Multi-Interface Channel Allocation in Fog Computing Systems using Thompson Sampling

Junge Zhu, Xi Huang, Xin Gao, Ziyu Shao, Yang Yang
School of Information Science and Technology, ShanghaiTech University
Email: {zhujg, huangxi, gaoxin, shaozy, yangyang}@shanghaitech.edu.cn

***Abstract*—In fog computing systems, fog nodes often maintain multiple interfaces to achieve simultaneous transmission to end devices. To maximize the utilization of network capacities and avoid interference, a critical mission for each fog node is to allocate distinct channels to its interfaces, *a.k.a. multi-interface channel allocation*, with the aim to maximize the total throughput induced by successful transmissions over finite time horizon. The lack of *priori* knowledge about channel state dynamics and the complex interplay among different interfaces' channel allocation make the allocation scheme design extremely challenging. To this end, channel allocation must be conducted in a dynamic fashion while integrated with an online learning procedure that proactively learns the statistics through transmission feedback with a well-balanced exploration-exploitation trade-off, while minimizing the throughput loss due to decision making under uncertainties, *a.k.a.*, the *regret*. This paper addresses the above challenges. By formulating the channel allocation problem in multi-armed bandit setting with multiple plays and leveraging Thompson sampling techniques, we propose a Multi-Interface Channel Allocation (MICA) scheme, which makes online channel allocation decisions by effective learning from feedback. Our theoretical analysis shows that MICA achieves a mild-value of regret upper bounded by $O(\log T)$ over a finite time horizon T. Further, we propose MICA-M which extends MICA to handle more general multi-level transmission feedback. Our simulation results verify the effectiveness and robustness of both MICA and MICA-M in achieving a notable reduction in regret compared to other baseline schemes.**

I. Introduction

In fog computing systems, heterogeneous end devices enjoy various on-demand services from their nearby processing nodes with enhanced computing and storage capacities, *a.k.a. fog nodes*, via wireless communication. Each fog node serves a fixed set of end devices within a particular time period; to achieve simultaneous transmissions to such end devices, a fog node usually maintains multiple interfaces [1]. To avoid communication interference among different interfaces, a fog node should allocate each interface with a distinct wireless channel from a fixed set of wireless channels [2], *a.k.a. multi-interface channel allocation*, wherein each channel maintains a constant transmission rate [3] and stationary channel dynamics. To make best use of wireless channel capacities, the allocation scheme design often aims at maximizing the total throughput induced by successful transmissions within the period.

To date, channel allocation scheme design still remains an unsolved and challenging problem in fog computing systems. Basically, such systems often manifest various uncertainties in practice; *e.g.*, channel state dynamics such as successful transmission probabilities (indicated by binary signals) are usually unattainable and can only be inferred by the feedback after each transmission. Without knowing such information *a priori*, channel allocation procedure must be conducted in a *dynamic* fashion and integrated with an effective online *learning* procedure that learns the unknown statistics from collected feedback over time, so that the throughput loss due to sub-optimal channel allocations, *a.k.a.*, the *regret*, can be minimized. To this end, two challenges need be to resolved. *One challenge* lies in the design of dynamic channel allocation must carefully handle the *exploration-exploitation* trade-off in the decision making process. On one hand, the allocation scheme, if overly relying on temporally learned estimates, *i.e.*, over-exploitation, may induce constantly enhanced bias towards some sub-optimal allocation decisions and result in increasing throughput loss; on the other hand, if exploring among different allocations too frequently, it can lead to even more throughput loss. *The other challenge* regards the fact that each time the designed scheme should allocate channels amongst multiple interfaces. The innate couplings among interfaces' allocation decisions results in a non-trivial interplay between the learning of different channels' statistics, making the design even more challenging.

Existing works have studied related problems and developed solutions in various wireless scenarios [4]–[9]; they basically focused on rate selection problem between a pair of transmitter and receiver over a single wireless channel. Essentially, the rate selection problem can be equivalent to a special case of multi-interface channel allocation problem, where there is only one interface and each distinct rate corresponds to one channel. Under such an equivalence, some of them assume the attainability of different channels' statistics such as the successful transmission probabilities [4]–[6]; however, as aforementioned, there is no empirical evidence to show such *priori* knowledge is attainable in fog computing systems. The other line of works considered the scenarios in which such channel statistics are unavailable with binary feedback signals to indicate the successfulness of the transmission [7]–[9], formulated the problem under multi-armed bandits settings, and proposed various learning-aided rate selection schemes. Existing schemes, through proven effective, are designed for channel allocation for only one interface, and incompetent to handle the more complicated case with non-trivial interplay among multiple interfaces' allocation.

This paper addresses the above challenges. In particular, we consider the problem of multi-interface channel selection in fog computing systems with unknown wireless channel statistics. Our key results and contributions are summarized as follows.

- ◇ **Problem Formulation:** We formulate the problem as a multi-armed bandit problem with multiple play, which aims to maximize the expectation of total throughput incurred by successful transmission within a finite period of time. To our best knowledge, this paper is the first to study the multi-interface channel allocation problem and propose effective solutions in fog computing systems.
- ◇ **Algorithm Design:** Inspired by recent work [10], we employ multi-play Thompson sampling techniques and propose a Multi-Interface Channel Allocation (MICA) scheme, which makes online and effective channel allocation decisions by achieving a balance between exploration and exploitation and proactive learning from collected feedback. Considering that in practice, channel states are characterized by multi-level (more than binary) feedback [11], we further propose MICA-M which extends MICA to handling such multi-level feedback.
- ◇ **Performance Analysis:** Our theoretical analysis shows that MICA is able to induce expected regret (throughput loss) that is limited within an $O(\log T)$ upper bound, which suggests that MICA's effectiveness in achieving high throughput in face of unknown system dynamics.
- ◇ **Numerical simulations:** We conduct extensive simulations to evaluate the performance of MICA and MICA-M, in which the results show the effectiveness and robustness of MICA compared to baseline schemes by achieving high throughput with a notable regret reduction under various settings.

The rest of the paper is organized as follows. Section II presents the system model and problem formulation. Section III elaborates the algorithm design and its extension, followed by its performance analysis. Section IV shows the simulation results and Section V concludes the paper.

II. System Model and Problem Formulation

In this section, we first introduce our system model, then formulate the channel allocation problem under multiple-play multi-armed bandit (MP-MAB) settings. We summarize the key notations in Table I.

A. System Model

We consider a fog node equipped with M interfaces to communicate with its associated end devices over a period of T time slots. Basically, the fog node maintains a fixed set of N independent channels, denoted by set $\mathcal{C} \triangleq \{1, \dots, N\}$. The transmission rate remains constant over the period, denoted by r_i for each channel $i \in \mathcal{C}$. Accordingly, the transmission on channel i succeeds with a probability of p_i. We define $\mathcal{R} \triangleq \{r_1, \dots, r_N\}$ as the set of transmission rates of all channels.

B. Problem Formulation

At each time slot t, the fog node allocates M channels to interfaces for simultaneous transmissions, denoted by I_t. The general goal of channel allocation is to maximize the total expected throughput induced by successful transmissions, *i.e.* the summation of all chosen channels' expected rate $r_i p_i$. In other words, for the fog node, it needs to solve the following optimization problem,

$$\underset{I \in \mathcal{I}}{\text{Maximize}} \quad \sum_{i \in I} r_i p_i, \tag{1}$$

where $\mathcal{I}$ denotes the set of all feasible channel allocations, *i.e.* $\mathcal{I} \triangleq \{J \subset \mathcal{C} : |J| = M\}$.

TABLE I: Key Notations

Notation	Description
M	Number of wireless interfaces of the fog node
N	Number of available channels
$\mathcal{C}$	Available channel set, $\mathcal{C} = \{1, \dots, N\}$
$\mathcal{I}$	Set of available channel subsets of size M
r_i	Transmission rate of channel i
p_i	Probability of successful transmission at channel i
I_t	Set of selected channels at time slot t
$X_i(t)$	Transmission feedback of channel i at time slot t

Note that, if the successful probabilities p_i of all channels are known *a priori*, Problem (1) can be directly solved by allocating the top M channels with the largest expected rate to its interfaces, denoted by set I^*. However, the successful transmission probability p_i is usually unattainable in practice, making the problem extremely challenging to solve. To cope with such challenge, the channel allocation procedure must be conducted in a *dynamic* fashion and integrated with online *learning* to learn the unknown successful transmission probabilities based on feedbacks from different allocations. Such a dynamic decision making procedure must strike a balance between *exploration* and *exploitation*, since over-exploration, *i.e.*, the fog node proactively switches among different allocations, can constantly incur inferior throughput; on the other hand, over-exploitation, *i.e.*, the fog node overly resorts to empirical estimates based on successful transmission, may miss other allocations with a potentially better throughput.

Inspired by recent work [10], we formulate the problem of multi-interface channel allocation under the settings of multiple-play multi-armed bandits (MP-MAB) model. A typical MAB problem considers the situation in which an agent makes a series of interactions with its residing environment. In each round, the agent picks one action (arm) from a given set of actions (arm set); according to the chosen action, the environment reveals a reward to the agent which is sampled from an unknown distribution. Through rounds of interaction, the agent aims to find an optimal policy to maximize the expected cumulative reward. MP-MAB extends MAB by enabling the agent to play multiple arms at each round. In the MP-MAB model, the reward of each selected arm is sampled from their distributions independently.

To recast Problem (1) to the MP-MAB settings, we view the fog node as the agent and the available channel set $\mathcal{C}$ as the arm set. At time slot t, the agent selects a subset of M arms (channels), denoted by $I_t \in \mathcal{I}$. Then, the fog node carries out the transmission over the chosen channels to end devices. We denote the observed rewards by $X_i(t), \forall i \in I_t$, in which $X_i(t) = 1$ if the transmission succeeds and zero otherwise. Equivalent to Problem (1), the fog node aims to solve the following optimization problem over a finite time horizon of T time slots,

$$\underset{[I_t^\pi]_t}{\text{Maximize}} \quad \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in I_t^\pi} r_i X_i(t)\right]. \tag{2}$$

We define the regret to measure the performance of channel allocation.

$$R_\pi(T) \triangleq \sum_{t=1}^{T}\sum_{j\in I^*} r_j p_j - \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in I_t^\pi} r_i X_i(t)\right], \tag{3}$$

where I^* denotes the optimal channel allocation, *i.e.*, the solution to Problem (1). Note that the regret is defined as the gap between policy π's expected cumulative reward and the optimal cumulative reward, so minimizing the regret is equivalent to maximizing the expected throughput induced by successful transmissions.

III. Algorithm Design & Performance Analysis

In this section, we elaborate the design of the channel allocation scheme, MICA (Channel Allocation using Multiple-Play Thompson Sampling), followed by a theoretical guarantee on the performance of MICA.

A. Algorithm Design

In section II, we recast the channel allocation problem into a multiple-play multi-armed bandit (MP-MAB) problem which aims at maximizing the expected weighted reward, as shown in (2). Inspired by the work of Harsh *et al.* [9], we design a modified Thompson Sampling algorithm to deal with the MP-MAB problem.

The feedback of the transmission over each channel is a binary signal, which follows a Bernoulli distribution with an unknown mean, *i.e.*, $X_i(t) \sim \text{Bern}(p_i)$, $\forall i \in I_t$. Based on Statistics, a common prior distribution is Beta distribution [12]. Over each channel i, we assume the successful transmission probability p_i to be a random variable that follows a Beta distribution with parameters a_i and b_i, *i.e.*, $p_i \sim \text{Beta}(a_i, b_i)$. Then according to Beta-Binomial conjugacy, the posterior distribution of p_i follows a beta distribution with parameters $a_i + S_i(t)$ and $b_i + F_i(t)$, *i.e.*, $\text{Beta}(a_i + S_i(t), b_i + F_i(t))$, where $S_i(t)$ and $F_i(t)$ are the numbers of successful and failed transmissions on channel i during the first t time slots, respectively.

By employing Thompson sampling techniques with Beta-Binomial conjugacy, we propose a Multi-Interface Channel Allocation (MICA) scheme. We show its pseudocode in Algorithm 1. Specifically, the prior distribution of $p_i, \forall i \in \mathcal{C}$ are set to $\text{Beta}(1, 1)$, *i.e.* $a_i = b_i = 1$, and parameters a_i and b_i are updated after each sampling of channel i. Setting

Algorithm 1 Multi-Interface Channel Allocation (MICA)

Input: Transmission rate set $\mathcal{R} = \{r_1, \ldots, r_N\}$
    Beta parameter $(a_i, b_i) \leftarrow (1, 1), \forall i \in \mathcal{C}$
Output: Sequence of channel allocations $\{I_t\}_{t=1}^T$

1: **for** t in $\{1, ..., T\}$ **do**
2:     **for** i in $\{1, ..., N\}$ **do**
3:         Draw $\theta_i(t) \sim \text{Beta}(a_i, b_i)$
4:     **end for**
5:     $I_t \leftarrow$ first M channels with the largest $r_i * \theta_i(t)$
6:     Observe feedback $X_i(t)$ for each i in I_t
7:     **for** $i \in I_t$ **do**
8:         **if** successful, *i.e.*, $X_i(t) = 1$ **then**
9:             $a_i \leftarrow a_i + 1$
10:         **else**
11:             $b_i \leftarrow b_i + 1$
12:         **end if**
13:     **end for**
14: **end for**

the prior distribution to $\text{Beta}(1, 1)$ means that p_i is uniformly distributed at $(0, 1)$, which is reasonable since the fog node doesn't have any prior knowledge of p_i. At each time slot t, the fog node first draws $\theta_i(t)$ from distribution $\text{Beta}(a_i, b_i)$ for all channels in line 2-4. Then in line 5, the fog node selects the set of M channels (I_t) with the largest estimated throughput. After all transmissions over the selected channels, the fog node receives the feedback signal $X_i(t)$ for each $i \in I_t$ in line 6. At the end of time slot t, the posterior distribution of the successful transmission probability p_i is renewed by updating a_i and b_i based on the feedback signal $X_i(t)$, as shown in line 7-13.

In addition, we note that MICA can be run in a computationally-efficient manner in each time slot t. Particularly in Algorithm 1, the for-loop (lines 2-4) requires N iteration and the for-loop (lines 7-13) requires M ($M < N$) iteration. Each iteration requires a constant complexity. When the fog node selects the set of M channels with the largest estimated throughput (line 5), radix sort can be run with a $O(N)$ computational complexity [13]. Hence, the computational complexity for MICA is $O(N)$, linear in the number of channels, implying MICA can make channel allocation decisions with low overheads.

Remark: MICA effectively strikes a balance between exploration and exploitation by effectively using the feedback information acquired after transmissions. Particularly, without any priori knowledge, by initializing the prior distribution of each p_i as $\text{Beta}(1, 1)$, *i.e.*, $\text{Uniform}(0, 1)$, MICA takes proactive exploration at the beginning. MICA extracts the empirical information from successively collected feedbacks and exploits them to update posterior distributions of different p_i. As a result, the decision making would incline to choosing the channels with empirically higher successful transmission probabilities (exploitation), while still accommodating exploration since the estimate $\theta_i(t)$ for each p_i is based on random sampling rather than a deterministic manner.

B. Performance Analysis

For the multi-interface channel allocation problem with a homogeneous transmission rate r, we show that the regret induced by MICA over finite time horizon is guaranteed within an $O(\log T)$ upper bound, which is specified by the following theorem. [1]

Theorem 1. *Over T time slots, the expected regret incurred by MICA is upper bounded by*

$$\mathbb{E}(R_{\widetilde{\pi}}(T)) \leq r \sum_{i \in \mathcal{C} \setminus I^*} \frac{\Delta_{i,m} \log T}{D\left(p_i, p_m\right)} + O((\log T)^{2/3}), \tag{4}$$

where 1) r is the homogeneous transmission rate; 2) $m \triangleq \arg\min_{j \in I^*} p_j$, i.e., the channel whose p_j is the minimum in the optimal channel set I^*; 3) $\Delta_{i,m} = p_m - p_i$; and 4) $D\left(x, y\right)$ is the Kullback-Leibler divergence [14] *between two Bernoulli distributions Bern(x) and Bern(y), i.e., $D\left(x, y\right) \triangleq x \log \frac{x}{y} + (1-x) \log \frac{1-x}{1-y}$.*

The proof is delegated to Appendix-A.

C. Extension of MICA

Under previous settings, the fog node is assumed to receive a binary feedback after each transmission, which indicates whether the transmission was successful or not. However, in practice, the transmission feedback may be quantified at a finer granularity, *a.k.a.* multi-level feedback [11]. For example, the feedback can be quantified as the observed channel quality after the transmission, which may be ranked by three levels, *good*, *medium*, *bad*. Such multi-level feedback setting is more general but also more complicated to cope with. In this section, we propose MICA-M, an extended version of MICA that deals with Multi-level feedback.

In the channel allocation system with L-level feedback, each channel i is associated with a probability vector $\boldsymbol{p}_i = (p_l^i)_{l=1}^L$ with $\sum_{l=1}^L p_l^i = 1$, which is the distribution of channel i's quality state. Specifically, channel i's quality is in level l with probability p_l^i.

We use an L-dimensional vector $\boldsymbol{Q}_i(t) = (n_l^i(t))_{l=1}^L$ to represent the quality state of channel i in time slot t. When the channel quality is in level l, the l-th element of $\boldsymbol{Q}_i(t)$ is 1 while other elements are 0, *i.e.*, $n_l^i(t) = 1$ and $n_{l'}^i(t) = 0 \ \forall l' \neq l$. Then $\boldsymbol{Q}_i(t)$ has a Multinomial distribution with parameters $n = 1$ and $\boldsymbol{p}_i$, *i.e.*, $\boldsymbol{Q}_i(t) \sim \text{Multinomial}(1, \boldsymbol{p}_i)$.

For each channel, when its quality is in level l, the transmission succeeds with probability α_l (α_l is a known constant can be obtained by empirical observations). Under this settings, the channel allocation problem becomes

$$\underset{[I_t^\pi]_t}{\text{Maximize}} \ \ \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i \in I_t^\pi} r_i \left(\sum_{l=1}^{L} n_l^i\left(t\right) \alpha_l\right)\right], \tag{5}$$

where $\sum_{l=1}^L n_l^i(t) \alpha_l$ is the probability of successful transmission on channel i at time slot t.

We assume the unknown probability vector $\boldsymbol{p}_i$ to be a Dirichlet distribution, *i.e.*, $\boldsymbol{p}_i \sim \text{Direchlet}(\boldsymbol{a}_i)$ and $\boldsymbol{a}_i = (a_l^i)_{l=1}^L$ is an L dimensional vector. According to Dirichlet-Multinomial conjugacy, the distribution of $\boldsymbol{p}_i$ can be updated based on observed quality state $\boldsymbol{Q}_i(t)$ of channel i at time slot t as follows:

$$\text{Dirichlet}(\boldsymbol{a}_i) \leftarrow \text{Dirichlet}(\boldsymbol{a}_i + \boldsymbol{Q}_i(t)). \tag{6}$$

By Dirichlet-Multinomial conjugacy [12] , we develop an extended version of MICA, Multi-Interface Channel Allocation with Multi-level feedback (MICA-M), to solve the channel allocation problem with multi-level feedback. The pseudocode of MICA-M is shown in Algorithm 2.

Algorithm 2 Multi-Interface Channel Allocation with Multi-level feedback (MICA-M)

Input: Transmission rate set $\mathcal{R} = \{r_1, \ldots, r_N\}$
  Dirichlet parameter $\boldsymbol{a}_i \leftarrow \mathbf{1}, \forall i \in \mathcal{C}$
Output: Action sequence $\{I_t\}_{t=1}^T$.

```
for t in {1, ..., T} do
    for i in {1, ..., N} do
        Draw θ_i(t) ~ Dirichlet(a_i), θ_i(t) = (θ_l^i(t))_{l=1}^L
    end for
    I_t ← top M channels with the largest r_i Σ_{l=1}^L θ_l^i(t) α_l
    Observe the channel state Q_i(t) for i in I_t
    for i ∈ I_t do
        a_i ← a_i + Q_i(t)
    end for
end for
```

IV. Simulation

A. Baselines

In this section, we conduct simulations to verify the effectiveness of MICA and MICA-M by comparing it with three state-of-the-art multiple-play multi-armed bandit algorithms: CUCB [15], MP-KL-UCB [16] and Bayes-UCB [17].

MP-KL-UCB: At time slot t, the fog node computes $q_i(t), \forall i \in \mathcal{C}$ as follows

$$q_i(t) = \max\Big\{q \in [0, r_i] : N_i(t) \cdot D\left(\tfrac{\hat{\mu}_i(t)}{r_i}, \tfrac{q}{r_i}\right) \leq \log(t) + c \log\log(t)\Big\}, \tag{7}$$

where $N_i(t)$ denotes the number of times the channel i has been selected by time slot t, $\hat{\mu}_i(t)$ denotes the empirical average throughput at channel i. $D(x, y)$ is the KL divergence between two Bernoulli distributions parametrized by x and y. After obtaining all values, MP-KL-UCB selects the top M channels ranked by $q_i(t)$.

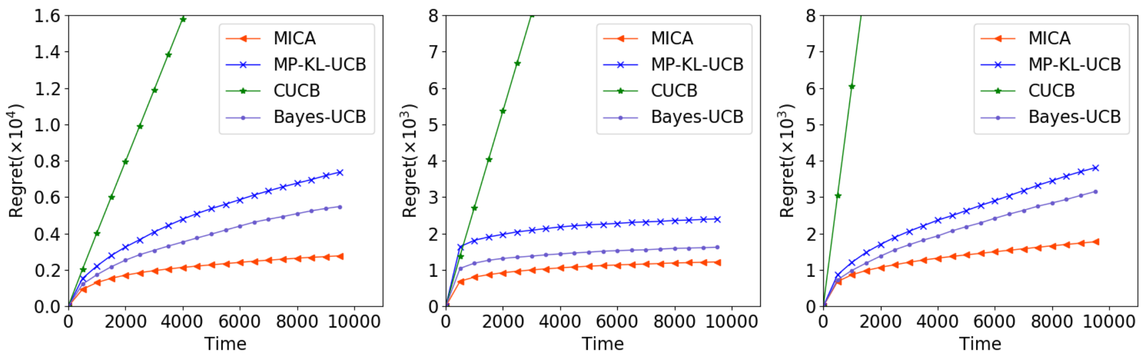
CUCB: At each time slot t, CUCB selects the top M channels ranked by $\hat{\mu}_i(t) + \sqrt{3\log(t)/2N_i(t)}$.

Bayes-UCB: In Bayes-UCB, the successful transmission probability p_i is assumed to be a random variable that follows Beta(a_i, b_i). At each time slot t, the fog node computes $q_i(t), \forall i \in \mathcal{C}$ as follows

$$q_i(t) = Q\left(1 - \tfrac{1}{t(\log T)^c}, \text{Beta}(a_{i(t)}, b_{i(t)})\right), \tag{8}$$

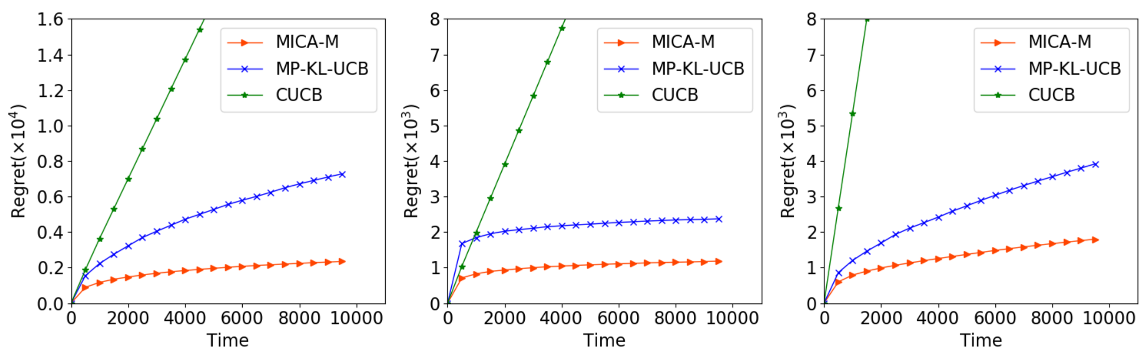
where c is a constant, and $Q(r, \rho)$ is the quantile function associated to the distribution ρ such that $\Pr(X \leq Q(r, \rho)) =$

[1]The regret analysis for the case with heterogenous transmission rates is an interesting future work.



(a) Gradual (b) Steep (c) Lossy

Fig. 1: Comparison of regret in binary-feedback setting



(a) Gradual (b) Steep (c) Lossy

Fig. 2: Comparison of regret in three-level feedback setting

r, for any random variable $X \sim \rho$. Then Bayes-UCB selects the top M channels ranked by $q_i(t)$, and updates the posterior distribution of p_i.

B. Basic Settings

We simulate a fog node with three wireless interfaces under commonly adopted settings [7]. Particularly, the fog node keeps eight available channels with corresponding transmission rates $\mathcal{R} = \{6, 9, 12, 18, 24, 36, 48, 54\}$ Mbps. Meanwhile, we consider the cases of channel allocation with different dimensions of feedback, which is denoted by L. We set $L = 2$ (binary), 3, and 6, respectively. Such settings are detailed as follows. Note that all simulation results are acquired by averaging the results from 50 runs of simulations.

Scenario 1: Binary feedback. We set the successful transmission probabilities in three different settings as follows [7].

- *Gradual*: $p = \{0.95, 0.9, 0.8, 0.65, 0.45, 0.25, 0.15, 0.1\}$
- *Steep*: $p = \{0.99, 0.98, 0.96, 0.93, 0.90, 0.1, 0.06, 0.04\}$
- *Lossy*: $p = \{0.9, 0.8, 0.7, 0.55, 0.45, 0.35, 0.2, 0.1\}$

Scenario 2: Multi-level feedback with $L = 3$. In this case, we set the probability vectors (of 3 dimensions) of channels' quality states in three different settings, as shown in Table II, III and IV, respectively. The successful transmission probabilities associated to the three states are set to be 1, 1/2, and 0, respectively.

TABLE II: Probability Vector in Gradual

	r	6	9	12	18	24	36	48	54
$\boldsymbol{p}$	p_1	0.94	0.85	0.75	0.55	0.35	0.20	0.10	0.06
	p_2	0.02	0.10	0.10	0.20	0.20	0.10	0.10	0.08
	p_3	0.04	0.05	0.15	0.25	0.45	0.70	0.80	0.86

TABLE III: Probability Vector in Steep

	r	6	9	12	18	24	36	48	54
$\boldsymbol{p}$	p_1	0.98	0.97	0.95	0.90	0.85	0.06	0.05	0.02
	p_2	0.02	0.02	0.02	0.06	0.10	0.08	0.02	0.04
	p_3	0.00	0.01	0.03	0.04	0.05	0.86	0.93	0.94

TABLE IV: Probability Vector in Lossy

	r	6	9	12	18	24	36	48	54
$\boldsymbol{p}$	p_1	0.85	0.70	0.60	0.50	0.35	0.20	0.10	0.05
	p_2	0.10	0.20	0.20	0.10	0.20	0.30	0.20	0.10
	p_3	0.05	0.10	0.20	0.40	0.45	0.50	0.70	0.85

Scenario 3: Multi-level feedback with $L = 6$. In this case, we compare MICP, MP-KL-UCB, and CUCB. We fix the successful transmission probabilities of different channels in the case of binary feedback scenario, three-level feedback sce-

nario and six-level feedback scenario. We omit the probability vectors of channels' quality states due to page limit.

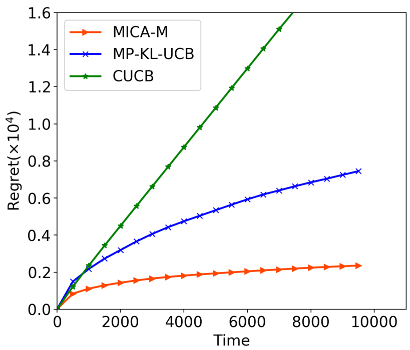


Fig. 3: Comparison of regret in six-level feedback setting

Scenario 4: Empirical information. Note that both MICA and Bayes-UCB employ Bayes optimization techniques. In practice, such schemes, if provided with some empirical *priori* information, such as estimated successful transmission probabilities, may be helpful to accelerate the learning procedure and achieve even higher throughput. Therefore, in this case, we consider the impact of such empirical information on MICA and Bayes-UCB. Such empirical information may have various accuracies. In our simulations, we replace the original prior distributions Beta$(1, 1)$ (a non-informative prior distribution and equivalent to Unif$(0, 1)$) for channel 5 - 8 with empirical information of different estimate accuracies: 1) accurate: Beta(4, 5), Beta(1, 3), Beta(2, 11) and Beta(1, 9); 2) inaccurate: Beta(5, 4), Beta(3, 1), Beta(11, 2) and Beta(9, 1). We use Kullback-Leibler divergence to measure the accuracy of empirical information. We set a threshold $h = 10^{-4}$. If the KL divergence between the distribution of the estimate and the distribution of the real successful transmission probability exceed h, the estimate is accurate; otherwise it is inaccurate.

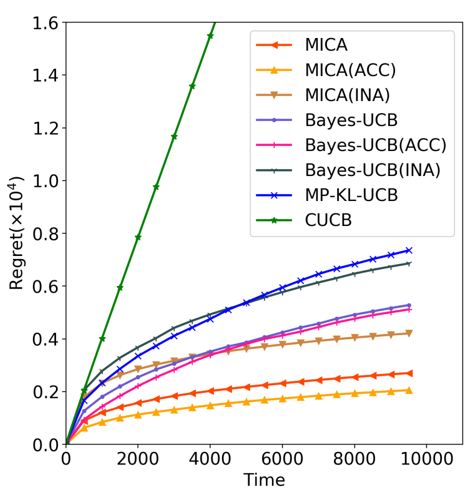


Fig. 4: Comparison of regret with empirical information

C. Simulation Results

Recall that the regret with respect to some algorithms is the gap between the expected cumulative throughput incurred by the algorithms and the optimal cumulative throughput. Therefore, a smaller regret implies a higher throughput. In Figures 1-4, we evaluate the performance of the algorithms by comparing their regret under various scenarios.

(1) From the results in Figure 1, we see that MICA outperforms MP-KL-UCB, CUCB and Bayes-UCB with regret reduction for channel allocation problem with binary feedback. Particularly, with a prior distribution of unknown successful transmission probabilities of channels, Bayes-UCB induces a smaller regret than MP-KL-UCB and CUCB. Nonetheless, MICA still notably outperforms Bayes-UCB with an even lower regret.

(2) From the results in Figure 2, we see that MICA-M has a smaller regret than MP-KL-UCB and CUCB for channel allocation problem with three-level feedback. We do not include Bayes-UCB in this setting since it is unable to handle channel allocation problem with multiple unknown parameters.

(3) Comparing Figure 3 with Figures 1 (a) and 1 (b), we observe that as the number of feedback levels increases, the regrets of MICP-M and MP-KL-UCB remain stable (approximately 2×10^3 for MICP-M and 7×10^3 for MP-KL-UCB at $t = 10^4$). The CUCB always performs worst, but its regret decreases as feedback is quantified at a finer granularity (approximately 1.6×10^4 at $t = 4 \times 10^3$ with binary feedback, 1.3×10^4 at $t = 4 \times 10^3$ with three-level feedback and 0.9×10^4 at $t = 4 \times 10^3$ with six-level feedback).

(4) Note that the curve MIAC(ACC) shows the regret of MACI with accurate empirical information and MICA(INA) shows the regret of MACI with inaccurate empirical information in Figure 4. We see that MICA performs better with accurate empirical information. With inaccurate empirical information, MICA performs a little worse, but still induces a smaller regret than Bayes-UCB, MP-KL-UCB and CUCB, which illustrates the robustness of MICA.

V. Conclusion

In this paper, we studied the multi-interface channel allocation problem in fog computing systems. By formulating the channel allocation problem with binary feedback as a MP-MAB problem, we proposed MICA, a channel allocation scheme using Thompson sampling techniques. Our theoretical analysis shows MICA effectively reduces throughput loss due to decision making under uncertainties with limited regret within an $O(\log T)$ upper bound. We further propose MICA-M, an extension of MICA to handle the channel allocation with multi-level feedback. We conduct extensive simulations to verify the effectiveness and robustness of MICA and MICA-M against state-of-the-art schemes.

References

[1] P. Kyasanur and N. H. Vaidya, "Routing and interface assignment in multi-channel multi-interface wireless networks," in *Proceedings of IEEE WCNC*, 2005.

[2] Y. Wu, J. A. Stankovic, T. He, and S. Lin, "Realistic and efficient multi-channel communications in wireless sensor networks," in *Proceedings of IEEE INFOCOM*, 2008.

[3] J. Qiao, X. S. Shen, J. W. Mark, and L. Lei, "Video quality provisioning for millimeter wave 5g cellular networks with link outage," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5692–5703, 2015.
[4] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive mac protocol for multi-hop wireless networks," in *Proceedings of ACM MobiCom*, 2001.
[5] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," in *Proceedings of ACM MobiCom*, 2002.
[6] G. Judd, X. Wang, and P. Steenkiste, "Efficient channel-aware rate adaptation in dynamic environments," in *Proceedings of ACM MobiSys*, 2008.
[7] R. Combes, A. Proutiere, D. Yun, J. Ok, and Y. Yi, "Optimal rate sampling in 802.11 systems," in *Proceedings of IEEE INFOCOM*, 2014.
[8] R. Combes and A. Proutiere, "Dynamic rate and channel selection in cognitive radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 5, pp. 910–921, 2015.
[9] H. Gupta, A. Eryilmaz, and R. Srikant, "Low-complexity, low-regret link rate selection in rapidly-varying wireless channels," in *Proceedings of IEEE INFOCOM*, 2018.
[10] J. Komiyama, J. Honda, and H. Nakagawa, "Optimal regret analysis of thompson sampling in stochastic multi-armed bandit problem with multiple plays," *arXiv preprint arXiv:1506.00779*, 2015.
[11] J. Li, X. Wu, and R. Laroia, *OFDMA Mobile Broadband Communications: A Systems Approach*. Cambridge University Press, 2013.
[12] P. D. Hoff, *A First Course in Bayesian Statistical Methods*. Springer, 2009, vol. 580.
[13] N. Satish, C. Kim, J. Chhugani, A. D. Nguyen, V. W. Lee, D. Kim, and P. Dubey, "Fast sort on cpus and gpus: a case for bandwidth oblivious simd sort," in *Proceedings of ACM SIGMOD*, 2010.
[14] A. Garivier and O. Cappé, "The kl-ucb algorithm for bounded stochastic bandits and beyond," in *Proceedings of COLT*, 2011.
[15] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial multi-armed bandit: General framework and applications," in *Proceedings of ICML*, 2013.
[16] H. Gupta, A. Eryilmaz, and R. Srikant, "Link rate selection using constrained thompson sampling," in *Proceedings of IEEE INFOCOM*, 2019.
[17] E. Kaufmann, O. Cappé, and A. Garivier, "On bayesian upper confidence bounds for bandit problems," in *Proceedings of JMLR AISTATS*, 2012.

Appendix A
Proof of theorem 1

with a homogeneous transmission rate, from (3), we obtain that

$$\begin{aligned} R_{\widetilde{\pi}}(T) &= \sum_{t=1}^{T}\sum_{j\in I^*} r_j p_j - \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in I_t^{\widetilde{\pi}}} r_i X_i(t)\right] \\ &= r\sum_{t=1}^{T}\left(\sum_{j\in I^*} p_j - \sum_{i\in I_t^{\widetilde{\pi}}} p_i\right) \end{aligned} \tag{9}$$

where $\widetilde{\pi}$ indicts the policy of MICA. For simplicity of notation, we omit this superscript in the following proof.

Without loss of generality, we assume that $p_1 > p_2 > p_3 > \ldots > p_n$. Then $I^* = \{1, 2 \ldots, M\}$. Note that the algorithm can't exploit the ordering. We decompose the regret into the contribution of each channel. For each suboptimal channel i ($i \in \mathcal{C}\backslash I^*$) , let

$$\Delta_i(t) = \begin{cases} \max\limits_{j\in I^*\backslash I_t} p_j - p_i & \text{if } I_t \neq I^*, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

and

$$R_i(T) = \sum_{t=1}^{T} \mathbf{1}\{i \in I_t\}\Delta_i(t). \tag{11}$$

Then we derive the following inequality

$$\begin{aligned} R(T) &= r\sum_{t=1}^{T}\left(\sum_{j\in I^*} p_j - \sum_{i\in I_t} p_i\right) \\ &= r\sum_{t=1}^{T}\left(\sum_{j\in I^*\backslash I_t} p_j - \sum_{i\in I_t\backslash I^*} p_i\right) \\ &\le r\sum_{t=1}^{T}\sum_{i\in I_t\backslash I^*}\left(\max_{j\in I^*\backslash I_t} p_j - p_i\right) \\ &= r\sum_{t=1}^{T}\sum_{i\in I_t\backslash I^*}\Delta_i(t) \\ &= r\sum_{i\in\mathcal{C}\backslash I^*} R_i(T). \end{aligned} \tag{12}$$

Now we define some additional notations for the analysis of $R_i(T)$. Let $N_i(t)$ denote the number of times the channel i has been selected by time slot t. Let $p_M^{(-)} = p_M - \delta$ and $p_i^{(+)} = p_i + \delta$, $i \in \mathcal{C}\backslash I^*$ and $\delta > 0$. δ is sufficiently small such that $p_M^{(-)} \in (p_{M+1}, p_M)$ and $p_i^{(+)} \in (p_i, p_M)$. We also define $\theta^*(t) = \max_{i\in\mathcal{C}}^{(M)} \theta_i(t)$ as the M-th largest posterior sample at time slot t, and $\theta^{**}_{\backslash i,j}(t) = \max_{k\in\mathcal{C}}^{(M-1)} \theta_k(t)$ as the $M-1$-th largest posterior sample except for channel i and j at time slot t. Let $\nu = \frac{p_{M-1}+p_M}{2}$. We define $N_i^{suf}(T) = \frac{\log T}{d(p_i^{(+)},\ p_M^{(-)})}$. Intuitively, $N_i^{suf}(T)$ is the sufficient number of explorations to make sure that channel i is not as good as channel M [10].

Besides, we define some events as follows:

$$\begin{aligned} &A_i(t) = \{i \in I_t\}, \\ &B(t) = \{\theta^*(t) \ge p_M^{(-)}\}, \\ &C_i(t) = \bigcap_{j\in\mathcal{C}\backslash[(I^*\backslash M)\cup\{i\}]} \{\theta^{**}_{\backslash i,j}(t) \ge \nu\}, \\ &D_i(t) = \{N_i(t) < N_i^{suf}(T)\}. \end{aligned}$$

Lemma 1. *The regret of channel i, $i \in \mathcal{C}\backslash I^*$ can be decomposed as :*

$$\begin{aligned} R_i(T) \le &\sum_{t=1}^{T}\mathbf{1}\{B^c(t)\} + \sum_{t=1}^{T}\mathbf{1}\{A_i(t), C_i^c(t)\} \\ &+ \sum_{j\in\mathcal{C}\backslash[(I^*\backslash M)\cup\{i\}]}\sum_{t=1}^{T}\mathbf{1}\{A_i(t), C_i(t), D_i(t), A_j(t)\} \\ &+ \sum_{t=1}^{T}\mathbf{1}\{A_i(t), B(t), D_i^c(t)\} + N_i^{suf}(T)\Delta_{i,M}, \end{aligned} \tag{13}$$

where, for instance, $B^c(t)$ is the complement of $B(t)$, $\{A_i(t), C_i^c(t)\}$ abbreviates $\{A_i(t) \cap C_i^c(t)\}$ and $\Delta_{i,M} = p_M - p_i$.

Proof. The event $A_i(t)$ can be decomposed as

$$\begin{aligned}A_i(t) &\subset B^c(t) \cup \{A_i(t), C_i^c(t)\} \cup \{A_i(t), B(t), C_i(t)\} \\ &\subset B^c(t) \cup \{A_i(t), C_i^c(t)\} \cup \{A_i(t), B(t), D_i^c(t)\} \\ &\quad \cup \{A_i(t), C_i(t), D_i(t)\}. \qquad (14)\end{aligned}$$

Combining (11) and (14), we have

$$\begin{aligned}R_i(T) &= \sum_{t=1}^{T} \mathbf{1}\{i \in I_t\}\Delta_i(t) \\ &= \sum_{t=1}^{T} \mathbf{1}\{A_i(t)\}\Delta_i(t) \\ &\leq \sum_{t=1}^{T} \mathbf{1}\{B^c(t)\} + \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i^c(t)\} \\ &\quad + \sum_{t=1}^{T} \mathbf{1}\{A_i(t), B(t), D_i^c(t)\} \\ &\quad + \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t)\}\Delta_i(t). \qquad (15)\end{aligned}$$

$\Delta_i(t)$ is defined in (10). At time slot t, if M and all suboptimal channels except for i are not selected, *i.e.*, $I_t = \{1, 2, \ldots, M-1, i\}$, then $\Delta_i(t) = \Delta_{i,M}$. Therefore,

$$\begin{aligned}&\sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t)\}\Delta_i(t) \\ \leq\ &\sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t)\}\Delta_{i,M} \\ &+ \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t), \bigcup_{j\in\mathcal{C}\setminus[(I^*\setminus M)\cup\{i\}]} A_j(t)\} \\ \leq\ &\sum_{t=1}^{T} \mathbf{1}\{A_i(t), D_i(t)\}\Delta_{i,M} \\ &+ \sum_{j\in\mathcal{C}\setminus[(I^*\setminus M)\cup\{i\}]} \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t), A_j(t)\} \\ \leq\ &N_i^{suf}(T)\Delta_{i,M} \\ &+ \sum_{j\in\mathcal{C}\setminus[(I^*\setminus M)\cup\{i\}]} \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t), A_j(t)\}. \qquad (16)\end{aligned}$$

Combining (15)-(16), we complete the proof of Lemma 1. $\square$

Then we bound the terms in the right-hand side of Lemma 1. From Lemma 3 and Lemma 4 in [10], the first four terms are bounded in expectation as:

$$\mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{B^c(t)\}\right) = O\left(\frac{1}{(p_M - p_M^{(-)})^2}\right) = O\left(\frac{1}{\delta^2}\right), \qquad (17)$$

$$\mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i^c(t)\}\right) = O(\log\log T), \qquad (18)$$

$$\begin{aligned}&\mathbb{E}\left(\sum_{j\in\mathcal{C}\setminus[(I^*\setminus M)\cup\{i\}]} \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t), A_j(t)\}\right) \\ &= O(\log\log T), \qquad (19)\end{aligned}$$

$$\mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{A_i(t), B(t), D_i^c(t)\}\right) = O\left(\frac{1}{\delta^2}\right). \qquad (20)$$

Now we just need to evaluate $N_i^{suf}(T) = \frac{\log T}{d(p_i^{(+)},\ p_M^{(-)})}$. From the convexity of KL divergence, there exists a constant $a_i = a_i(p_i, p_M) > 0$,such that

$$d(p_i^{(+)}, p_M^{(-)}) = d(p_i + \delta, p_M - \delta) \geq (1 - a_i\delta)d(p_i, p_M). \qquad (21)$$

Therefore, combining (17)-(21) with (12)-(13), we have

$$\begin{aligned}&\mathbb{E}[R(T)] \\ &\leq r \sum_{i\in\mathcal{C}\setminus I^*} \mathbb{E}[R_i(T)] \\ &\leq r \sum_{i\in\mathcal{C}\setminus I^*} \mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{B^c(t)\}\right) + \mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i^c(t)\}\right) \\ &\quad + \mathbb{E}\left(\sum_{j\in\mathcal{C}\setminus[(I^*\setminus M)\cup\{i\}]} \sum_{t=1}^{T} \mathbf{1}\{A_i(t), C_i(t), D_i(t), A_j(t)\}\right) \\ &\quad + \mathbb{E}\left(\sum_{t=1}^{T} \mathbf{1}\{A_i(t), B(t), D_i^c(t)\}\right) + N_i^{suf}(T)\Delta_{i,M} \\ &\leq r \sum_{i\in\mathcal{C}\setminus I^*} \frac{\log T \Delta_{i,M}}{(1 - a_i\delta)d(p_i, p_M)} + O\left(\frac{1}{\delta^2}\right) + O(\log\log T). \qquad (22)\end{aligned}$$

If $a_i\delta \leq 1/2$, $(1 - a_i\delta)^{-1} \leq 1 + 2a_i\delta$. Let $\epsilon < \frac{1}{2}$, and $\delta = \epsilon / \max_{i\in\mathcal{C}\setminus I^*} a_i = \Theta(\epsilon)$. Then we can rewrite (20) as

$$\begin{aligned}&\mathbb{E}[R(T)] \\ &\leq r \sum_{i\in\mathcal{C}\setminus I^*} \frac{(1 + 2\epsilon)\log T \Delta_{i,M}}{d(p_i, p_M)} + O(\epsilon^{-2}) + O(\log\log T). \qquad (23)\end{aligned}$$

By letting $\epsilon = O((\log T)^{-1/3})$, we obtain that

$$\mathbb{E}[R(T)] \leq r \sum_{i\in\mathcal{C}\setminus I^*} \frac{\log T \Delta_{i,M}}{d(p_i, p_M)} + O((\log T)^{2/3}). \qquad (24)$$

Without the ordering assumption in the proof, let $m \triangleq \arg\min_{j\in I^*} p_j$, *i.e.*, the channel whose p_j is the minimum in the optimal channel set I^* and $\Delta_{i,m} = p_m - p_i$. The upper bound of the regret is

$$\mathbb{E}[R(T)] \leq r \sum_{i\in\mathcal{C}\setminus I^*} \frac{\log T \Delta_{i,m}}{d(p_i, p_m)} + O((\log T)^{2/3}). \qquad (25)$$

Here we complete the proof of Theorem 1.