OPTIMAL TASK OFFLOADING IN FOG-ENABLED NETWORKS VIA INDEX POLICIES

Fuqian Yang, Zhaowei Zhu, Shangshu Zhao, Yang Yang, and Xiliang Luo

ShanghaiTech University, Shanghai, China
E-mail: {yangfq, zhuzhw, zhaoshsh, yangyang, luoxl}@shanghaitech.edu.cn

ABSTRACT

Fog computing has been considered to be a potential solution to enable computation-intensive and latency-critical application at the battery-empowered mobile devices. Task offloading could take advantages of available neighboring computational resources to achieve low latency. By exploiting the temporal correlation of the states at fog nodes, a sequential decision-making problem that aims to optimize the long-term task offloading performance is considered in this paper. Such a problem is a partially observed Markov decision process with an exploitation-exploration tradeoff that is difficult to analyze. We address this tradeoff in task offloading under the framework of restless multi-armed bandits (RMAB). The indexability analysis of this task offloading problem is then provided. Meanwhile, an index policy, which is asymptotically optimal and has remarkably low computation complexity, is established based on the Whittle's index to solve the task offloading problem. Numerical results show the superiority of the proposed task offloading method.

Index Terms— Fog computing, task offloading, restless multi-armed bandits problem, RMAB, index policy.

1. INTRODUCTION

Widespread applications of the Internet of Things (IoT), 5G wireless systems, and the embedded artificial intelligence in recent years require energy-efficient data processing capability at user equipments (UEs) [1, 2]. For battery-empowered mobile devices, fog computing (or mobile edge computing) has been considered to be a potential solution to achieve low latency [2]. To exploit the benefits of all the available computational resources, fog computing is employed to distribute computing, storage, control, and communication services along the cloud-to-thing continuum [1, 2].

Task offloading is one important problem in fog computing [2, 3, 5, 6, 8–10]. Among the literatures, some researchers addressed the energy issues and formulated task offloading as deterministic optimization problems [2–4]. When the real-time states, e.g. the computation queue length, of UEs and servers are considered, the task offloading problem becomes a typical stochastic optimization problem. To tackle the problem that the future system information was difficult to predict, the Lyapunov optimization method was invoked in [5–9] to transform the challenging stochastic optimization problem to a sequential decision problem, which included a series of deterministic problems in each time slot and required only the current system information. Assuming system parameters were unknown, the tradeoff between learning the system parameters and pursuing the empirically optimal offloading strategy was investigated under the bandit model in [10].

These task offloading methods ignored the temporal correlation of fog node states. In this paper, we assume the parameters of fog nodes are not available at UEs but the states at fog nodes are temporally correlated. The state of a fog node indicates whether the computational resources are sufficient or not. To characterize the temporal correlations of states at fog nodes, we employ a two-state Markov chain, which is popular in studying time-correlated processes [11], to model the evolution of the state at a fog node. If a task is offloaded to a fog node with sufficient computational resources, the UE can obtain a high reward. At the beginning of each slot, the UE needs to select some fog nodes for task offloading such that the long-term reward over an infinite horizon is maximized. The task offloading decision in each slot is associated with a fundamental conflict between taking actions that yield high current rewards and taking actions that sacrifice current gains with the prospect of reaping better future returns [12], which is the classical exploitation-exploration tradeoff and is difficult to analyze. To tackle this problem, we analyze the problem under the framework of restless multi-armed bandits (RMAB) [13]. Specifically, the paper establishes an RMAB framework of task offloading from a UE to multiple fog nodes in fog-enabled networks. In order to solve the task offloading problem via an index policy, we further analyze the indexability of the task offloading problem and obtained the closed-form Whittle's index expressions in different cases. Numerical results show that our proposed Whittle's index policy for task offloading can achieve better performance than other existing policies.

2. SYSTEM MODEL AND PROBLEM STATEMENT

We consider a fog-enabled network where the M fog nodes can offer computational resources and a UE can simultaneously offload K tasks to K out of M fog nodes in each time slot. Note we have assumed K orthogonal channels are available for a UE to offload tasks. Consider that the UE does not know the accurate information about the parameters of fog nodes, but can only obtain the knowledge about fog node states which indicate whether the computational resources at fog nodes are sufficient or not. For example, if the fog nodes are equipped with energy harvesting devices [14], the energy harvesting states can be treated as the fog node states, which can be predicted by the UE. At the beginning of each slot, the UE selects K out of N fog nodes to offload K tasks with knowledge of sates at fog nodes.

The fog node states are assumed to be temporally correlated. Specifically, the evolution of the state at a fog node is considered to be a two-state Markov chain. Each chain has two states denoted by states 1 (good) and 0 (bad), respectively. See Fig. 1. Let $X_m(t)$ be the state of fog node m in slot t. The probability transition matrix of the two-state Markov chain in fog node m is given by

$$P_m = \begin{bmatrix} p_m & 1 - p_m \\ q_m & 1 - q_m \end{bmatrix}, \tag{1}$$

where $p_m := \Pr\{X_m(t+1) = 1 | X_m(t) = 1\}$ and $q_m :=$

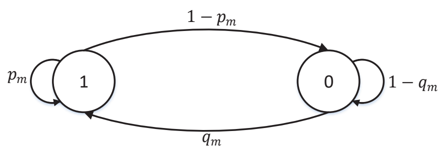

Fig. 1. A two-state Markov chain.

$\Pr\{X_m(t+1) = 1|X_m(t) = 0\}$. If the state of a fog node is 1, the task is processed with sufficient computational resources, which contributes to a low latency and a high reward. On the contrary, if the state of a fog node is 0, the deficiency of computational resources leads to a high latency and a low reward. Specifically, the immediate reward is defined as

$$r_m(X_m(t)) = \begin{cases} 1, & X_m(t) = 1, \\ \delta, & X_m(t) = 0, \end{cases} \tag{2}$$

where $0 \le \delta < 1$ indicates a low reward. Let $b_m(t)$ be a belief value, which represents the probability that fog node m is in state 1 in slot t. Thus the expected immediate reward of fog node m is given by

$$R_m(b_m(t)) = [b_m(t) + (1 - b_m(t))\delta]. \tag{3}$$

Let $a_m(t)$ be the action in task offloading, i.e. $a_m(t) = 1$ (active action) if fog node m is selected to offload a task and $a_m(t) = 0$ (passive action) otherwise. The belief states $b_m(t)$ evolve as

$$b_m(t+1) = \begin{cases} p_m, & \text{if } a_m(t) = 1, X_m(t) = 1, \\ q_m, & \text{if } a_m(t) = 1, X_m(t) = 0, \\ Q_m(b_m(t)), & \text{if } a_m(t) = 0, \end{cases} \tag{4}$$

where $Q_m(b_m(t)) := b_m(t)p_m + (1-b_m(t))q_m$. Consider that the evolution of the states of fog nodes has positive autocorrelation, i.e. $p_m > q_m$. Thus, we have $q_m \le b_m(t) \le p_m$.

At the beginning of each slot, the UE only maintains a belief value of each fog node instead of the exact knowledge of the current states. With knowledge of these belief values, the UE selects K out of M fog nodes to offload K tasks. Before the end of each slot, the selected fog nodes feed back the computation results and their current states. Meanwhile, the belief values are updated according to (4). We assume that the expected latencies due to data transmissions, e.g. transmitting tasks and results between the UE and fog nodes, are almost the same. In order to obtain computation results with low latency, the UE aims to find an optimal policy for task offloading which maximizes the infinite-horizon expected discounted sum of rewards. Specifically, we consider the following task offloading optimization problem:

$$\begin{aligned} \underset{\boldsymbol{\phi}\in\boldsymbol{\Phi}}{\text{maximize}} \quad & \mathsf{E}^{\boldsymbol{\phi}}_{\boldsymbol{b}(0)} \left[\sum_{t=0}^{\infty} \sum_{m=1}^{M} \beta^t R_m(b_m(t)) a_m(t) \right], \\ \text{subject to} \quad & \sum_{m=1}^{M} a_m(t) \le K, \end{aligned} \tag{5}$$

where $0 \le \beta < 1$ is the discounted factor, $\boldsymbol{b}(0)$ represents initial state vector of M fog nodes, $\boldsymbol{\Phi}$ is the set of admissible policies for task offloading, and $\mathsf{E}^{\boldsymbol{\phi}}_{\boldsymbol{b}(0)}$ denotes the expectation under a policy $\boldsymbol{\phi}$ conditional on the initial belief state vector being equal to $\boldsymbol{b}(0)$.

The task offloading problem in (5) is actually an RMAB problem where each fog node can be regarded as an arm and the state of each arm is the corresponding belief value of the fog node. To tackle the exploitation-exploration tradeoff in such problems, we make use of the well-established theory behind RMAB. In the next section, we will first overview the RMAB problem briefly and then solve the task offloading problem by exploiting the Whittle's index policy.

3. RMAB PROBLEM AND INDEX APPROACH

The conventional deterministic Markov multi-armed bandits (MAB) problems could be solved optimally by the Gittins index policy [15], which could significantly reduce the complexity of finding the optimal solution. As an extension of a conventional MAB problem, an RMAB problem allows multiple arms to be activated simultaneously and allows passive arms to change states [13]. In [16], the authors proved that a general RMAB was PSPACE-hard and its optimal solution was out of reach. To solve such an RMAB problem, by Lagrangian relaxation, Whittle heuristically proposed a priority index policy in [13], which was known as the Whittle's index policy and it was asymptotically optimal under certain conditions [17]. However, the existence of this index is not always guaranteed. In fact, to ensure the existence of a Whittle's index policy, the RMAB problem should be indexable and the indexability of an RMAB problem is unfortunately hard to establish [18]. Meanwhile, the computing of the Whittle's index can be complex and often relies on numerical approximations [18]. Addressing RMAB problems, in [12, 19], the author has introduced and explored the concept of the marginal productivity (MP) index, which is an extension of the Whittle's index. With an economic interpretation, the MP index policies aim to dynamically allocate resources to the arms that can make better uses of them [20]. Note that the MP index gives a general method for the calculation of the Whittle's index. In the sequel, we will solve the task offloading òptimization problem using the Lagrangian relaxation and the Whittle's MP index approach.

3.1. Lagrangian Relaxation

The Lagrangian relaxation technique relaxes the deterministic constraint in (5) to the following statistical one:

$$\mathsf{E}\left[\sum_{t=0}^{\infty}\sum_{m=1}^{M}\beta^t a_m(t)\right] \le \frac{K}{1-\beta}. \tag{6}$$

According to (5), the following Lagrangian can be formulated:

$$\mathcal{L}(\lambda, \{\phi_m\}_{m=1}^M) = \lambda\frac{K}{1-\beta} + \sum_{m=1}^{M}\mathcal{L}_m(\lambda, \phi_m), \tag{7}$$

where ϕ_m denotes the task offloading policy for fog node m and

$$\mathcal{L}_m(\lambda, \phi_m) = \mathsf{E}^{\phi_m}_{b_m(0)}\left[\sum_{t=0}^{\infty}\beta^t\left\{R_m(b_m(t))a_m(t) - \lambda a_m(t)\right\}\right].$$

As a result, we can approximate the constrained optimization problem in (5) by solving the following problem:

$$\underset{\lambda\ge 0}{\text{minimize}} \ \underset{\boldsymbol{\phi}\in\boldsymbol{\Phi}}{\text{maximize}} \ \mathcal{L}(\lambda, \{\phi_m\}_{m=1}^M), \tag{8}$$

where λ is the Lagrangian multiplier. Fixing λ, the relaxed optimization problem (8) can be decomposed into M separate single-UE subproblems, i.e.

$$\underset{\phi_m\in\Phi_m}{\text{maximize}} \ \mathsf{E}^{\phi_m}_{b_m(0)}\left[\sum_{t=0}^{\infty}\beta^t\left\{R_m(b_m(t))a_m(t) - \lambda a_m(t)\right\}\right], \tag{9}$$

where Φ_m is the set of admissible task offloading policies for fog node m. Note that λ can be interpreted as the cost for task offloading. As λ increases, the passive action should be more attractive. Let $\mathcal{D}_m(\lambda)$ be the set of belief states of fog node m in which it is optimal to stay passive. Let $\mathcal{S}_k$ denote the state space of fog node k. Then we can define the Whittle's indexability using the following definition.
Definition: A fog node is indexable if and only if, the set $\mathcal{D}_m(\lambda)$ monotonically expands from $\emptyset$ to $\mathcal{S}_k$ as λ increases from $-\infty$ to ∞. The RMAB problem is indexable if and only if all the fog nodes are indexable.

For each belief value, b_m, of fog node m, the Whittle's index, $\lambda(b_m)$, is defined as the infimum of the set of λ with which $b_m \in \mathcal{D}_k(\lambda)$. Specifically, the Whittle's index is given by $\lambda(b_m) = \inf\{\lambda | b_m \in \mathcal{D}_m(\lambda)\}$. If the RMAB problem is indexable, the Whittle's indices $\{\lambda_m(b_m)\}_m^M$ can be used as priority indices for solving (5). In [22], Niño-Mora established sufficient indexability conditions for continuous-state restless bandits based on partial conservation laws (PCLs) and provided a general framework for the Whittle's index calculation based on the MP index theory. In the following, we will address the indexability of the task offloading problem and calculate the Whittle's index by using the MP index theory.

3.2. Index Policy for Task Offloading

In this section, we focus on a generic single arm subproblem in (9), and thus drop the superscript m from the above notations. In order to simplify the index formulation, we let the immediate reward be $r'(X(t)) = [r(X(t)) - \delta]/(1-\delta)$. Thus, the expected immediate reward is $R(b(t)) = b(t)$. According to the MP index theory in [23], to judge the task offloading performance, for a given initial belief state $b := b(0)$, we define the work measure of a fog node as $g^{\phi}(b) := \mathsf{E}_b^{\phi}[\sum_{t=0}^{\infty} \beta^t a(t)]$, which gives the expected total discounted number of the offloaded tasks. Correspondingly, we define the reward measure as $f^{\phi}(b) := \mathsf{E}_b^{\phi}[\sum_{t=0}^{\infty} \beta^t R(b(t))a(t)]$, which gives the expected discounted reward.

It has been shown in [11] that there exists a threshold policy which is the optimal policy for the optimization problem in (9). Specifically, there exists a threshold $z \in (q,p)$, such that the optimal action $a(t) = 1$ if the current belief value $b > z$ and the optimal action $a(t) = 0$, otherwise. We define z as a threshold level for the threshold policy. Thus, if the belief state exceeds threshold level z, the fog node will be offloaded a task. We denote the z-threshold policy as $\phi(z)$. We can characterize the measures as follows:

$$g^{\phi(z)}(b) = \begin{cases} 1 + \beta b g^{\phi(z)}(p) + \beta(1-b)g^{\phi(z)}(q), & b \in (z,p], \\ \beta g^{\phi(z)}(q + (p-q)b), & b \in [q,z), \end{cases} \tag{10}$$

and

$$f^{\phi(z)}(b) = \begin{cases} b + \beta b f^{\phi(z)}(p) + \beta(1-b)f^{\phi(z)}(q), & b \in (z,p], \\ \beta f^{\phi(z)}(q + (p-q)b), & b \in [q,z). \end{cases} \tag{11}$$

For a given threshold z and action a, we let $\phi(1,z)$ denote the policy that takes action a in the initial slot and adopts the z-threshold policy thereafter. We further define the marginal work measure as

$$\begin{aligned} w^{\phi(z)}(b) &:= g^{\phi(1,z)}(b) - g^{\phi(1,z)}(b) \\ &= 1 + \beta b g^{\phi(z)}(p) + \beta(1-b)g^{\phi(z)}(q) - \beta g^{\phi(z)}(q + (p-q)b), \end{aligned}$$

and define the marginal reward measure as

$$\begin{aligned} r^{\phi(z)}(b) &:= f^{\phi(1,z)}(b) - f^{\phi(1,z)}(b) \\ &= b + \beta b f^{\phi(z)}(p) + \beta(1-b)f^{\phi(z)}(q) - \beta f^{\phi(z)}(q + (p-q)b). \end{aligned}$$

If $w^{\phi(z)}(b) \neq 0$, the marginal productivity rate is given by

$$\gamma^{\phi(z)}(b) := r^{\phi(z)}(b)/w^{\phi(z)}(b). \tag{12}$$

According to the sufficient conditions for PCL-indexable in [22], the fog node is PCL-indexable if $w^{\phi(z)}(b) > 0$ and the index:

$$\lambda^*(b) := \gamma^{\phi(b)}(b), \tag{13}$$

is monotonically nondecreasing in b. Furthermore, according to Theorem 1 in [22], if a fog node is PCL-indexable, it is Whittle's indexable and $\lambda^*(b)$ is its Whittle's index, which is also called the Whittle's MP index.

In order to verify the conditions for PCL-indexable and compute the Whittle's MP index, in the sequel, we address the calculation of the Whittle's MP index. For example, in order to calculate marginal work measure, i.e. $w^{\phi(z)}(b)$, we need to evaluate $g^{\phi(z)}(p)$, $g^{\phi(z)}(q)$ and $g^{\phi(z)}(bp + (1-b)q)$. Note that, for any $z \in (q,p)$, we have

$$\begin{aligned} g^{\phi(z)}(p) &= 1 + \beta p g^{\phi(z)}(p) + \beta(1-p)g^{\phi(z)}(q), \\ g^{\phi(z)}(q) &= \beta g^{\phi(z)}(q'), \end{aligned} \tag{14}$$

$$\begin{aligned} f^{\phi(z)}(p) &= p + \beta p f^{\phi(z)}(p) + \beta(1-p)f^{\phi(z)}(q), \\ f^{\phi(z)}(q) &= \beta f^{\phi(z)}(q'), \end{aligned} \tag{15}$$

where $q' := q + (p-q)q$. According to (14) and (15), we obtain

$$\begin{aligned} g^{\phi(z)}(p) &= \frac{1}{1-\beta p}[1 + \beta^2(1-p)g^{\phi(z)}(q')], \\ f^{\phi(z)}(p) &= \frac{1}{1-\beta p}[p + \beta^2(1-p)f^{\phi(z)}(q')]. \end{aligned} \tag{16}$$

According to (14), (15) and (16), we can further calculate the index value of $\gamma^{\phi(b)}(b)$ in different cases via (12) and (13). By calculating the index value, it can be verified that the subproblem is PCL-indexable, and $\lambda^*(b)$ is the Whittle MP index. In summary, we can establish the following proposition.

Proposition 1. *For a belief value $b \in [q,p]$, the Whittle's MP index $\lambda^*(b)$ for the task offloading optimization problem in (9) can be obtained as follows.*

For case 1, $q \leq b < q + q(p-q)$, the Whittle's MP index is given by

$$\lambda^*(b) = \frac{b + \beta(b-q)}{1 + \beta(b-q)}; \tag{17}$$

For case 2, $q/(1-(p-q)) \leq b \leq p$, the Whittle's MP index is given by

$$\lambda^*(b) = \frac{b}{1 + \beta(b-p)}; \tag{18}$$

For case 3, $q + q(p-q) \leq b < q/(1-(p-q))$, the Whittle's MP index is given by

$$\lambda^*(b) = \frac{b + \beta b f^{\phi(b)}(p) + \beta(1-b)f^{\phi(b)}(q) - \beta f^{\phi(b)}(q + (p-q)b)}{1 + \beta b g^{\phi(b)}(p) + \beta(1-b)g^{\phi(b)}(q) - \beta g^{\phi(b)}(q + (p-q)b)}. \tag{19}$$

Proof. The calculation of the MP index value is similar to [23]. However, we consider that $q \leq b \leq p$. Specifically, we calculate the MP index value in three different cases. For case 1, i.e. $q \leq b < q + q(p-q)$, and case 2, i.e. $q/(1-(p-q)) \leq b \leq p$, we can obtain the simplified index value easily and the details are omitted here. For case 3, i.e. $q+q(p-q) \leq z < q/(1-(p-q))$, we cannot explicitly obtain $g^z(q')$ and $f^z(q')$. Let $h_0(b) = b$ and $h_1(b) = q + (p-q)b$. Considering the iteration $h_n(b) = h_n(h_{n-1}(b))$, for $n \geq 1$, we have $h_n(b) = \frac{1-(p-q)^n}{1-(p-q)} q + (p-q)^n b$. Note that $\lim_{n\to\infty} h_n(b) = q/(1-(p-q))$. Meanwhile, we note that $h_n(b)$ is increasing in n when $b < q/(1-(p-q))$. Thus, for any $b \leq z$, there exists a $n := n_{b,z}$ such that $h_{n-1}(b) \leq z$ and $h_n(b) > z$. Note that $q' = q + q(p-q) \leq z$, thus we have

$$g^{\phi(z)}(q') = \beta^{n^*} g^{\phi(z)}(h_{n^*}(q')), \;\; f^{\phi(z)}(q') = \beta^{n^*} f^{\phi(z)}(h_{n^*}(q')). \tag{20}$$

where $n^* = n_{q',z}$, $g^{\phi(z)}(h_{n^*}(q'))$, $f^{\phi(z)}(h_{n^*}(q'))$ are given as

$$g^{\phi(z)}(h_{n^*}(q')) = 1 + \beta h_{n^*}(q') g^{\phi(z)}(p) + \beta(1-h_{n^*}(q')) g^{\phi(z)}(q),$$
$$f^{\phi(z)}(h_{n^*}(q')) = h_{n^*}(q') + \beta h_{n^*}(q') f^{\phi(z)}(p) + \beta(1-h_{n^*}(q')) f^{\phi(z)}(q).$$

From (16) and (20), we can solve $g^{\phi(z)}(q')$ and $f^{\phi(z)}(q')$ and obtain $g^{\phi(z)}(p)$, $g^{\phi(z)}(q)$, $f^{\phi(z)}(p)$ and $f^{\phi(z)}(q)$. For any $b \in [q, z]$, the marginal work measure is given by

$$w^{\phi(z)}(b) = 1 + \beta b g^{\phi(z)}(p) + \beta(1-b) g^{\phi(z)}(q) - \beta^{n^*} g^{\phi(z)}(h_{n^*}(b)), \tag{21}$$

where $h_{n^*}(b) > z$ and $n^* = n_{b,z}$. For $b \in (z, p]$, we have

$$w^{\phi(z)}(b) = 1 + \beta b g^{\phi(z)}(p) + \beta(1-b) g^{\phi(z)}(q) - \beta g^{\phi(z)}(b), \tag{22}$$

where $g^{\phi(z)}(b)$ is given by (10) with $b > z$. Note that $w^{\phi(z)}(b) > 0$. Similarly, for $b \in [q, z]$, the marginal reward measure is given by

$$r^{\phi(z)}(b) = b + \beta b f^{\phi(z)}(p) + \beta(1-b) f^{\phi(z)}(q) - \beta^{n^*} f^{\phi(z)}(h_{n^*}(b)), \tag{23}$$

where $h_{n^*}(b) > z$ and $n^* = n_{b,z}$. For $b \in (z, p]$, we have

$$r^{\phi(z)}(b) = b + \beta b f^{\phi(z)}(p) + \beta(1-b) f^{\phi(z)}(q) - \beta f^{\phi(z)}(b). \tag{24}$$

The index value in (19) can be obtained from (21), (22), (23) and (24). The derived closed-form index value expressions can be used to verify the PCL indexability conditions directly. Thus, we can claim that the index values are the Whittle's MP index. □

According to Proposition 1, we can calculate the values of the Whittle's MP indices $\{\lambda_m^*(b_m(t))\}_{m=1}^M$ of all fog nodes in each slot using the closed-form index expressions with remarkably low complexity. The task offloading decisions are further made according to the index values. In summery, we can establish the following task offloading strategy based on the Whittle's MP index policy: *At the beginning of time slot t, with knowledge of belief values $\{b_m(t)\}_{m=1}^M$, the K fog nodes with the highest Whittle's MP indices $\lambda_m^*(b_m(t))$ are selected from the M fog nodes for task offloading.*

Note that since the Whittle's MP index value is monotonically increases with belief value $b_m(t)$, when the fog nodes have the same Markov structure, i.e. the Markov probability transition matrices are the same, and vary independently across fog nodes, the Whittle's index policy essentially becomes the myopic policy, which aims to maximize the expected immediate reward. For myopic policy, the priority index value for fog node m is given by $\lambda_m^{\text{myopic}} = R_m(b_m(t))$ in time slot t. The myopic policy is thus *offloading tasks to the K fog nodes with the K largest myopic indices λ_m^{myopic}.*

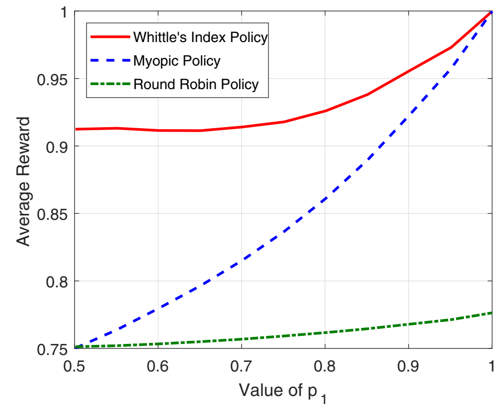

Fig. 2. Performance comparison of different task offloading policies. ("Whittle's Index Policy": the index policy proposed in Section 3.2; "Myopic Policy": the myopic policy that aims to maximize the expected immediate reward; "Round Robin Policy": the round robin policy that allocates tasks in a circular order.)

4. NUMERICAL RESULTS

In this section, we compare our proposed task offloading strategy based on the Whittle's MP index policy with the myopic policy and the round robin policy by numerical results. We consider a fog computing system with $M = 10$ fog nodes and a UE offloads $K = 1$ task in each time slot to a fog node. The states of fog nodes are modeled as a two-state Markov chain. The low reward value is set to $\delta = 0.5$ and the time horizon is set to $T = 10^3$.

In Fig. 2, with different transition probabilities, we compare the average rewards of different index policies by setting $\beta = 0.999$. We let $q_1 = 0.5, p_2 = 0.99, q_2 = 0.01$ and let $p_m = 0.8, q_m = 0.2$ for $m \neq 1, 2$. The value of p_1 varies in the interval $[0.5, 1]$. To highlight the superiority of our method, we set $q_1 = 0.5, p_2 = 0.99, q_2 = 0.01$ and let p_1 varies in the interval $[0.5, 1]$. The transition probabilities of the other UEs are set as $p_m = 0.8, q_m = 0.2, m = 3, 4, \cdots, 10$. Note that node-2 has a large autocorrelation, while fog node-1 has relative small autocorrelation. The results show that the proposed task offloading strategy based on the Whittle's index policy outperforms the other two conventional index policies and the corresponding performance gain increases as p_1 approaches 0.5. We can see the Whittle's index policy can achieve better performance when the differences between fog nodes are larger.

5. CONCLUSIONS

In this paper, we have addressed the task offloading problem in a fog-enabled network and have proposed an RMAB framework to offload multiple tasks from one UE to multiple neighbor fog nodes. In particular, by exploiting the temporal correlation of fog node states, we have analyzed the indexability of the task offloading problem. Meanwhile, we have obtained the corresponding closed-form Whittle's index expressions. Furthermore, an index policy based on Whittle's index has been proposed to solve the optimal task offloading problem in a fog-enabled network. Numerical results demonstrate the superiority of the Whittle's index policy for task offloading problem.

6. REFERENCES

[1] M. Chiang and T. Zhang, “Fog and IoT: An overview of research opportunities,” *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

[2] T. Q. Dinh, J. Tang, Q. D. La, and T. Q. S. Quek, “Offloading in mobile edge computing: Task allocation and computational frequency scaling,” *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3571–3584, Aug. 2017.

[3] C. You, K. Huang, H. Chae, and B.-H. Kim, “Energy-efficient resource allocation for mobile-edge computation offloading,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1397–1411, Mar. 2017.

[4] Y. Yang, K. L. Wang, G. W. Zhang, X. Chen, X. Luo, and M. T. Zhou, “MEETS: Maximal energy efficient task scheduling in homogeneous fog networks,” *IEEE Internet Things J.*, in print, Jun. 2018.

[5] J. Kwak, Y. Kim, J. Lee, and S. Chong, “DREAM: Dynamic resource and task allocation for energy minimization in mobile cloud systems,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2510–2523, Dec. 2015.

[6] Y. Mao, J. Zhang, S. H. Song, and K. B. Letaief, “Stochastic joint radio and computational resource management for multi-UE mobile-edge computing systems,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5994–6009, Sept. 2017.

[7] S. Zhao, Y. Yang, Z. Shao, X. Yang, H. Qian, and C. X. Wang, “FEMOS: Fog-enabled multi-tier operations scheduling in dynamic wireless networks,” *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1169–1183, Apr. 2018.

[8] Y. Yang, S. Zhao, W. Zhang, Y. Chen, X. Luo, and J. Wang, “DEBTS: Delay energy balanced task scheduling in homogeneous fog networks,” *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2094–2106, Jun. 2018.

[9] L. Pu, X. Chen, J. Xu, and X. Fu, “D2D fogging: An energy-efficient and incentive-aware task offloading framework via network-assisted D2D collaboration,” *IEEE J. Sel. Areas Commun.*, vol. 34, no.12, pp. 3887–3901, Dec. 2016.

[10] Z. Zhu, T. Liu, S. Jin, and X. Luo, “Learn and pick right nodes to offload”, *arXiv preprint arXiv:1804.08416*, 2018.

[11] W. Ouyang, S. Murugesan, A. Eryilmaz, and N. B. Shroff, “Exploiting channel memory for joint estimation and scheduling in downlink networks—a Whittles indexability analysis,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1702–1719, Apr. 2015.

[12] J. Niño-Mora, “Restless bandits, partial conservation laws and indexability,” *Adv. Appl. Probab.*, vol. 33, no. 1, pp. 76–98, Mar. 2001.

[13] P. Whittle,“Restless bandits: Activity allocation in a changing world,” *A Celebration of Applied Probability, Ed. J. Gani, J. Appl. Probab.*, vol. 25A, pp. 287–298, 1988.

[14] P. Blasco and D. Gündüz, “Multi-access communications with energy harvesting: A multi-armed bandit model and the optimality of the myopic policy,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 585–597, Mar. 2015.

[15] J. C. Gittins, “Bandit processes and dynamic allocation indices,” *J. Roy. Statist. Soc. Ser. B*, vol. 41, no. 2, pp. 148–149, 1979.

[16] C. H. Papadimitriou and J. N. Tsitsiklis, “The complexity of optimal queueing network control,” in *Proc. IEEE Struct. Complexity Theory Conf.*, Amsterdam, the Netherlands, Jun./Jul. 1994, pp. 318–322.

[17] R.R. Weber and G. Weiss, “On an index policy for restless bandits,” *J. Appl. Probab.*, vol. 27, no. 3, pp. 637–648, Sept. 1990.

[18] K. Liu and Q. Zhao, “Indexability of restless bandit problems and optimality of Whttile index for dynamic multichannel access,“ *IEEE Trans. Inf. Theory*, vol.56, no. 11, pp. 5547–5567, Nov. 2010.

[19] J. Niño-Mora, “Dynamic priority allocation via restless bandit marginal productivity indices,” *TOP*, vol. 15, no. 2, pp. 161–198, Dec. 2007.

[20] P. Whittle, “Comments on: Dynamic priority allocation via restless bandit marginal productivity indices,” *TOP*, vol. 15, no. 2, pp. 217–219, Dec. 2007.

[21] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed Bandit Allocation Indices.* 2th ed. Wiley, 2011.

[22] J. Niño-Mora and S. S. Villar, “Sensor scheduling for hunting elusive hiding targets via Whittle’s restless bandit index policy,” in *Proc. IEEE Netw. Games, Control Optim.*, Paris, France, Oct. 2011, pp. 1–8.

[23] J. Niño-Mora, “An index policy for dynamic fading-channel allocation to heterogeneous mobile UEs with partial observations,” in *Proc. IEEE Next Generation Internet Networks*, Krakow, Poland, Apr. 2008, pp. 231–238.