorithm 1.

The lines 1 and 2 represent the original input of the algorithm. From line 3, the ϵ_i can be updated according to the Equation (19)-(23). The $E_1(\epsilon_{i,k}^{'})$ or $E_2(\epsilon_{i,k}^{'})$ can be achieved by Equations (24) and (25). From line 13 to 22 of the algorithm, the extrema $\Phi'(i|\omega)$ is reached by iterating the value of $\Phi(i|\omega)$, which is corresponding to the Equations of (28)-(35) in the above derivations. The meaning details of the lines can also be referenced in the models and mechanism.

VII. Experimental Results and Analysis

In this section, we provide numerical results to evaluate the effects of critical parameters to the strategy of UAVs scheduling, by which the fluctuations of hyper-parameters (such as energy cost, reputation etc.) can be indicated to the mechanism. The scenario is as follows: 1 control center was defined, two UAVs were pointed in working situation and under the control of the center. Only one task request was broadcast to UAVs by control center at each time. The number of tasks of each service was 200, and the original energy of each UAV had 2000 J.

Our work can be divided into the sections: 1) the system belief determination, 2) the effects of dominant parameters to the energy cost and payoff, and 3) the relationships among important paired hyper-parameters including energy consumption and reputation, tasks and reputation, tasks achievement and reputation increment, as well as the parameters that directly affect the hyper-parameters.

Algorithm 1 Game of Energy Consumption and Reputation

Input:
Task M_j;
The original energy capacity of player i, $E^o_{i,j}$;
The average unit price of energy used to task M_j, p_j;
The residual energy after working in j,$E^r_{i,j}$;
The energy cost for player transfer,$E^f_{j,k}$;
The unit cost of player i working on task $M_{i,j}$, $c_{i,j}$;
The time of player i working in task j, t;
The energy cost of i working in j until time t, $c_{i,j,t}$;
Task M_k;
The cost of player i working on task $M_{i,k}$, $c_{i,k}$;
The average unit price of energy used to task M_k, p_k;
α and β: weight parameters
Adjust coefficient of reputation, μ;
Adjust coefficient of evaluation value, ψ;
Coefficient value, λ.
Weights, $\omega = (\omega_1, \omega_2)$
Output:
Payoff $\Psi_{i,j}$, $\Psi_{i,k}$

1: $(E^o_i - E^r_{i,j}) \leftarrow 0$, $P_{i,j,t} \leftarrow 0$, $P_{i,j} \leftarrow 0$, $o_j \leftarrow 0$, $o_k \leftarrow 0$
2: $U \leftarrow P_{i,j} - P_{i,j,t,k}$
3: **if** $E^r_i > \xi_1(E_{i,k} + E^f_i)$ **then**
4: $\quad \epsilon_i \leftarrow \eta o_{i,k} + \beta o'_{i,k}$
5: **else**
6: $\quad \epsilon_i \leftarrow o_{i,k}$
7: **end if**
8: $r'_i \leftarrow r_i + \epsilon_i$
9: **if** $\bar{E}_k < \xi_2 E_k$ **then**
10: $\quad E_1(\epsilon'_{i,k}) \leftarrow 0$
11: **else**
12: $\quad E_2(\epsilon'_{i,k}) \leftarrow 0$
13: $\quad f_{i,k} \leftarrow \lambda log_2(1 + \frac{r'_i}{C})v - \lambda log_2(1 + \frac{r_i}{C})v$
14: $\quad c*_{i,k} \leftarrow f_{i,k}$
15: $\quad \Psi_{i,k} \leftarrow c*_{i,k}$
16: $\quad \Phi'(i|\omega) \leftarrow \Psi(i|\omega)$
17: $\quad$ **if** $\Phi'(i|\omega)$ is extrema **then**
18: $\quad\quad$ extrema is the equilibrium point
19: $\quad$ **else**
20: $\quad\quad$ Go back to Line 3 until converging to an extrema of $\Phi'(i|\omega)$
21: $\quad$ **end if**
22: **end if**

A. Experimental Settings and Belief Determination

We assume that all UAVs are randomly distributed in a working region, and a control center can interact with them constantly. A large number of iterations are set for 10000 to demonstrate the convergence process of belief. The parameters are selected as follows:

- Number of UAVs: 2
- Tasks $M_j = M_k = 200$
- The original energy capacity of player i, $E^o_{i,j} = 2000$;
- The average unit price of energy used to task M_j, $p_j = 10$
- The energy cost for player transfer, $E^f_{j,k} = 200$;

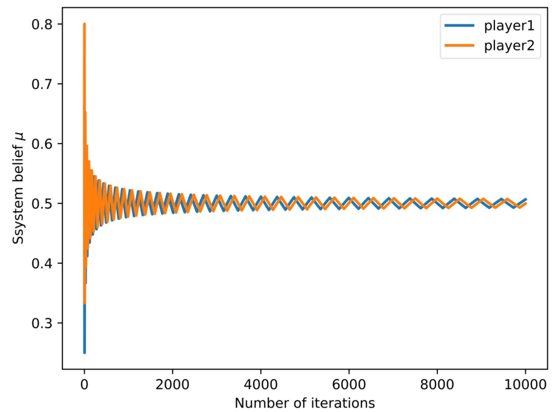


Fig. 5. Belief determination.

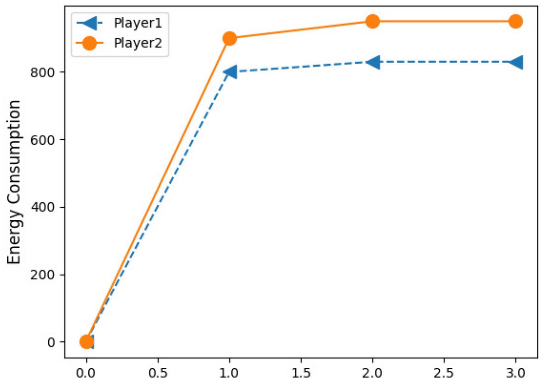
- The unit cost of player i working on task $M_{i,j}$, $c_{i,j} = 10$.
- Adjust coefficient of evaluation value, $\psi = 0.5$
- Weights, $\omega = (\omega_1, \omega_2) = (0.5, 0.5)$

The detailed parameters for comparison are set in subsections. Therefore, each UAV has 200 tasks to run. We set the initial payoff to 400, the reputation weight η to 0.8, and coefficient of reputation compensation λ to 0.2. Based on these settings, we can adjust the parameters according to the obvious fluctuation of energy consumption and payoff caused by reputations in the following experiments. The detailed parameters settings and analysis on ξ_1 and ξ_2, and other parameters are presented in the following subsections. Fig. 5 shows the fluctuations of system beliefs of both players along with the increasing number of iterations, where the approximation value of 0.5 means the equal probability of choosing the two strategies.

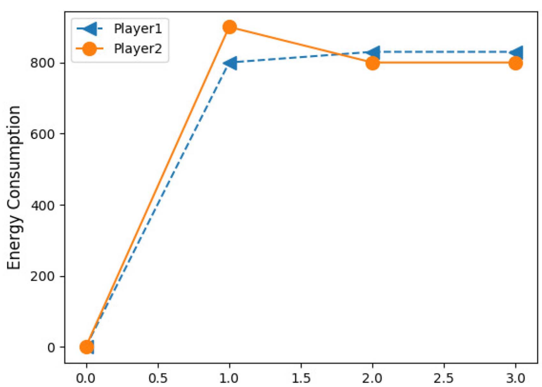
As we described in the preliminary work, Bayes rule implies the perfect equilibrium in belief mechanism, where the probability distribution of possible payoffs of players can be derived with incomplete information of other players [36]. Different system beliefs will produce different payoffs of players, which can cause the uncertainty of the ultimate payoff. Sarangi [37] analysed the beliefs in the epistemic logic sense and argued the evident importance of networks with incomplete information. Belief updating rule was introduced for sequential games in [38], and the equilibrium predictions differing from the original rule were compared to the new one. The Belief structure of novel matrix game was described in [39], and heterogeneous cooperative belief for social dilemma in multi-agent system was clarified in [40]. Therefore, a reliable belief in a game process is of great importance.

B. Numerical Results

According to the above result, the belief is set up as $\mu = 0.5$. This setting of belief will show different payoffs of players under the two strategic options with the fluctuation of reputation. In our work, an UAV faces two strategic options (stay and leave), which will bring the variation of energy consumption and the change of reputation. These two metrics are treated as the main factors affecting payoffs of players. We first set related parameters and then adjust them accordingly. The

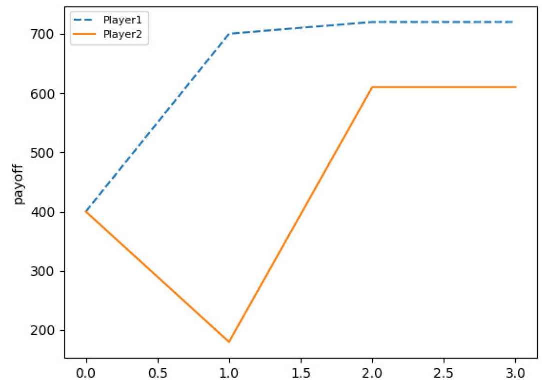


(a) The energy consumption obtained by the proposed model under the parameter $\eta = 0.8$

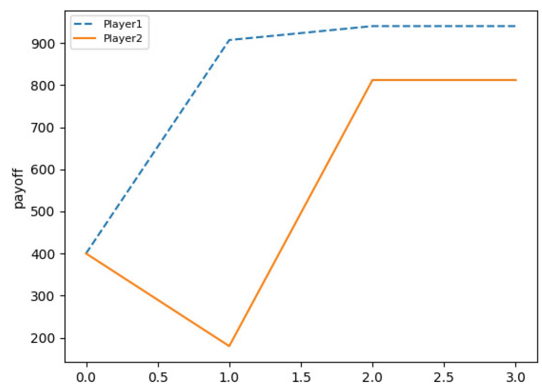


(b) The energy consumption obtained by the proposed model under the parameter $\eta = 0.5$

Fig. 6. The effect of the parameter η on energy consumption.



(a) The payoff obtained from the proposed model under the parameter $\lambda = 0.2$



(b) The payoff obtained from the proposed model under the parameter $\lambda = 0.5$

Fig. 7. The effect of the parameter λ on payoff.

reputation compensation $\epsilon = 0.2$, the weight of the action preference of players $\alpha = 0.5$, coefficient of trust value of players $\delta = 0.5$, the minimum weight of required energy $\xi_1 = 0.5$, the maximum weight of required energy $\xi_2 = 1$, the weight of expected payoff $\omega_1 = \omega_2 = 0.5$, and the original energy of a player $E_i^o = 2000$. The required energy for task *j* and *k* are 600. The average unit price of energy of an UAV staying in the task *j* is 10, which is equal to the average price of unit energy for an UAV leaving for a new task *k*. In what follows, the effects of different parameters are investigated.

The effect of parameter η on energy consumption. As shown in Fig. 6, the parameter η indicates the leading weight of an increment in reputation of a player in task *k*. Compared to the weight of the reputation increment β, it should be more than half of the weight, so we set $\eta = 0.8$. Thus, we have $\beta = 0.2$ because $\eta+\beta = 1$ is associated with the reputation expectation in Eq. (13). Figs. 6(a) and 6(b) show the effect of parameter η on energy consumption. We set $\eta = 0.5$ and $\eta = 0.8$ to compare different energy consumption under different parameter values. Operators always intend to reduce the energy consumption so as to maximize their profits. When $\eta = 0.8$ or 0.5, the energy consumption of two players are rising. The reduction of energy consumption shown at inflection point in Fig. 6(b), is because the UAV stops working for the current task and transfers to another site for a new task. The larger value of η than 0.8 lead to more energy consumption but the smaller of η than 0.5 brings about the unstable reputation. The latter η shown in Fig. 6(b) presents the energy efficiency apparently.

The effect of parameter λ on payoff. As shown in Figs. 7(a) and 7(b), the payoff is designed from the perspective of the energy consumption of executing task *j* (or partial executions of tasks *j* and *k* in the case that the UAV moves to perform the new task *k* after this UAV leaving the working site of task *j*), as well as reputation changes. We notice that the payoff is increasing according to the figures. The reputation shown in x-axis means that an UAV has chosen a strategy, the reputation is going up but the payoff would not change for its maximum. From the previous model, we know that the payoff is not only affected by parameter λ, but also by parameter η which associates with the compensation of reputation. We investigate the effect of payoff by setting $\lambda = 0.2$ and 0.5, respectively. The distinct difference of the payoff in Figs. 7(a) and 7(b) lies in the amount of payoff, even though they are in similar trends. In Fig. 7(b), the payoff of player 1 can achieve approximate 950, after which it stops growing up. This means it is in the equilibrium point. However, in Fig. 7(a), the setting of λ contributes to reach the equilibrium point but the payoff is less

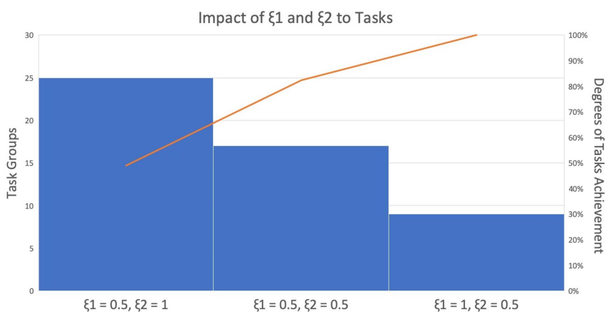


Fig. 8. The effects of parameters ξ_1 and ξ_2 on tasks.

than that shown in Fig. 7(b). The difference between players 1 and 2 in Fig. 7(a) is more significant than that in Fig. 7(b). It is necessary to clarify that in our experiment, equilibrium point dose not always exist when we adjust the parameters with extreme values. In this case, the payoff results do not converge to stable values. From our results, we are noticed that the energy consumption and the reputation associated payoff contribute to UAV scheduling strategies.

The effect of parameter μ on reputation fluctuation. This parameter affects player's reputation increment, and it integrates η of players reputations into two different working situations (finish or not). As we know, μ is in range of 0 $\sim$ 1. The fluctuation of reputation is shown in Figs. 6(a) and 6(b), from which a discount exists for the energy consumption so that the stable fluctuations are from 1.0 to 3.0. With the extensive experiments, it is observed that when η is 0.8, the lines indicate that the energy consumption and reputation are positively correlated. However, when the μ equals to 0.5, the energy consumption of player 2 goes down and then maintains its steady in a level.

The effects of ξ_1 and ξ_2 on tasks. As shown in Fig. 8, these two parameters stand for the minimum and maximum coefficients of energy consumption required to finish the task k. We keep the other parameters fixed and similar with the discussion above to set the value $\xi_1 = 0.5$, $\xi_2 = 1$. When an UAV needs to take a strategy option of stay or leave, the required energy consumption of finishing the tasks is oscillatory due to task loads. According to Equation (33), when we predict the energy expectation, the maximum coefficient with less than 0.5 will increase the payoff with the purpose of Ψ_i maximization, but it will stagnate the reputation increment which is described in Equation (31). The setting of $\xi_1 = 0.5$, $\xi_2 = 1$ derives the expectation value which satisfies the purpose of dependability of energy requirement by reducing errors. We divide our experimental results into 30 groups to indicate the effects of ξ_1 and ξ_2 on tasks. The right y-axis represents the percentage that different value combinations of ξ_1 and ξ_2. As shown in Fig. 8, we can see that when $\xi_1 = 0.5$, $\xi_2 = 1$, the accomplishment degree is the highest.

The compensation trend of reputation. In Fig. 9, we demonstrate the compensation trend of reputation under the time frame based on the prior parameters' setting. According to the performance of energy consumption in Fig. 6 (a) and Fig. 6(b), and the payoffs achievement in Fig. 7 (a) and Fig. 7(b), the critical parameters $\eta = 0.5$ and $\lambda = 0.5$ are appointed to

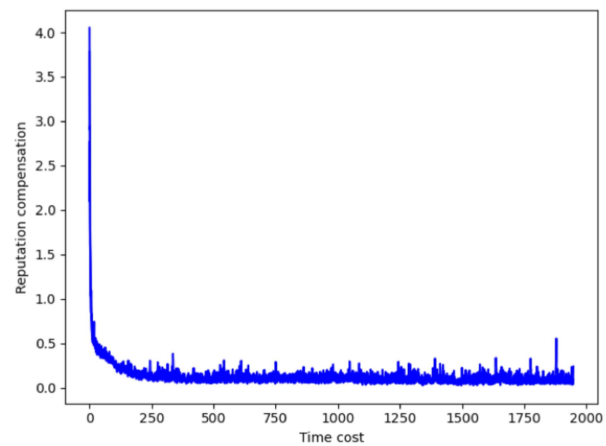


Fig. 9. The compensation trend of reputation with time frame.

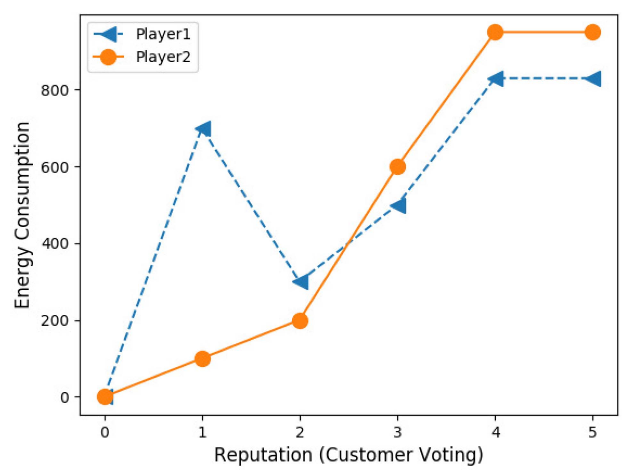


Fig. 10. Energy consumption vs. reputation increase.

conduct this experiment for the low energy consumption and the high payoffs. The trend from the beginning to 250s shows a sharp decline and then tends to be stable in the following time slots. The slight fluctuations of the compensation performance in the following time slots imply the restrictions of energy to the compensation change which have been theoretically supported by Equations (16), (18), (24) and Equation (27).

The relationships of hyper parameters.

- *Energy Consumption* and *Reputation:* As shown in Fig. 10, based on the performance of the UAV and the degree of tasks fulfilment, the score from customers are assigned according to the historic records. In our experiments, we input the initial score randomly from 0 $\sim$ 5 and depict them in X-axis. The energy consumption which reflects the achievement situation of tasks is described in Y-axis. Their relationships with given reputations are displayed in Fig. 10. The lowest point of blue line is provided with value 2 of reputation score, which means that the UAV changes its task upon the control center request. The rising tread of energy consumption of this player indicates its flying action and the execution of the new job. We notice that the main reason of player 1 changing its working site is due to the result of the game process. However, player 2 sticking in the current working site

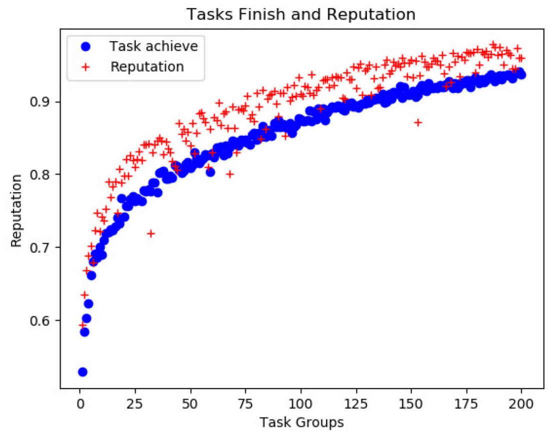


Fig. 11. Tasks finish vs. reputation increase.

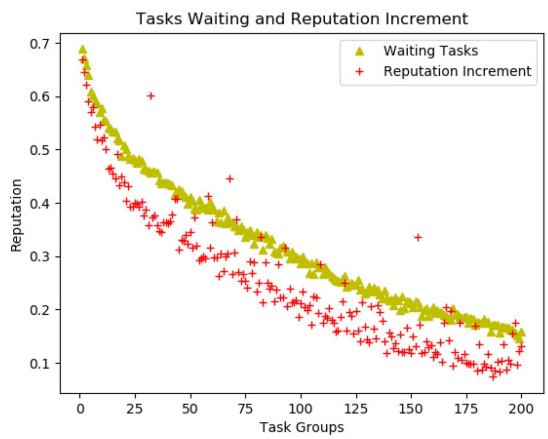


Fig. 12. Task Waiting vs. Reputation Increment.

causes the continuous increase in energy consumption, which is explicated in the orange line.

- *Tasks* and *Reputation:* As shown in Fig. 11, we consider that each UAV has 200 tasks to run, which is described in the x-axis. The normalization of reputation is represented by y-axis. The blue line indicates that the achievement of tasks is in steady rise. During the process, the reputation, expressed by red cross, goes up with the extent of tasks fulfilment. Above the blue line, the red cross points show the reputation change of the UAV due to the shift of working tasks.
- *Tasks Finish* and *Reputation Increment:* Based on the above analysis, the tasks waiting to be executed and the reputation increment can be seen in Fig. 12. The downward trend of yellow line of waiting tasks means that the waiting tasks decrease along with the time passes. The trend of yellow scatters demonstrate the decrease of waiting tasks and the red scatters stand for the falling of reputation increment. This figure demonstrates that the more tasks completed, the less of reputation increment obtained.

VIII. Conclusion

UAVs scheduling strategy is one of the typical mobile edge computing in IoT applications. In this paper, we analyzed an UAVs application scenario and investigated specific challenges in strategic scheduling. The great requirement in challenging is that the decrease of energy consumption and increase of the reputation should be guaranteed simultaneously. In order to address the problem, we formulate the energy consumption and reputation model in a multiple tasks scenario with respect to the commercial profit maximization of network operators. The payoff-credits model, energy model and reputation model have been built respectively for the basis of the mechanism. Through the game framework combined with reputation mechanism, the equilibrium point is reached in our experiment. This mechanism includes energy consumption and the change of reputation, therefore, we have analyzed the performance of the UAV by a few groups of hyper-parameters (such as energy consumption and reputation, tasks finish and reputation increment etc.). The significant parameters impacting the performance of energy cost, reputation, payoff and tasks are also analyzed.

Through the performance analysis of energy consumption and the reputation, it can be clearly seen that the increment of the reputation does not always cause the continuous increase of energy, which is the opposite of the intrinsic view. The trend and performance of the energy consumption and payoff show the superiority of the selected UAV. This advantage implies that the mechanism can be considered for the long-term profit maximization of operators. In addition, the developed model can be further utilized in similar optimal payoff issues in the future.

References

[1] S. Koulali, E. Sabir, T. Taleb, and M. Azizi, "A green strategic activity scheduling for UAV networks: A sub-modular game perspective," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 58–64, May 2016.
[2] M. Campion, P. Ranganathan, and S. Faruque, "UAV swarm communication and control architectures: A review," *J. Unmanned Veh. Syst.*, vol. 7, no. 2, pp. 93–106, 2019.
[3] H. Hildmann and E. Kovacs, "Using unmanned aerial vehicles (UAVs) as mobile sensing platforms (MSPs) for disaster response, civil security and public safety," *Drones*, vol. 3, no. 3, p. 59, 2019.
[4] H. Ghazzai, A. Kadri, M. B. Ghorbel, H. Menouar, and Y. Massoud, "A generic spatiotemporal UAV scheduling framework for multi-event applications," *IEEE Access*, vol. 7, pp. 215–229, 2018.
[5] R. Trestian, O. Ormond, and G.-M. Muntean, "Reputation-based network selection mechanism using game theory," *Phys. Commun.*, vol. 4, no. 3, pp. 156–171, 2011.
[6] H. Wu, H. Li, R. Xiao, and J. Liu, "Modeling and simulation of dynamic ant colony's labor division for task allocation of UAV swarm," *Physica A, Stat. Mech. Appl.*, vol. 491, pp. 127–141, Feb. 2018.
[7] S. Waharte and N. Trigoni, "Supporting search and rescue operations with UAVs," in *Proc. Int. Conf. Emerg. Security Technol.*, Canterbury, U.K., 2010, pp. 142–147.
[8] G. Li, X. Zhou, J. Yin, and Q. Xiao, "An UAV scheduling and planning method for post-disaster survey," *Int. Archives Photogrammetry Remote Sens. Spat. Inf. Sci.*, vol. 40, no. 2, pp. 169–172, 2014.
[9] L. Bertuccelli, W. Beckers, and M. Cummings, "Developing operator models for UAV search scheduling," in *Proc. AIAA Guid. Navig. Control Conf.*, 2010, p. 7863.
[10] R. Dai, S. Fotedar, M. Radmanesh, and M. Kumar, "Quality-aware UAV coverage and path planning in geometrically complex environments," *Ad Hoc Netw.*, vol. 73, pp. 95–105, May 2018.
[11] S.-Y. Lin and P.-H. Chou, "A dynamic data driven-based semi-distributed reputation mechanism in unknown networks," *Electron. Commer. Res. Appl.*, vol. 14, no. 6, pp. 532–541, Oct./Nov. 2015.

[12] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du, and M. Guizani, "Evaluating reputation management schemes of Internet of vehicles based on evolutionary game theory," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5971–5980, Jun. 2019.

[13] M. E. Mkiramweni, C. Yang, J. Li, and Z. Han, "Game-theoretic approaches for wireless communications with unmanned aerial vehicles," *IEEE Wireless Commun.*, vol. 25, no. 6, pp. 104–112, Dec. 2018.

[14] S. Mousavi, F. Afghah, J. D. Ashdown, and K. Turck, "Use of a quantum genetic algorithm for coalition formation in large-scale UAV networks," *Ad Hoc Netw.*, vol. 87, pp. 26–36, May 2019.

[15] V. Crespi, W. W. Chung, and A. B. Jordan, "Decentralized sensing and tracking for UAV scheduling," in *Proc. Sens. Command Control Commun. Intell. (C3I) Technol. Homeland Security Homeland Defense III*, vol. 5403. Orlando, FL, USA, 2004, pp. 451–463.

[16] F. B. Sorbelli, C. M. Pinotti, and V. Ravelomanana, "Range-free localization algorithm using a customary drone," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Taormina, Italy, 2018, pp. 9–16.

[17] M. Hua, Y. Wang, Z. Zhang, C. Li, Y. Huang, and L. Yang, "Power-efficient communication in UAV-aided wireless sensor networks," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1264–1267, Jun. 2018.

[18] S. K. K. Hari, K. Sundar, S. Rathinam, and S. Darbha, "Scheduling tasks for human operators in monitoring and surveillance applications," *IFAC-PapersOnLine*, vol. 49, no. 32, pp. 54–59, 2016.

[19] M. Rosalie, G. Danoy, S. Chaumette, and P. Bouvry, "Chaos-enhanced mobility models for multilevel swarms of UAVs," *Swarm Evol. Comput.*, vol. 41, pp. 36–48, Aug. 2018.

[20] G. Cai and J. Huang, "A new finance chaotic attractor," *Int. J. Nonlinear Sci.*, vol. 3, no. 3, pp. 213–220, 2007.

[21] M.-A. Messous, H. Sedjelmaci, and S.-M. Senouci, "Implementing an emerging mobility model for a fleet of UAVs based on a fuzzy logic inference system," *Pervasive Mobile Comput.*, vol. 42, pp. 393–410, Dec. 2017.

[22] D.-K. Phung and P. Morin, "Modeling and energy evaluation of small convertible UAVs," *IFAC Proc. Vol.*, vol. 46, no. 30, pp. 212–219, 2013.

[23] V. Sharma, F. Song, I. You, and M. Atiquzzaman, "Energy efficient device discovery for reliable communication in 5G-based IoT and BSNs using unmanned aerial vehicles," *J. Netw. Comput. Appl.*, vol. 97, pp. 79–95, Nov. 2017.

[24] C. Tang, L. Wu, G. Wen, and Z. Zheng, "Incentivizing honest mining in blockchain networks: A reputation approach," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 1, pp. 117–121, Jan. 2020.

[25] Y. Ren, G. Wang, L. Yu, B. Shi, W. Hu, and Z. Wang, "Rigorous or tolerant: The effect of different reputation attitudes in complex networks," *Future Gener. Comput. Syst.*, vol. 83, pp. 476–484, Jun. 2018.

[26] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, "Toward secure blockchain-enabled Internet of vehicles: Optimizing consensus management using reputation and contract theory," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2906–2920, Mar. 2019.

[27] W. Dai, H. Lu, J. Xiao, and Z. Zheng, "Task allocation without communication based on incomplete information game theory for multi-robot systems," *J. Intell. Robotic Syst.*, vol. 94, nos. 3–4, pp. 841–856, 2019.

[28] D. Levin and L. Zhang. (2019). *Bridging Level-K to Nash Equilibrium*. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2934696

[29] J. J. Jaramillo and R. Srikant, "DARWIN: Distributed and adaptive reputation mechanism for wireless ad-hoc networks," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw.*, 2007, pp. 87–98.

[30] Z. Su, Q. Qi, Q. Xu, S. Guo, and X. Wang, "Incentive scheme for cyber physical social systems based on user behaviors," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 1, pp. 92–103, Jan.–Mar. 2020.

[31] J. Guo, Z. Song, Y. Cui, Z. Liu, and Y. Ji, "Energy-efficient resource allocation for multi-user mobile edge computing," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, 2017, pp. 1–7.

[32] K. Kaur, S. Garg, G. S. Aujla, N. Kumar, J. J. Rodrigues, and M. Guizani, "Edge computing in the industrial Internet of Things environment: Software-defined-networks-based edge-cloud interplay," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 44–51, Feb. 2018.

[33] O. Väänänen and T. Hämäläinen, "Requirements for energy efficient edge computing: A survey," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Cham, Switzerland: Springer, 2018, pp. 3–15.

[34] Z. Li, T. Yang, and Z. Feng, "The multiobjective trajectory optimization for hypersonic glide vehicle based on normal boundary intersection method," *Math. Problems Eng.*, vol. 2016, Jun. 2016, Art. no. 9407238.

[35] I. Akhtar, L. Li, X. Li, W. Chen, and Z. Han, "Weighted sum-rate maximization for cooperative multiceli multiuser massive MIMO systems based on discretized pareto boundary approximation," in *Proc. 27th Wireless Opt. Commun. Conf. (WOCC)*, Hualien, Taiwan, 2018, pp. 1–5.

[36] F. Li and P. Norman, "On Bayesian persuasion with multiple senders," *Econ. Lett.*, vol. 170, pp. 66–70, Sep. 2018.

[37] S. Sarangi, "Exploring payoffs and beliefs in game theory," Ph.D. dissertation, Dept. Doctor Philos. Econ., Virginia Polytechn. Inst. State Univ., Blacksburg, VA, USA, 2000.

[38] L. Jiang and J. Wu, "Belief-updating rule and sequential reciprocity," *Games Econ. Behav.*, vol. 113, pp. 770–780, Jan. 2019.

[39] Y. Han and Y. Deng, "A novel matrix game with payoffs of maxitive belief structure," *Int. J. Intell. Syst.*, vol. 34, no. 4, pp. 690–706, 2019.

[40] K. Huang, X. Chen, Z. Yu, C. Yang, and W. Gui, "Heterogeneous cooperative belief for social dilemma in multi-agent system," *Appl. Math. Comput.*, vol. 320, pp. 572–579, Mar. 2018.

Juan Zhang received the B.Eng. degree in mechanical design, manufacture and its automation and the M.Sc. degree in control science and engineering from Beihang University, China, in 2010, and in 2014, respectively. She is currently pursuing the Ph.D. degree with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, U.K. Before joining the current group, she worked as a Research Assistant with the Department of Computer Science, University of Otago, New Zealand, from 2015 to 2017. Her research is in mobile edge computing for future mobile networks, Internet of Things, and decision-making strategies in mobile networks.

Yulei Wu (Senior Member, IEEE) received the B.Sc. degree (First Class Hons.) in computer science and the Ph.D. degree in computing and mathematics from the University of Bradford, U.K., in 2006 and 2010, respectively. He is a Senior Lecturer with the Department of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, U.K. His research has been supported by Engineering and Physical Sciences Research Council of United Kingdom, the National Natural Science Foundation of China, and the University's Innovation Platform and industry. His expertise is on networking and his main research interests include computer networks, networked systems, software-defined networks and systems, network management, and network security and privacy. He contributes to major conferences on networking and networked systems as various roles, including the Steering Committee Chair, the General Chair, and the Program Chair. He is an Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, *Computer Networks* (Elsevier), and IEEE ACCESS. He is a Fellow of the Higher Education Academy.

Geyong Min (Member, IEEE) received the B.Sc. degree in computer science from the Huazhong University of Science and Technology, China, in 1995, and the Ph.D. degree in computing science from the University of Glasgow, U.K., in 2003. He is a Professor of high performance computing and networking with the Department of Computer Science within the College of Engineering, Mathematics and Physical Sciences, University of Exeter, U.K. His research interests include computer networks, wireless communications, parallel and distributed computing, ubiquitous computing, multimedia systems, modeling, and performance engineering.

Fei Hao received the B.Sc. degree in information and computing science and the M.Sc. degree in computer software and theory from Xihua University, China, in 2005 and 2008, respectively, and the Ph.D. degree in computer science and engineering from Soonchunhyang University, South Korea, in 2016. Since 2016, he has been with Shaanxi Normal University, Xi'an, China, where he is an Associate Professor. He is currently taking a Marie Sklodowska-Curie Individual Fellowship at the University of Exeter, Exeter, U.K. His research interests include social computing, ubiquitous computing, big data analysis and processing, and mobile cloud computing.

Laizhong Cui (Senior Member, IEEE) received the B.S. degree from Jilin University, Changchun, China, in 2007, and the Ph.D. degree in computer science and technology from Tsinghua University, Beijing, China, in 2012. He is currently a Professor with the College of Computer Science and Software Engineering, Shenzhen University, China. His research interests include Future Internet architecture and protocols, edge computing, multimedia systems and applications, blockchain, Internet of Things, cloud and big data computing, software-defined network, social network, computational intelligence, and machine learning. He led more than 10 scientific research projects, including the National Key Research and Development Plan of China, the National Natural Science Foundation of China, the Guangdong Natural Science Foundation of China, and Shenzhen Basic Research Plan. He has published more than 70 papers, including the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE INTERNET OF THINGS JOURNAL, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, the *ACM Transactions on Internet Technology*, the IEEE TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, and IEEE NETWORK. He serves as an Associate Editor or a Member of Editorial Board for several international journals, including the *International Journal of Machine Learning and Cybernetics*, the *International Journal of Bio-Inspired Computation*, *Ad Hoc and Sensor Wireless Networks*, and the *Journal of Central South University*. He is a Senior Member of the CCF.