

Index-based Update Policy for Minimizing Information Mismatch with Markovian Sources

Sunjung Kang and Changhee Joo

Abstract—We consider a scenario where a base station collects time-varying state information from multiple sources, and makes system decisions based on the collected information. When the information update is constrained to one source at a time, the state information at the base station can be stale and different from actual state of the sources, in which case the base station can make a false decision due to the information mismatch (or error). In this paper, we assume that the update decisions are made at the base station without current state information, and consider the problem of minimizing the information mismatch under limited communication capability. For two-state Markovian source, we consider two different types of estimators at the base station, and characterize the optimal update policy. For the symmetric case, we can obtain the closed-form average cost. Further, with multiple symmetric sources, we show that the problem is indexable and obtain the close-form Whittle's index for the two different types of estimator.

Index Terms—Remote estimation, restless multi-armed bandit, wireless networks, Whittle's index.

I. INTRODUCTION

IN recent years, the applications of real-time monitoring systems and cyberphysical systems, such as drone monitoring systems and medical monitoring systems, have been increasing. In these applications, sensors transmit update packets with time-varying sensing information to a remote monitor (or a base station/receiver) [1]. The remote monitor estimates the sources' states from the transmitted information and makes system decisions as needed. To this end, keeping the information fresh through timely updates is critical in these applications. Although it is ideal to update fresh information continuously, this is often impractical due to the limitation of the network resources, which causes the information at the monitor to be stale. In a commercial IoT systems, a base station can take the role of the remote monitor. We use the terms of base station and remote monitor interchangeably, similarly for the terms of source and sensor.

Regarding the performance metric for information update, recently the age of information (AoI) has attracted much

attention [2], [3]. AoI is a performance metric measuring the freshness of information, which is the time elapsed since the most recently generated packet received by the receiver was sampled at the source. In [4], [5], the authors investigate the optimal update rate at a source to minimize the AoI and show that the *zero-waiting* update policy is not optimal in some scenarios (e.g., random transmission time). The state-space collapse of the AoI under heavy traffic has been observed in [6]. In [7], the authors develop optimal update policies for minimizing the re-defined AoI to address the normal and alarm states of a stochastic process in an energy harvesting status update system. A scenario where a base station transmits timely information to multiple nodes is studied in [8], where the AoI minimization problem with multiple nodes is formulated as a restless multi-armed bandit (RMAB) problem, and a low-complexity algorithm using Whittle's index theory has been developed. In [9], the authors develop the centralized scheduling policies including Whittle's index policy to minimize the average AoI under several system constraints. In [10], the authors analytically show the optimality of Whittle's index policy for minimizing the AoI in many-source networks.

Remote estimation has been studied [11]–[18], where the freshness of information is measured as how close the information state between a source and the remote monitor. In [11], the authors claim that the sampling strategy that minimizes the AoI does not always minimize the estimation error by taking an example of a source where information state follows a two-state Markov chain. In [12], the authors investigate sampling and scheduling policies to minimize the average AoI in networks with multiple sources updating a common receiver. In particular, for two-state Markovian sources, the authors employ the *sample-at-change* sampling policy and Whittle's index policy for minimizing the AoI as a scheduling policy. Through simulations, the authors compare the proposed update policy with the optimal policy for minimizing the estimation error and show that the proposed update policy is near optimal, especially with a large number of sources. In [13], the authors develop Whittle's index policy for minimizing the uncertainty of information measured by Shannon's entropy in networks where a common receiver is updated by distributed sources of which states evolve according to a two-state Markov chain. The authors in [14], [15] study the optimal sampling rate to minimize the mean square error of information state of the Wiener-process and Ornstein-Uhlenbeck, respectively. In [16], [17], the authors consider a scenario where information state is an independent and identically distributed (i.i.d.) process and there is a communication cost to transmit information to a remote estimator. They showed that an optimal policy for

Manuscript received February 25, 2021; revised June 28, 2021; approved for publication by Nikolaos Papas, Division III Editor, August 10, 2021.

This work is supported in part by the NRF grant funded by the Korea government (MSIT) (No. NRF-2021R1A2C2013065), and by the MSIT, Korea, under the ICT Creative Consilience program (IITP-2021-2020-0-01819) supervised by the IITP.

S. Kang is with ECE, The Ohio State University, USA, email: kang.853@osu.edu.

C. Joo is with the Department of CSE, Korea University, Korea, email: changhee@korea.ac.kr.

C. Joo is the corresponding author.

Digital Object Identifier: 10.23919/JCN.2021.000027

Creative Commons Attribution-NonCommercial (CC BY-NC).

This is an Open Access article distributed under the terms of Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided that the original work is properly cited.

minimizing the estimation error is of threshold type.

In this paper, we consider a scenario where a base station collects state information from multiple two-state Markovian sources, and the information update is constrained to one source at a time due to the limitation of networks resources. This model is somewhat simple but captures the essential aspects of how the base station has to sequentially select a source to minimize the information mismatch (or estimation error) accounting for the Markovian behavior and the communication cost. Our problem is different from [9], [10], [13] in that this paper considers the estimation error as a metric to measure the freshness of information instead of the AoI [9], [10] or the uncertainty of information [13], and also different from [11], [14]–[17] in that the base station has uncertainty about not only future evolution of the information state, but also its history. Further, unlike [12], where the proposed scheduling policy minimizes the AoI with the rationale that the estimation error can be minimized accordingly under the assumption that the estimation error increases as the AoI increases, this paper does not use the assumption between the AoI and the estimation error and aims to design a scheduling policy that minimizes the estimation error.

We formulate our problem as a partially observable Markov decision process (POMDP), where the belief state calculated by the Bayes rule is sufficient statistics. Some POMDP problem can be solved using multi-armed bandit (MAB) techniques, in which a scheduler selects one arm (or bandit), and receives a reward from the chosen arm. The state of arms can change in a Markovian manner. There are two types of the MAB process: rested and restless. For the former, only the chosen arm changes its state and the states of the other do not, in which case, it is shown that the Gitten's index policy is optimal [19]. For the latter, the states of all arms change, and the transition law of a state can be different between the chosen arm and the others. It is shown that finding an optimal solution to this RMAB problem is PSPACE-hard [20]. In [21], Whittle relaxes the RMAB problem, and shows that if the problem satisfies some condition (or the problem is *indexable*), then an optimal solution is an index-type policy and the index is called *Whittle's index*. In [22], the authors study the sufficient conditions for the indexability for general source models and develop an algorithm to compute Whittle's indices for indexable restless bandits.

It has been shown that Whittle's index policy that activates an arm with the highest Whittle's index (or K highest when multi-arm activation is allowed), is asymptotically optimal when the arms are stochastically identical [23], [24]. Further, it has been shown that Whittle's index achieves heuristically near-optimal performance in many applications, including AoI [8]–[10], dynamic spectrum allocation [24], queueing system [25], and resource allocation problem [26]. In [27], the authors study a scenario where each bandit is modeled as a two-state Markov chain and when an arm is activated, the player receives a signal with noise, i.e., there is a probability that the actual state is 0 even when the received signal is 1.

In this paper, we address a wireless scheduling problem to minimize the information mismatch between the base station and distributed sources by exploiting Whittle's index. The

most related studies are [28], [29] that apply Whittle's index to an estimation problem with two-state Markov chain model. Our work is different from [28] in that we aim to minimize the penalty incurred when the estimated states are different from the actual states of sources and thus the penalty depends on both the source state and the estimation, while [28] aims to maximize the rewards that are incurred when an activated arm is in a specific state (i.e., rewards depend only on the source state).

Also, our work is different from [29] in that we investigate the average cost criterion over infinite horizon while [29] investigates the discount reward criterion over infinite horizon. It is known that the infinite sum of discount cost/reward always converges if the one-step cost/reward is bounded, while the average cost/reward over infinite horizon can diverge even when the one-step cost/reward is bounded [30]. Further, the results of [29] are limited to the case of discount factor $\beta \leq 1/2$, which implies that the results are applicable only to severe discounting scenarios, and cannot be extended (e.g., by taking $\beta \rightarrow 1$) to our problem with the average cost criterion.

Our contributions can be summarized as follows.

- We analytically show that an optimal solution to the problem of single source with communication cost is of threshold type, for two different types of estimators. Further, for the special case of symmetric source, we obtain the minimum expected average penalty in the closed form.
- For multi-source scenarios, we model our problem as an RMAB, and show that our problem of minimizing the average cost is indexable for symmetric sources, and obtain the closed-form Whittle index. As such, we develop a low-complexity scheduling algorithm using the Whittle's index theory for the information error minimization.
- We show through simulations that our policy achieves near-optimal performance.

The rest of paper is organized as follows. In Section II, we describe the system model and formulate our problem. In Section III, we characterize the optimal policy in a single source network with communication cost. In Section IV, we show that our problem is indexable, and obtain a closed-form Whittle index. We verify through simulations that our algorithm performs well in Section V, and conclude our work in Section VI.

II. SYSTEM MODEL

We consider the scenario where a base station collects time-varying state information from N sources, and makes system decisions based on the collected information. We consider a time-slotted system, and at each time slot, the base station can request state information from at most one source. We assume that sources are numbered from 1 to N , and each source can have two states 0 and 1. The state of source i varies following a Markov process with transition probability $P_i = [p_{00}^{(i)}, p_{01}^{(i)}; p_{10}^{(i)}, p_{11}^{(i)}]$ and it is assumed that the transition probability is fixed and known to the base station.

Let $S_i(t) \in \{0, 1\}$ denote the state of source i at the beginning of time slot t , and let $D_i(t) \in \{0, 1\}$ denote the

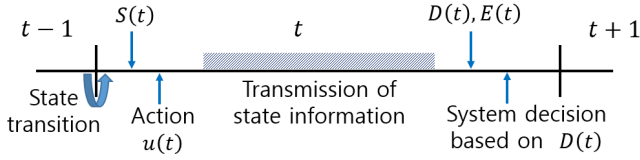


Fig. 1. Structure of time slot.

latest updated state of source i at the base station at time slot t . $\mathbf{S}(t) = [S_1(t), \dots, S_N(t)]$ and $\mathbf{D}(t) = [D_1(t), \dots, D_N(t)]$ denote their vector, respectively. Let $\mathbf{u}(t) = [u_1(t), \dots, u_N(t)]$ denote the action vector in time slot t with $u_i(t) \in \{0, 1\}$, where $u_i(t) = 1$ implies that the base station requests state information to source i . The base station can request the information to at most one source¹ in a time slot, i.e., $\sum_{i=1}^N u_i(t) \leq 1$ for all t . We assume that the communication between the base station and the sources is reliable, and thus there is neither loss nor error in the information delivery. If the base station obtains the information from source i at time slot t , we have $D_i(t) = S_i(t)$, and otherwise, $D_i(t) = D_i(t-1)$. Let $E_i(t)$ denote the difference between actual state and latest updated state of source i at time slot t , i.e., $E_i(t) = |S_i(t) - D_i(t)|$. For source i with $u_i(t) = 1$, we have $E_i(t) = 0$ since $D_i(t) = S_i(t)$. We define the total error $E(t)$ over N sources at time slot t as $E(t) = \sum_{i=1}^N E_i(t)$. We note that the state transition occurs on the boundary of time slot, and thus $S_i(t)$ is determined at the beginning of time slot. Then the base station makes decision $u(t)$, which determines $D(t)$ and thus $E(t)$, as shown in Fig. 1.

Let $\mathcal{H}(t)$ denote the history of the base station up to time slot t , i.e., $\mathcal{H}(t) = \{\mathbf{u}(1), S_{a(1)}(1), \mathbf{D}(1), \dots, \mathbf{u}(t), S_{a(t)}(t), \mathbf{D}(t)\}$ with $\mathcal{H}(0) = \emptyset$, where $a(t) = i$ such that $u_i(t) = 1$. A policy $\pi = (\pi(t))_{t=1}^\infty$ is a sequence of mappings $\pi(t) : \mathcal{H}(t-1) \rightarrow \mathbf{u}(t)$. Our objective is to find policy π that minimizes the expected average error (or mismatch) over infinite time horizon:

$$\begin{aligned} & \text{minimize } \mathbb{E}_\pi \left[\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E(t) \mid \mathbf{D}(0) \right] \\ & \text{subject to } \sum_{i=1}^N u_i(t) \leq 1 \text{ for all } t, \end{aligned} \quad (1)$$

where $\mathbf{D}(0)$ is an initial value at the beginning of the first time slot.

We consider a belief error vector $\mathbf{e}(t) = [e_1(t), \dots, e_N(t)]$, where $e_i(t)$ denotes the belief error probability for source i at time slot t if its state is not updated in time slot t . Specifically, the belief error evolves according to the Bayes rule as

$$e_i(t) = \begin{cases} p_{D\bar{D}}^{(i)}, & \text{if } u_i(t-1) = 1, \\ \mathcal{T}_D^{(i)}(e_i(t-1)), & \text{if } u_i(t-1) = 0, \end{cases} \quad (2)$$

¹Since we are investigating strongly decomposable index policies (e.g., myopic index and Whittle's index), this constraint can be relaxed and our results can be applied to more general scenarios where $K \geq 1$ out of N sources can update the receiver simultaneously [23].

where $D = D_i(t-1)$, $\bar{D} = 1 - D$, and $\mathcal{T}_D^{(i)}(e)$ denotes the one-step belief error

$$\mathcal{T}_D^{(i)}(e) = \begin{cases} ep_{11}^{(i)} + (1-e)p_{01}^{(i)}, & \text{if } D = 0, \\ ep_{00}^{(i)} + (1-e)p_{10}^{(i)}, & \text{if } D = 1. \end{cases} \quad (3)$$

Note that the base station can compute $e_i(t)$ from $u_i(t-1)$, $D_i(t-1)$, and $e_i(t-1)$. Also since it makes an update decision without full knowledge of system states, it can be modeled as a POMDP. It is known that sufficient statistics for making optimal decision under POMDP is the belief states, which is the probability distribution over states and can be updated using the Bayes rule [31] as in (3).

We formulate the error minimization problem as a RMAB problem with tuple $(D_i(t-1), e_i(t))$ as a state of arm i at t th round. If active action $u_i(t) = 1$ is applied to arm i , state $(D_i(t), e_i(t+1))$ is set as $D_i(t) = S_i(t)$, and $e_i(t+1) = p_{D_i\bar{D}_i}^{(i)}$. Otherwise, $e_i(t+1)$ evolves according to (3). At each time slot, consider a policy that assigns an index to each arm, which is a real-valued function of a state of an arm, and plays (or activates) the arm with the highest index. Such a policy is called as an index policy. Further, the index policy is called *strongly decomposable* if the index of each arm depends only on the characteristics of the arm. A simple example of strongly decomposable index policy is the myopic policy that directly uses $e_i(t)$ as an index of arm i at the time slot t to decide $u_i(t)$. This myopic policy generally does not perform well since it ignores the impact of current action on future errors. Another strongly decomposable index policy can be obtained using Whittle's index [21], which achieves near-optimal performance in a wide range of applications [8], [12], [24]–[26], [32]. The main challenge to adopt the Whittle's index is the indexability of the problem. In this work, we show that our problem is indexable when the source dynamics can be represented by a symmetric Markov chain, i.e., $p_{01} = p_{10}$.

III. THRESHOLD POLICY FOR SINGLE SOURCE

As a first step, we obtain an optimal solution to single source problem with communication (or transmission) cost. Note that, in the original problem (1), multiple sources are coupled through the constraint. By considering a dual problem, we can relax the constraint and decompose the problem to N per-source problems with communication cost, which can be interpreted as the Lagrangian multiplier associated with the time portion that packets are transmitted. We will use the cost as the index in our decision.

At each time slot, the base station has an option of either collecting the state information at communication cost (active action) or skipping to the next time at the expense of information inaccuracy (passive action). For notational convenience, we omit subscript (or superscript) i for the single source case. Let $c \geq 0$ denote the communication cost that occurs when the base station collects the state information from the source. Then, the penalty at time slot t is $P(t) = u(t)c + (1-u(t))E(t)$, and we can define the problem

of minimizing the expected average penalty over infinite time horizon as

$$\text{minimize } \mathbb{E}_\pi \left[\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T P(t) \mid D(0) \right]. \quad (4)$$

It is known that a stationary optimal solution to (4) does not necessarily exist under average cost criterion over infinite time horizon [33]. In this work, we show the existence of a stationary optimal policy that solves (4) by considering the problem with vanishing discounts [28], [33], [34].

To this end, we consider the problem with discounted total penalty:

$$\text{minimize } \mathbb{E}_\pi \left[\lim_{T \rightarrow \infty} \sum_{t=1}^T \beta^{t-1} P(t) \mid D(0) \right], \quad (5)$$

with discount factor $\beta \in (0, 1)$, and show that it always admits a stationary optimal solution of threshold type. Then, we show that, for symmetric source, there exists a stationary optimal policy that solves the original long-term average penalty problem.

A. Threshold Policy

We first describe a discrete-time Markov Decision Process (MDP), whose state at time slot t is represented by (D, e) . For each state, the base station has two possible actions, $u = 0$ (passive action) and $u = 1$ (active action). If the passive action is taken, then the state evolves to $(D, \mathcal{T}_D(e))$ with probability 1. If the active action is taken with communication cost c , the state evolves to $(D, \mathcal{T}_D(0))$ with probability $1 - e$, and to $(\bar{D}, \mathcal{T}_{\bar{D}}(0))$ with probability e , where $\bar{D} = 1 - D$ (i.e., the flipped state of D). The penalty under state (D, e) and action u is given as $uc + (1 - u)e$.

Let $V_{\beta,c}(D, e)$ denote the minimum expected total penalty given initial state (D, e) for the problem with discount factor β and communication cost c . Let $V_{\beta,c}(D, e; u)$ denote the expected total penalty obtained by one-time deviation, i.e., by taking action u at this state and then following the optimal policy. Then, we can write the Bellman optimality equations as

$$\begin{aligned} V_{\beta,c}(D, e) &= \min\{V_{\beta,c}(D, e; u=0), V_{\beta,c}(D, e; u=1)\}, \\ V_{\beta,c}(D, e; u=0) &= e + \beta V_{\beta,c}(D, \mathcal{T}_D(e)), \\ V_{\beta,c}(0, e; u=1) &= c + \beta[eV_{\beta,c}(1, p_{10}) + (1-e)V_{\beta,c}(0, p_{01})], \\ V_{\beta,c}(1, e; u=1) &= c + \beta[eV_{\beta,c}(0, p_{01}) + (1-e)V_{\beta,c}(1, p_{10})]. \end{aligned} \quad (6)$$

It is known that under the discounted total penalty criterion over infinite time horizon, there always exists a stationary optimal policy that satisfies the above Bellman equations (6) [30].

Remark 1: We implicitly assume that once the based station is updated with a new status information D , it holds and uses D as an estimate of the source's status until the next update occurs. Besides this update-and-hold (UAH) estimator, another promising estimation of the source's status is maximum a posteriori (MAP) [29]. Given (D, e) , if $e > 0.5$, the MAP estimator uses \bar{D} as an estimate of the source's status. Thus, the Bellman optimality equations for MAP estimator can be written as (6) using $\tilde{V}_{\beta,c}(D, e) = \mathbb{I}\{e \leq 0.5\}V_{\beta,c}(D, e) +$

$\mathbb{I}\{e > 0.5\}V_{\beta,c}(\bar{D}, 1 - e)$ instead of $V_{\beta,c}(D, e)$ in 2–4 lines. Throughout the paper, we assume the UAH estimator if otherwise stated, and remark on how the techniques can be also applied to the MAP estimator.

Let $u_c^*(D, e)$ denote the optimal action given state (D, e) with a fixed cost c . It can be written as

$$u_c^*(D, e) = \arg \min_{u \in \{0, 1\}} V_{\beta,c}(D, e; u). \quad (7)$$

In order to prove that an optimal policy is of threshold type, we first show that the value function $V_{\beta,c}(D, e)$ is concave in e (Lemma 1), and that there exists a cost range such that a deterministic decision is an optimal action outside of the range (Lemma 2).

Lemma 1: For fixed c , the value function $V_{\beta,c}(D, e)$ is concave in e .

Proof 1: For notational convenience, we drop β, c in $V_{\beta,c}$, and let $V_n(D, e)$ denote the value function at n th iteration under value iterations starting at (D, e) . We have

$$\begin{aligned} V_0(D, e) &= 0 \text{ for all } D \in \{0, 1\}, e \in [0, 1], \\ V_1(D, e) &= \min\{c, e\}, \\ V_{n+1}(0, e) &= \min\{c + \beta[eV_n(1, p_{10}) + (1-e)V_n(0, p_{01})], \\ &\quad e + \beta V_n(0, \mathcal{T}_0(e))\}, \\ V_{n+1}(1, e) &= \min\{c + \beta[eV_n(0, p_{01}) + (1-e)V_n(1, p_{10})], \\ &\quad e + \beta V_n(1, \mathcal{T}_1(e))\}. \end{aligned} \quad (8)$$

Clearly, $V_1(0, e)$ and $V_1(1, e)$ are concave in e . We now show the concavity of $V_n(0, e)$ and $V_n(1, e)$ by induction. Suppose that $V_n(0, e)$ and $V_n(1, e)$ are concave in e . Then, the first term in (8) is linear in e , and the second term is concave² in e since $V_n(0, \mathcal{T}_0(e)) = V_n(0, ep_{11} + (1-e)p_{01})$. Thus, $V_{n+1}(0, e)$, the minimum of two concave functions, is also concave in e . Similarly, we can obtain the same result for $V_{n+1}(1, e)$. From the convergence theorem of value iterations [30], we have $V_n(D, e) \rightarrow V(D, e)$ as $n \rightarrow \infty$, and thus $V(D, e)$ is concave in e for fixed c . This completes the proof.

Remark 2 (MAP): Under the MAP estimator, the second term of (8) will be replaced with $e + \beta \tilde{V}_n(0, \mathcal{T}_0(e))$, where $\tilde{V}_n(0, \mathcal{T}_0(e)) = \mathbb{I}\{\mathcal{T}_0(e) \leq 0.5\}V_n(0, \mathcal{T}_0(e)) + \mathbb{I}\{\mathcal{T}_0(e) > 0.5\}V_n(1, 1 - \mathcal{T}_0(e))$. Since both $V_n(0, \mathcal{T}_0(e))$ and $V_n(1, 1 - \mathcal{T}_0(e))$ are concave, their non-negative weighted sum of concave functions is also concave. Thus, Lemma 1 holds for the MAP estimator.

We now consider a range $[c_L, c_H]$, beyond which an optimal policy is deterministic. Specifically, when $c < c_L$, the optimal policy has active action for all (D, e) , and when $c \geq c_H$, it has passive action for all (D, e) . In the following lemma, we show such c_L and c_H exist.

Lemma 2: For the single source network problem with communication cost c , there exists an optimal policy that makes the deterministic decision beyond range $[c_L, c_H]$ with

$$\begin{aligned} c_L &= 0, \\ c_H &= \max\left\{\frac{1-\beta+\beta p_{01}-\beta p_{10}}{(1-\beta)(1-\beta p_{11}+\beta p_{01})}, \frac{1-\beta+\beta p_{10}-\beta p_{01}}{(1-\beta)(1-\beta p_{00}+\beta p_{10})}\right\}. \end{aligned} \quad (9)$$

²If $f(x)$ is concave in x , then $f(ax + b)$ is concave in x for $a \neq 0$.

Proof 2: Note that when c is very small, an optimal policy has $u = 1$ for all $(D, e) \in \{0, 1\} \times [0, 1]$, and when c is very large, an optimal policy has $u = 0$ for all (D, e) . Under this assumption, we solve Bellman equations (6) to obtain c_L and c_H , and then show that the result satisfies the Bellman optimality equations.

Case 1) When c is very small, updating is an optimal decision, and we will have

$$V_{\beta,c}(D, e; u = 1) < V_{\beta,c}(D, e; u = 0) \text{ for all } (D, e). \quad (10)$$

This and (6) lead to

$$\begin{aligned} V_{\beta,c}(0, p_{01}) &= V_{\beta,c}(0, p_{01}; u = 1) \\ &= c + \beta [p_{01} V_{\beta,c}(1, p_{10}) + (1 - p_{01}) V_{\beta,c}(0, p_{01})], \end{aligned} \quad (11)$$

$$\begin{aligned} V_{\beta,c}(1, p_{10}) &= V_{\beta,c}(1, p_{10}; u = 1) \\ &= c + \beta [p_{10} V_{\beta,c}(0, p_{01}) + (1 - p_{10}) V_{\beta,c}(1, p_{10})]. \end{aligned} \quad (12)$$

Combining (11) and (12), it can be easily shown that $V_{\beta,c}(0, p_{01}) = V_{\beta,c}(1, p_{10}) = c/(1 - \beta)$, with which and (6), we obtain that, for any $(D, e) \in \{0, 1\} \times [0, 1]$,

$$\begin{aligned} V_{\beta,c}(D, e; u = 1) &= \frac{c}{1 - \beta}, \\ V_{\beta,c}(D, e; u = 0) &= e + \beta V_{\beta,c}(D, \mathcal{T}_D(e)) \\ &= e + \beta V_{\beta,c}(D, \mathcal{T}_D(e); u = 1) \\ &= e + \frac{\beta c}{1 - \beta}. \end{aligned} \quad (13)$$

To satisfy (10), it is required that $c \leq e$ for all $e \in [0, 1]$. Thus, we can conclude that for all $c < c_L = 0$, (13) satisfies the Bellman optimality equations. In addition, in this case, the optimal action is $u = 1$.

Case 2) Let $\mathcal{T}_D^k(e)$ denote the k -step belief error update of e given $D \in \{0, 1\}$, i.e., the belief of the base station when it does not collect the state information from the source for k consecutive slots. It is known in [35] that the belief error evolves as

$$\begin{aligned} \mathcal{T}_0^k(e) &= \frac{p_{01} - (p_{11} - p_{01})^k (p_{01} - (1 + p_{01} - p_{11})e)}{1 + p_{01} - p_{11}}, \\ \mathcal{T}_1^k(e) &= \frac{p_{10} - (p_{00} - p_{10})^k (p_{10} - (1 + p_{10} - p_{00})e)}{1 + p_{10} - p_{00}}. \end{aligned} \quad (14)$$

When c is very large, from (6) and (14), we can obtain

$$V_{\beta,c}(D, e; u = 0) \leq V_{\beta,c}(D, e; u = 1) \text{ for all } (D, e), \quad (15)$$

$$\begin{aligned} V_{\beta,c}(0, p_{01}) &= V_{\beta,c}(0, p_{01}; u = 0) \\ &= \sum_{t=0}^{\infty} \beta^t \mathcal{T}_0^t(p_{01}) = \frac{p_{01}}{(1 - \beta)(1 - \beta p_{11} + \beta p_{01})}, \end{aligned} \quad (16)$$

$$\begin{aligned} V_{\beta,c}(1, p_{10}) &= V_{\beta,c}(1, p_{10}; u = 0) \\ &= \sum_{t=0}^{\infty} \beta^t \mathcal{T}_1^t(p_{10}) = \frac{p_{10}}{(1 - \beta)(1 - \beta p_{00} + \beta p_{10})}. \end{aligned} \quad (17)$$

Using (6), (16) and (17), we have that, for any $e \in [0, 1]$,

$$\begin{aligned} V_{\beta,c}(0, e; u = 1) &= \frac{\beta(p_{10} - p_{01})e + \beta p_{01}}{(1 - \beta)(1 - \beta p_{11} + \beta p_{01})} + c, \\ V_{\beta,c}(0, e; u = 0) &= \frac{(1 - \beta)e + \beta p_{01}}{(1 - \beta)(1 - \beta p_{11} + \beta p_{01})}, \\ V_{\beta,c}(1, e; u = 1) &= \frac{\beta(p_{01} - p_{10})e + \beta p_{10}}{(1 - \beta)(1 - \beta p_{00} + \beta p_{10})} + c, \\ V_{\beta,c}(1, e; u = 0) &= \frac{(1 - \beta)e + \beta p_{10}}{(1 - \beta)(1 - \beta p_{00} + \beta p_{10})}. \end{aligned} \quad (18)$$

Combining the above equations and (15), we have $c \geq \frac{(1 - \beta + \beta p_{10} - \beta p_{01})e}{(1 - \beta)(1 - \beta p_{11} + \beta p_{01})}$, and $c \geq \frac{(1 - \beta + \beta p_{01} - \beta p_{10})e}{(1 - \beta)(1 - \beta p_{00} + \beta p_{10})}$. Hence, for all $c \geq c_H$, where $c_H := \max\{\frac{1 - \beta + \beta p_{10} - \beta p_{01}}{(1 - \beta)(1 - \beta p_{11} + \beta p_{01})}, \frac{1 - \beta + \beta p_{01} - \beta p_{10}}{(1 - \beta)(1 - \beta p_{00} + \beta p_{10})}\}$, (18) satisfies the Bellman optimality equations, and in this case, the optimal action is $u = 0$.

Remark 3: Cost c_L makes the two actions indifferent at $e = 0$, and similarly, cost c_H provides an indifferent decision at $e = 1$.

Remark 4: If $p_{11} \geq p_{01}$, the k -step belief error $\mathcal{T}_0^k(e)$ monotonically increases to π_1 for $e < \pi_1$ with respect to k , and monotonically decreases to π_1 for $e > \pi_1$. Hence, it converges to π_1 for all e . If $p_{11} < p_{01}$, then $\mathcal{T}_0^{2k}(e)$ monotonically increases to π_1 and $\mathcal{T}_0^{2k+1}(e)$ monotonically decreases to π_1 for $e < \pi_1$ with respect to k , and vice versa for $e > \pi_1$. Thus, it also converges to π_1 for all e while it alternates around π_1 . The similar results hold for $\mathcal{T}_1^k(e)$, which converges to π_0 for all e .

Remark 5 (MAP): Note that, from (3), it can be easily shown that $\mathcal{T}_{\bar{D}}(1 - e) = 1 - \mathcal{T}_D(e)$. Hence, under the MAP estimator, the one-step evolution of state (D, e) is $\min\{(D, \mathcal{T}_D(e)), (\bar{D}, 1 - \mathcal{T}_D(e))\}$, where the minimum compares the second value e of the pair (D, e) . Given $(\bar{D}, 1 - e)$, the state evolves to $\min\{(\bar{D}, \mathcal{T}_{\bar{D}}(1 - e)), (D, 1 - \mathcal{T}_{\bar{D}}(1 - e))\} = \min\{(\bar{D}, 1 - \mathcal{T}_D(e)), (D, \mathcal{T}_D(e))\}$. Thus, k -step evolution of $\min\{(D, e), (\bar{D}, 1 - e)\}$ for the MAP estimator is $\min\{(D, \mathcal{T}_D^k(e)), (\bar{D}, 1 - \mathcal{T}_D^k(e))\}$. Due to the different k -step evolution, the upper bound (9) no longer holds for the MAP estimator. However, we can still show that the upper bound is finite, $c_H < \infty$, using its k -step error evolution of $\min\{\mathcal{T}_D^k(e), 1 - \mathcal{T}_D^k(e)\}$, as well as the same lower bound $c_L = 0$.

Using Lemmas 1 and 2, we show that an optimal policy is of threshold type.

Theorem 1: There is a threshold policy that solves the Bellman optimal equations (6). Specifically, given cost c , the policy has the decision

$$u_c^*(D, e) = \begin{cases} 0, & \text{if } e \leq e_D^*(\beta, c), \\ 1, & \text{if } e > e_D^*(\beta, c), \end{cases}$$

with a real-valued threshold $e_D^*(\beta, c)$.

Proof 3: From Lemma 2, $u_c^*(D, e) = 1$ for all (D, e) when $c \leq c_L$, and $u_c^*(D, e) = 0$ for all (D, e) when $c \geq c_H$. Consider the case when $c_L < c < c_H$. From Remark 3, we have $V_{\beta,c}(D, 0; u = 0) < V_{\beta,c}(D, 0; u = 1)$ and $V_{\beta,c}(D, 1; u = 0) > V_{\beta,c}(D, 1; u = 1)$. For fixed cost c , $V_{\beta,c}(D, e; u = 1)$ is linear in e from (6) and $V_{\beta,c}(D, e; u = 0)$ is concave in e from Lemma 1, and both of them are continuous in e from (18). Thus, there exists exactly one point $e \in [0, 1)$ such that $V_{\beta,c}(D, e; u = 0) = V_{\beta,c}(D, e; u = 1)$. Further, for $e \in [0, e^*]$, we have $V_{\beta,c}(D, e; u = 0) \leq V_{\beta,c}(D, e; u = 1)$, and for $e \in (e^*, 1]$, we have $V_{\beta,c}(D, e; u = 0) > V_{\beta,c}(D, e; u = 1)$.

Remark 6 (MAP): Theorem 1 also holds for the MAP estimator from Lemma 1 and the existence of range $[c_L, c_H]$ with $c_L = 0, c_H < \infty$.

B. Special Case of Symmetric Source

In this subsection, we focus on symmetric scenarios with $p_{01} = p_{10} = p$, under which we have $\mathcal{T}_0^k(e) = \mathcal{T}_1^k(e)$ and $V_{\beta,c}(0, e) = V_{\beta,c}(1, e)$ for all e, k , and thus we can simplify the state representation of the Markov chain with only e . We rewrite (6) by dropping D as

$$\begin{aligned} V_{\beta,c}(e) &= \min\{V_{\beta,c}(e; u = 0), V_{\beta,c}(e; u = 1)\}, \\ V_{\beta,c}(e; u = 0) &= e + \beta V_{\beta,c}(\mathcal{T}(e)), \\ V_{\beta,c}(e; u = 1) &= c + \beta V_{\beta,c}(p). \end{aligned} \quad (19)$$

In this case, we can show the existence of stationary optimal policy for the (undiscounted) long-run average penalty minimization problem. We use the following Dutta's theorem.

Theorem 2 (Dutta [36]): Let S be the state space and suppose that the dynamic programming problem is value bounded, i.e., there exist a state $z \in S$, a function $M(\cdot) : S \rightarrow \mathbb{R}$, and a constant $M \in \mathbb{R}$ such that

$$-\infty < M(s) < V_{\beta}(s) - V_{\beta}(z) < M < \infty \text{ for all } s \in S.$$

Then we have the following results.

- 1) There exists $\lambda \in \mathbb{R}$ such that $\lambda = \lim_{\beta \rightarrow 1} (1 - \beta)V_{\beta}(s)$ for all $s \in S$.
- 2) There exists a stationary optimal policy for the long-run expected average problem.
- 3) Let $\pi_{\beta}(\cdot)$ be a stationary optimal policy for the discounted penalty minimization problem (5). If $\pi_{\beta}(\cdot) \rightarrow \pi$ pointwise as $\beta \rightarrow 1$, then π is a stationary optimal policy for the long-run average problem (4).

The following lemma shows that Theorem 2 is applicable to our case.

Lemma 3: The single symmetric source network problem with communication cost under the discounted total penalty criterion is value bounded. Specifically, we have

$$|V_{\beta,c}(e) - V_{\beta,c}(e')| \leq \frac{1}{1-\beta+2\beta p} \text{ for all } e, e' \in [0, 1].$$

Proof 4: From Lemma 1 and the proof of Lemma 2, we have three cases for $V_{\beta,c}(e; u = 0)$ and $V_{\beta,c}(e; u = 1)$ as shown in Fig. 2. We mark $V_{\beta,c}(e) = \min\{V_{\beta,c}(e; u = 0), V_{\beta,c}(e; u = 1)\}$ as a solid line. When $c < c_L$, we have $|V_{\beta,c}(e) - V_{\beta,c}(e')| = 0$ for all $e, e' \in [0, 1]$. When $c_L \leq c < c_H$, from (6) and (9), we have $|V_{\beta,c}(e) - V_{\beta,c}(e')| \leq |V_{\beta,c}(1; u = 1) - V_{\beta,c}(0; u = 0)| = c < c_H = \frac{1}{1-\beta+2\beta p}$ for all $e, e' \in [0, 1]$. If $c \geq c_H$, then $|V_{\beta,c}(e) - V_{\beta,c}(e')| \leq |V_{\beta,c}(1; u = 0) - V_{\beta,c}(0; u = 0)| = \frac{1}{1-\beta+2\beta p}$ for all $e, e' \in [0, 1]$ from (18).

Remark 7 (MAP): Under the MAP estimator, the value boundedness can be also shown similarly, since $c_L = 0$ and $c_H < \infty$.

From Theorem 2 and Lemma 3, there exists a stationary optimal policy that minimizes the long-term expected average penalty for a symmetric Markovian source. Further, if we can obtain the closed-form expressions of value functions $V_{\beta,c}(p)$ and threshold $e^*(\beta, c)$, we may directly obtain the solution to the expected average penalty problem. Since it is difficulty to obtain the closed-form $e^*(\beta, c)$, we show that an optimal policy is of threshold type by using the existence of an optimal policy of Theorem 2, and then obtain the closed-form

expression of the minimum expected average penalty λ over infinite horizon. In the following, we study the problem (4) of the long-term expected average penalty and omit the discount factor β .

Given the existence of a stationary optimal policy, we have the following Bellman equations as

$$\lambda + \phi_c(e) = \min\{e + \phi_c(\mathcal{T}(e)), c + \phi_c(p)\}, \quad (20)$$

where $\phi_c(\cdot)$ denotes the cost-to-go function that is the differential penalty occurred by the transient change [30]. It is not difficult to show that $\phi_c(\cdot)$ is concave in e for fixed c , and $\phi_c(e)$ is nondecreasing and concave in c for fixed e : we construct value iteration for the n th iteration cost-to-go function $\phi_{n,c}(e)$, and use the techniques in the proof of Lemma 1 along with convergence $\phi_{n,c}(e) \rightarrow \phi_c(e)$ as $n \rightarrow \infty$. Also, we can obtain

$$c_L = 0, \quad c_H = \frac{1}{2p},$$

which can be obtained by either taking the limit $\beta \rightarrow 1$ in Lemma 2, or directly solving (20) as in Lemma 2. From the concavity of $\phi_c(\cdot)$ and the existence of boundary costs c_L and c_H , it can be shown that an optimal policy that solves (20) is of threshold type, following the same line of analysis of Theorem 1. We omit the proof, and using the fact that an optimal policy is of threshold type. We now obtain the minimum expected average penalty λ .

Lemma 4: When a source follows a symmetric Markov process, we have the minimum expected average penalty λ as

- When $p \leq 1/2$, we have

$$\lambda = \begin{cases} c, & \text{if } e^*(c) < p, \\ \frac{1}{\tilde{K}+1}(c + h_p(p, \tilde{K})), & \text{if } p \leq e^*(c) < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } e^*(c) \geq \frac{1}{2}, \end{cases} \quad (21)$$

where $\tilde{K} = \left\lfloor \log_{1-2p} \frac{1-2e^*(c)}{1-2p} \right\rfloor + 1$, and $h_p(x, y) = \frac{y}{2} - \frac{(1-2x)(1-(1-2p)^y)}{4p}$.

- When $p > 1/2$, we have

$$\lambda = \begin{cases} c, & \text{if } e^*(c) < p, \\ \frac{1}{2}, & \text{if } e^*(c) \geq p. \end{cases} \quad (22)$$

Proof 5: Since λ does not depend on the initial state [30], we solve (20) with initial belief state p under the stationary optimal policy of threshold type with threshold $e^*(c)$.

Suppose that $p \leq 1/2$. Then $\mathcal{T}^K(p) = (1-p)\mathcal{T}^{K-1}(p) + (1-\mathcal{T}^{K-1}(p))p$ is a monotonically increasing function of K converging to $\frac{1}{2}$ from Remark 4.

- If $e^*(c) < p$, the policy will take action $u = 1$ at $e = p$. From (20), we have $\lambda + \phi_c(p) = c + \phi_c(p)$, i.e., $\lambda = c$.
- If $e^*(c) \geq 1/2$, the policy will take action $u = 0$ at $e = \mathcal{T}^k(p)$ for all k , since $\mathcal{T}^k(p) < \frac{1}{2}$. Then we have $\lambda + \phi_c(p) = p + \phi_c(\mathcal{T}(p))$ and thus, $\phi_c(\mathcal{T}(p)) - \phi_c(p) = \lambda - p$. Similarly, we have $\phi_c(\mathcal{T}^k(p)) - \phi_c(\mathcal{T}^{k-1}(p)) = \lambda - \mathcal{T}^{k-1}(p)$ for all k . Letting $\phi_c(p) = 0$, we can obtain

$$\begin{aligned} \phi_c(\mathcal{T}^K(p)) &= K\lambda - \sum_{k=0}^{K-1} \mathcal{T}^k(p) \\ &= K\lambda - \frac{K}{2} + \frac{(1-2p)(1-(1-2p)^K)}{4p}, \end{aligned}$$

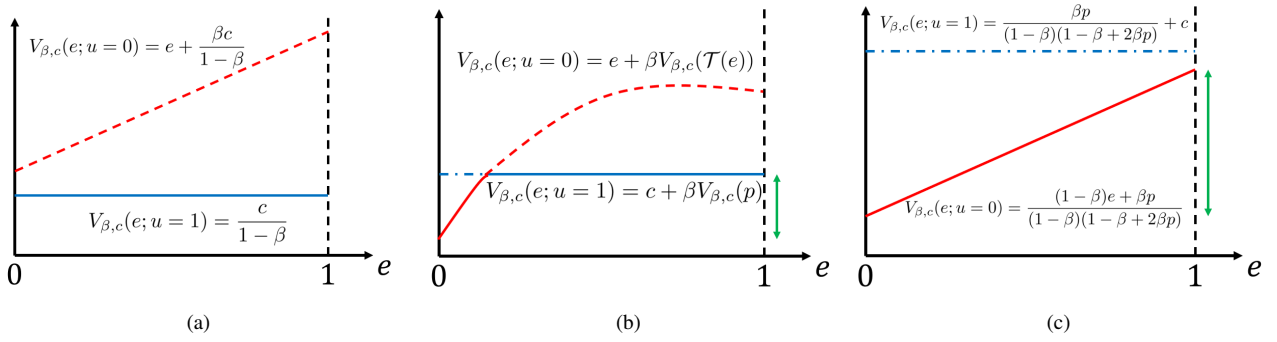


Fig. 2. Value function $V_{\beta,c}(e; u=0)$ and $V_{\beta,c}(e; u=1)$. Solid lines denote their minimum $V_{\beta,c}(e)$: (a) When $c \leq c_L$, (b) when $c_L < c < c_H$, and (c) when $c \geq c_H$.

where the last equation is from (14). By dividing both sides by K and letting $K \rightarrow \infty$, we can obtain $\lambda = 1/2$ since $\phi_c(\mathcal{T}^K(p))$ converges to finite $\phi_c(\frac{1}{2})$.

- If $p \leq e^*(c) < 1/2$, then from the monotonicity of $\mathcal{T}^k(p)$, there exists an integer $\tilde{K} \geq 0$ such that $\mathcal{T}^{\tilde{K}}(p) \leq e^*(c)$ and $\mathcal{T}^{\tilde{K}+1}(p) > e^*(c)$. From $\mathcal{T}^k(p) = \frac{1-(1-2p)^{k+1}}{2}$ and the integer constraint, we have $\tilde{K} = \left\lceil \log_{1-2p} \frac{1-2e^*(c)}{1-2p} \right\rceil + 1$. In this case, we can obtain

$$\phi_c(\mathcal{T}^{\tilde{K}+1}(p)) = \tilde{K}\lambda - \frac{\tilde{K}}{2} + \frac{(1-2p)(1-(1-2p)^{\tilde{K}})}{4p}. \quad (23)$$

Since $\mathcal{T}^{\tilde{K}+1}(p) > e^*(c)$ implies $\lambda + \phi_c(\mathcal{T}^{\tilde{K}+1}(p)) = c + \phi_c(p)$, we combine it with (23). Letting $\phi_c(p) = 0$, we obtain (21).

Now suppose that $p > 1/2$. Then, $\mathcal{T}^k(p)$ converges to $\frac{1}{2}$ as $k \rightarrow \infty$ while it alternates around $1/2$ from Remark 4. We note that $\mathcal{T}^k(p) < p$ for all k . Thus, if $e^*(c) < p$, the optimal policy will take action $u = 1$ at $e = p$ and we have $\lambda = c$. If $e^*(c) \geq p$, the policy will take action $u = 0$ at $e = \mathcal{T}^k(p)$ for all k , and we have $\lambda = 1/2$ as in the case that $p \leq 1/2$ and $e^*(c) \geq 1/2$.

Remark 8 (MAP): Suppose that $p_{01} = p_{10} = 1 - p$. Then, for the MAP estimator, the one-step evolution of the belief error e is $\min\{ep + (1-e)(1-p), 1-ep - (1-e)(1-p)\} = \min\{1-e(1-p) - (1-e)p, e(1-p) + (1-e)p\}$, which is the same as the one-step evolution of e for the Markov chain with $p_{01} = p_{10} = p$. By induction, it can be shown that k -step evolution of the belief error e for the transition probability p is the same as k -step evolution of e for $1-p$.

Remark 9 (MAP): Note that, for the MAP estimator, the error range is $[0, 0.5]$ since if $(D, e), e > 0.5$ then the MAP estimator uses $(\bar{D}, 1-e)$. Thus, when $p \leq \frac{1}{2}$, the minimum expected average penalty λ_{MAP} for the MAP estimator is the same as the first and second cases in (21) since $\mathcal{T}^k(e) = \min\{\mathcal{T}^k(e), 1-\mathcal{T}^k(e)\}$ for all $k \geq 0$ and $e \in [0, 0.5]$. When $p > 1/2$, from Remark 8, we have λ_{MAP} as in (21) using $q = 1-p$ instead of p .

IV. INDEX POLICY FOR MULTIPLE SYMMETRIC SOURCES

In this section, we show that the problem of single symmetric source with communication cost is *indexable*, and obtain the closed-form index, which can be considered as the amount

of communication cost that the base station is willing to pay for the information update. For multiple sources, we can minimize the information mismatch by updating the source with the highest index first.

We start from the definition of indexability. Let $\mathcal{P}(c)$ be the set of belief state e for which it is optimal to not update the state information, i.e., $\mathcal{P}(c) = \{e \in [0, 1] : u_c^*(e) = 0\}$.

Definition 1 (Indexability [21]): The problem of single source with communication cost is indexable if $\mathcal{P}(c)$ monotonically increases from \emptyset to the entire space $[0, 1]$ as the communication cost c increases from $-\infty$ to ∞ . The extended problem to multiple sources is indexable if each problem of source i is indexable for all sources i .

In our problem of single symmetric source, the indexability is equivalent to the monotonic increase of threshold $e^*(c)$ for $c \in (c_L, c_H]$. The following lemma shows a sufficient condition for the indexability of our single-source problem.

Lemma 5: Threshold $e^*(c)$ monotonically increases with respect to $c \in (c_L, c_H]$ if

$$\frac{\partial \phi_c(e; u=1)}{\partial c} \Big|_{e=e^*(c)} > \frac{\partial \phi_c(e; u=0)}{\partial c} \Big|_{e=e^*(c)}, \quad (24)$$

where $\phi_c(e; u=0) = e + \phi_c(\mathcal{T}(e))$ and $\phi_c(e; u=1) = c + \phi_c(p)$.

Proof 6: We prove by contraposition. Suppose that there exists a $\tilde{c} \in (c_L, c_H]$ such that $e^*(c)$ is decreasing at $c = \tilde{c}$. Then, there exists a $\delta > 0$ such that for all $\epsilon \in [0, \delta]$, $e^*(\tilde{c} + \epsilon) \leq e^*(\tilde{c})$. Thus, given cost $\tilde{c} + \epsilon$, the optimal decision at $e^*(\tilde{c})$ is $u = 1$, i.e.,

$$\phi_{\tilde{c}+\epsilon}(e^*(\tilde{c}); u=1) \leq \phi_{\tilde{c}+\epsilon}(e^*(\tilde{c}); u=0). \quad (25)$$

Similarly, we have

$$\phi_{\tilde{c}}(e^*(\tilde{c}); u=1) = \phi_{\tilde{c}}(e^*(\tilde{c}); u=0). \quad (26)$$

Combining (25) and (26), and dividing by ϵ , we have

$$\frac{\phi_{\tilde{c}+\epsilon}(e^*(\tilde{c}); u=1) - \phi_{\tilde{c}}(e^*(\tilde{c}); u=1)}{\epsilon} \leq \frac{\phi_{\tilde{c}+\epsilon}(e^*(\tilde{c}); u=0) - \phi_{\tilde{c}}(e^*(\tilde{c}); u=0)}{\epsilon},$$

which implies $\frac{\partial \phi_c(e; u=1)}{\partial c} \Big|_{e=e^*(\tilde{c})} \leq \frac{\partial \phi_c(e; u=0)}{\partial c} \Big|_{e=e^*(\tilde{c})}$. By contraposition, the lemma holds.

We now show the indexability of the single source network with communication cost using Lemma 5.

Theorem 3: For a source that follows a two-state symmetric Markov process, the single source problem with communication cost is indexable.

Proof 7: To show that $\mathcal{P}(c)$ monotonically increases in c , it suffices to show (24) for all $c \in (c_L, c_H]$. From (20), we have

$$\left. \frac{\partial \phi_c(e; u=1)}{\partial c} \right|_{e=e^*(c)} = 1 + \frac{\partial \phi_c(p)}{\partial c}, \quad (27)$$

$$\left. \frac{\partial \phi_c(e; u=0)}{\partial c} \right|_{e=e^*(c)} = \frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c}. \quad (28)$$

We are going to use a couple of intermediate results: (i) If $e > e^*(c)$, then the optimal policy has action $u = 1$, and we have $\lambda + \phi_c(e) = c + \phi_c(p)$ from (20). By taking the partial derivative with respect to c , we have

$$\frac{\partial \phi_c(e)}{\partial c} = 1 + \frac{\partial \phi_c(p)}{\partial c} - \frac{\partial \lambda}{\partial c}. \quad (29)$$

(ii) If $\mathcal{T}^k(e) \leq e^*(c)$ for all $k \geq 0$, the optimal action is $u = 0$ at $\mathcal{T}^k(e)$. Then, again from (20), we have $\lambda + \phi_c(\mathcal{T}^{k-1}(e)) = \mathcal{T}^{k-1}(e) + \phi_c(\mathcal{T}^k(e))$ for all k . Letting

$$h_p(x, y) = \frac{y}{2} - \frac{(1-2x)(1-(1-2p)^y)}{4p},$$

we can obtain

$$\begin{aligned} \phi_c(e) &= \sum_{k=0}^{K-1} \mathcal{T}^k(e) - K\lambda + \phi_c(\mathcal{T}^K(e)) \\ &= \frac{K}{2} - \frac{(1-2e)(1-(1-2p)^K)}{4p} - K\lambda + \phi_c(\mathcal{T}^K(e)). \\ &= h_p(e, K) - K\lambda + \phi_c(\mathcal{T}^K(e)). \end{aligned}$$

Note that $\phi_c(\mathcal{T}^K(e)) \rightarrow \phi_c(1/2)$ as $K \rightarrow \infty$. Dividing the equation by K and letting $K \rightarrow \infty$, we obtain $\lambda = 1/2$ (i.e., λ does not depend on c), and using this, we obtain $\phi_c(e) = \phi_c(1/2) - (1-2e)/4p$ when $K \rightarrow \infty$. This results in

$$\frac{\partial \phi_c(e)}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}. \quad (30)$$

We now show (24) using (27) – (30). Suppose that $p \leq 1/2$.

- If $e^*(c) < p$, then we have $\mathcal{T}(e^*(c)) > e^*(c)$ from Remark 4, and $\lambda = c$ from Lemma 4. Hence, from (29) with $e = \mathcal{T}(e^*(c)) > e^*(c)$, we have $\frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c} = \frac{\partial \phi_c(p)}{\partial c}$. From this and (27), (28), we have

$$\begin{aligned} \left. \frac{\partial \phi_c(e; u=1)}{\partial c} \right|_{e=e^*(c)} &= 1 + \frac{\partial \phi_c(p)}{\partial c} \\ &> \frac{\partial \phi_c(p)}{\partial c} = \left. \frac{\partial \phi_c(e; u=0)}{\partial c} \right|_{e=e^*(c)}. \end{aligned} \quad (31)$$

- If $p \leq e^*(c) < 1/2$, then we have $\mathcal{T}(e^*(c)) > e^*(c)$ and $\lambda = \frac{1}{K+1} \left[c + h_p(p, \tilde{K}) \right]$. Thus, from (29), we have

$$\begin{aligned} \left. \frac{\partial \phi_c(e; u=1)}{\partial c} \right|_{e=e^*(c)} &= 1 + \frac{\partial \phi_c(p)}{\partial c} \\ &> 1 + \frac{\partial \phi_c(p)}{\partial c} - \frac{1}{K+1} = \left. \frac{\partial \phi_c(e; u=0)}{\partial c} \right|_{e=e^*(c)}, \end{aligned}$$

where $\tilde{K} = \left\lfloor \log_{1-2p} \frac{1-2e^*(c)}{1-2p} \right\rfloor + 1 > 0$.

- If $e^*(c) \geq 1/2$, then we have $\mathcal{T}^k(e^*(c)) \leq e^*(c)$ and $\mathcal{T}^k(p) \leq e^*(c)$ for all k from Remark 4. Thus, we have $\frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}$ and $\frac{\partial \phi_c(p)}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}$ from (30), and

$$\begin{aligned} \left. \frac{\partial \phi_c(e; u=1)}{\partial c} \right|_{e=e^*(c)} &= 1 + \frac{\partial \phi_c(p)}{\partial c} \\ &> \frac{\partial \phi_c(p)}{\partial c} = \left. \frac{\partial \phi_c(e; u=0)}{\partial c} \right|_{e=e^*(c)}. \end{aligned} \quad (32)$$

Now suppose that $p > 1/2$.

- If $e^*(c) < 1/2$, then we have $\mathcal{T}(e^*(c)) > 1/2$ since $\mathcal{T}(\cdot)$ alternates around $1/2$, and $\lambda = c$ from Lemma 4. Hence, we have $\frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c} = \frac{\partial \phi_c(p)}{\partial c}$ from (29), and the sufficient condition holds as in (31).
- If $e^*(c) \geq p$, then we have $e^*(c) \geq \mathcal{T}^k(e^*(c))$ and $e^*(c) \geq \mathcal{T}^k(p)$ for all k from Remark 4. Thus, (30) holds for all e , we have $\frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c} = \frac{\partial \phi_c(p)}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}$ which satisfies the sufficient condition as in (32).
- If $1/2 \leq e^*(c) < p$, then the optimal action is $u = 1$ at $e = p$, and $u = 0$ at $e = 1/2$. Thus, we have

$$\lambda + \phi_c(p) = c + \phi_c(p), \text{ and } \lambda + \phi_c(\frac{1}{2}) = \frac{1}{2} + \phi_c(\frac{1}{2}).$$

Solving the above equations, we have $\lambda = c = 1/2$, which implies that if $c = 1/2$, then $u = 1$ for every time slot and $u = 0$ for every time slot are optimal solutions and give the time-average penalty of $1/2$. Further, it implies that if $c \neq 1/2$, then the case of $(1/2 \leq e^*(c) < p)$ does not occur.

Hence, we have $\lambda + \phi_c(1/2) = c + \phi_c(p)$. By taking partial derivative with respect to c , we obtain $\frac{\partial \phi_c(p)}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}$. Further, since $\mathcal{T}^k(e^*(c)) \leq e^*(c)$ for all k , we have $\frac{\partial \phi_c(\mathcal{T}(e^*(c)))}{\partial c} = \frac{\partial \phi_c(\frac{1}{2})}{\partial c}$ from (30). Thus, the sufficient condition is satisfied as in (32).

Given the indexability of the problem, we compute the index as follows.

Definition 2 (Whittle's index [21]): Given indexability, Whittle's index denoted by $W(e)$ is the infimum transmission cost c that makes both actions (i.e., $u = 0, 1$) equally desirable at e .

Theorem 4: Consider the error probability minimization problem with multiple symmetric Markovian sources. The closed-form expression of Whittle's index $W_i(e_i)$ for source i is given as follows.

- When $p^{(i)} \leq 1/2$:

$$W_i(e_i) = \begin{cases} e_i, & \text{if } e_i < p^{(i)}, \\ (\tilde{K} + 1)e_i - h_{p^{(i)}}(p^{(i)}, \tilde{K}), & \text{if } p^{(i)} \leq e_i < \frac{1}{2}, \\ \frac{e_i}{2p^{(i)}}, & \text{if } e_i \geq \frac{1}{2}, \end{cases} \quad (33)$$

where $\tilde{K} = \left\lfloor \log_{1-2p^{(i)}} \frac{1-2e_i}{1-2p^{(i)}} \right\rfloor + 1$.

- When $p^{(i)} > 1/2$:

$$W_i(e_i) = \begin{cases} e_i, & \text{if } e_i < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \frac{1}{2} \leq e_i < p^{(i)}, \\ \frac{e_i}{2p^{(i)}}, & \text{if } e_i \geq p^{(i)}. \end{cases} \quad (34)$$

Proof 8: We drop i for notational convenience. From the definition of Whittle index, for given e , we find cost $W(e)$ such that the two actions ($u = 0$ and 1) are indifferent under the threshold policy π . Let $\pi(e, c)$ denote the decision of policy π with cost c : $\pi(e', W(e)) = 0$ for all $e' \leq e$, and $\pi(e', W(e)) = 1$ for all $e' > e$ by definition. First, suppose that $p \leq \frac{1}{2}$.

- For $e \in [0, p)$, we have $e < \mathcal{T}(e)$ since $\mathcal{T}(\cdot)$ monotonically increases toward $1/2$. Thus, we have $\pi(\mathcal{T}(e), W(e)) = 1$, which leads to $\lambda + \phi_c(\mathcal{T}(e)) =$

$c + \phi_c(p)$ from (20) with $c = W(e)$. Similarly, we have $\pi(p, W(e)) = 1$ and $\lambda + \phi_c(p) = c + \phi_c(p)$. Also, since the policy is indifferent at e , we have $e + \phi_c(\mathcal{T}(e)) = c + \phi_c(p)$. By solving, the three equations, we can obtain that $c = W(e) = e$.

- For $e \in [p, 1/2]$, we have $e < \mathcal{T}(e) < 1/2$, and thus $\pi(\mathcal{T}(e), W(e)) = 1$, which leads to $\lambda + \phi_c(\mathcal{T}(e)) = c + \phi_c(p)$ from (20) with $c = W(e)$. By letting $\phi_c(p) = 0$ with $c = W(e)$, we have $\lambda + \phi_c(\mathcal{T}(e)) = c$. Also from (21), we have $\lambda = \frac{1}{K+1} \left[c + h_p(p, \tilde{K}) \right]$, where $\tilde{K} = \left\lfloor \log_{1-2p} \frac{1-2e}{1-2p} \right\rfloor + 1$. The indifference of the policy at e results in $e + \phi_c(\mathcal{T}(e)) = c$. Combining with $\lambda + \phi_c(\mathcal{T}(e)) = c$, we have $e = \lambda$ and thus (33).
- For $e \geq 1/2$, we have $1/2 \leq \mathcal{T}(e) \leq e$ since $\mathcal{T}(\cdot)$ monotonically decreases towards $1/2$ from Remark 4, and thus $\pi(\mathcal{T}(e), W(e)) = 0$, which leads to $\lambda + \phi_c(\mathcal{T}(e)) = \mathcal{T}(e) + \phi_c(\mathcal{T}^2(e))$ with $c = W(e)$. Also, from $p \leq 1/2$ and $\mathcal{T}(p) \leq 1/2 < W(e)$, we have $\pi(\mathcal{T}(p), W(e)) = 0$, implying $\lambda + \phi_c(p) = p + \phi_c(\mathcal{T}(p))$. Combining these with the indifference equation $e + \phi_c(\mathcal{T}(e)) = c + \phi_c(p)$, we obtain

$$e + \mathcal{T}(e) + \phi_c(\mathcal{T}^2(e)) = c + p + \phi_c(\mathcal{T}(p)). \quad (35)$$

We note that $\mathcal{T}^k(e) \leq e$ and $\mathcal{T}^k(p) \leq e$ for all $k \geq 0$, $p \leq 1/2$, $e \geq 1/2$. Then, from $\pi(\mathcal{T}^k(e), W(e)) = 0$ and $\pi(\mathcal{T}^k(p), W(e)) = 0$, we can obtain

$$\phi_c(\mathcal{T}^{k+1}(e)) - \phi_c(\mathcal{T}^k(e)) = \lambda - \mathcal{T}^k(e), \quad (36)$$

$$\phi_c(\mathcal{T}^{k+1}(p)) - \phi_c(\mathcal{T}^k(p)) = \lambda - \mathcal{T}^k(p), \quad (37)$$

respectively. Further, from (36) and (37), we can obtain

$$\begin{aligned} \phi_c(\mathcal{T}^{K+1}(e)) + \sum_{k=2}^K \mathcal{T}^k(e) &= \phi_c(\mathcal{T}^2(e)) + (K-1)\lambda, \\ \phi_c(\mathcal{T}^K(p)) + \sum_{k=1}^{K-1} \mathcal{T}^k(p) &= \phi_c(\mathcal{T}(p)) + (K-1)\lambda, \end{aligned}$$

respectively. By adding $e + \mathcal{T}(e)$ to the both sides of the first equation and $c + p$ to the both sides of the second equation, and by combining them with (35), we obtain

$$\sum_{k=0}^K \mathcal{T}^k(e) + \phi_c(\mathcal{T}^{K+1}(e)) = c + \sum_{k=0}^{K-1} \mathcal{T}^k(p) + \phi_c(\mathcal{T}^K(p)),$$

where

$$\begin{aligned} \sum_{k=0}^K \mathcal{T}^k(e) &= h_p(e, K+1), \\ \sum_{k=0}^{K-1} \mathcal{T}^k(p) &= h_p(p, K). \end{aligned} \quad (38)$$

By letting $K \rightarrow \infty$, we have $\phi_c(\mathcal{T}^{K+1}(e')) \rightarrow \phi_c(1/2)$ for any e' . From $p \in [0, 1]$, we can obtain $c = W(e) = e/2p$.

Now we consider the other case of $p > 1/2$.

- For $e \in [0, 1/2]$ and $e \in [p, 1]$, we can follow the same line of the analysis as the case $p \leq 1/2$ to obtain (34).
- For $e \in [1/2, p]$, we have $\pi(p, W(e)) = 1$, which leads to $\lambda + \phi_c(p) = c + \phi_c(p)$, i.e., $\lambda = c$, from (20). Also since $\mathcal{T}(e) < e$, we have $\pi(\mathcal{T}(e), W(e)) = 0$, and thus $\lambda + \phi_c(e) = e + \phi_c(\mathcal{T}(e))$. Letting $\phi_c(e) = 0$, and combining two equations, we have $\phi_c(\mathcal{T}(e)) = c - e$. We generalize this by using $\mathcal{T}^k(e) < e$ for all k . From $\pi(\mathcal{T}^k(e), W(e)) = 0$, we have $\lambda + \phi_c(\mathcal{T}^k(e)) =$

$\mathcal{T}^k(e) + \phi_c(\mathcal{T}^{k+1}(e))$ for all k with $c = W(e)$. Then using $\lambda = c$ and from (38),

$$\begin{aligned} \phi_c(\mathcal{T}^k(e)) &= c - \mathcal{T}^{K-1}(e) + \phi_c(\mathcal{T}^{K-1}(e)) \\ &= Kc - \sum_{k=0}^{K-1} \mathcal{T}^k(e) \\ &= Kc - \frac{K}{2} + \frac{(1-2e)(1-(1-2p)^K)}{4p}. \end{aligned}$$

From $\phi_c(\mathcal{T}^K(e)) \rightarrow \phi_c(1/2)$ as $K \rightarrow \infty$, we can obtain $c = 1/2$ by dividing both sides by K and letting $K \rightarrow \infty$.

Therefore, we can assign an index to each symmetric source i as in (33) or (34), depending on the value of p . Since a source with a higher index implies that the threshold policy is willing to take a higher communication cost for an update than the other sources, we can develop an index policy that updates the source with the highest index to minimize the long-run average error sum over all the sources.

Remark 10 (MAP): From Remarks 8 and 9, the indexability for the MAP estimator can be shown. Further, for $p \leq 0.5$, the Whittle's index for the MAP estimator is the same as the first and second cases in (33). For $p > 0.5$, the Whittle's index is the same as (33) by replacing $q = 1 - p$ instead of p .

V. SIMULATION RESULTS

In this section, we show the performance of Whittle's index policy through simulation results. We first compare five policies: dynamic programming (DP) over finite time horizon (for UAH and MAP estimators), Whittle's index policy (for UAH and MAP estimators), and myopic policy. Recall that the myopic policy updates source $i = \arg \max_i e_i(t)$ at each time t . Although the DP is an optimal policy, its high computational complexity prohibits its practical implementation [30]. Thus, in this simulation, we consider two symmetric sources over 1000 time-slots. In particular, the transition probability $p^{(1)}$ of source 1 is fixed at 0.1 or 0.5 while the transition probability $p^{(2)}$ of source 2 varies from 0.1 to 0.9. We recall that the communication channel is a noiseless channel and has no transmission delay.

Fig. 3 shows the average error of five policies with two symmetric sources when $p^{(1)} = 0.1$ (Fig. 3((a))) or $p^{(1)} = 0.5$ (Fig. 3((b))) while $p^{(2)} = \alpha$ varying from 0.1 to 0.9. The results are averaged over 10^4 repetitions. As seen in Fig. 3, the Whittle's index policy achieves near-optimal performance. Further, since the evolution of estimated error under MAP estimator for Markov chains with transition probabilities of p and $1 - p$ is the same, the performance of MAP estimator is symmetric around $\alpha = 0.5$. As seen in Fig. 3((a)), when $p^{(2)} = \alpha \geq 0.5$, the belief error of source 1 is no larger than 0.5. Thus, the myopic policy always choose source 2, and the average error (per source) becomes 0.25. The same holds for $\alpha \leq 0.5$ in Fig. 3((b)). We observed that the performance gap between the Whittle's index policies and the dynamic programming is insignificant, which confirms to the previous results of the optimality of the Whittle's index policy with many sources/arms [10]. Thus, we conjecture that the Whittle's

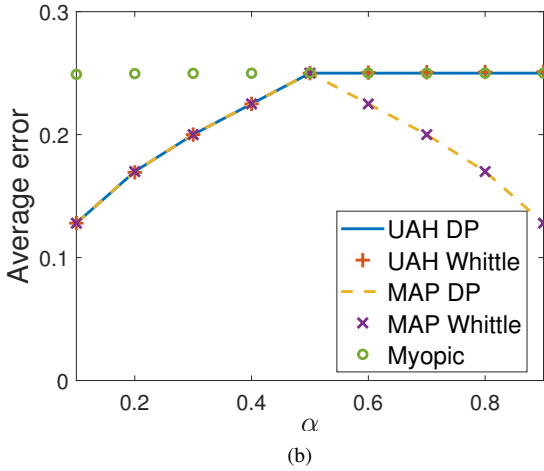
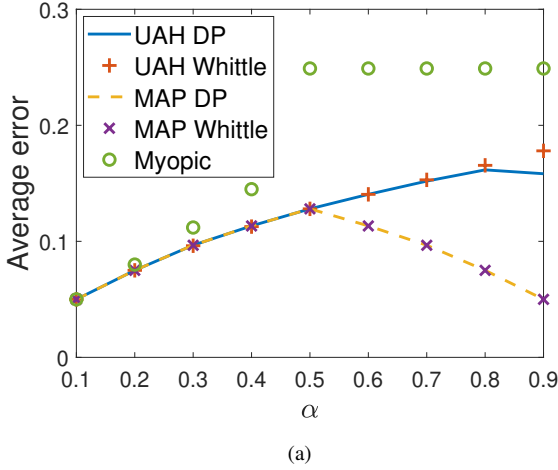


Fig. 3. Average error performance with two sources: (a) $p^{(1)} = 0.1$ and $p^{(2)} = \alpha$ and (b) $p^{(1)} = 0.5$ and $p^{(2)} = \alpha$.

index policy proposed in this paper performs very close to the optimal policy (DP).

Next, we compare the performance of Whittle's index policy and the myopic policy with a larger number of symmetric sources. The number of sources changes from 1 to 10, and the transition probabilities $p^{(i)} \in (0, 1)$ are chosen uniformly at random. The total time-slot is 10^5 and the result is averaged over 100 repetitions. As seen in Fig. 4, Whittle's index policy outperforms the myopic policy for all the cases.

VI. CONCLUSION

In this paper, we considered a status update system with multiple sources updating a common remote estimation and investigated the centralized scheduling policies to minimize the expected average estimation error over infinite time horizon. We showed that an optimal solution to the problem of single source with communication cost is of threshold type. For multi-source scenario, we modeled the error minimization problem as a restless multi-armed bandits problem and showed that our problem with symmetric sources is indexable. Under indexability, we developed low-complexity algorithm by obtaining a closed-form expression of the Whittle's index.

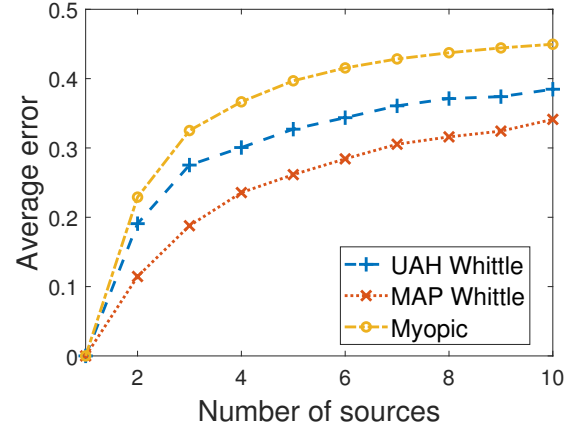


Fig. 4. Average error performance with multiple sources.

Through numerical simulations, we showed that our algorithm is near-optimal by comparing finite-time dynamic programming.

REFERENCES

- [1] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *Proc. IEEE WCSP*, 2011.
- [2] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Found. Trends Netw.*, vol. 12, no. 3, pp. 162–259, 2017.
- [3] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *J. Commun. Netw.*, vol. 21, no. 3, 2019.
- [4] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.
- [5] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," *IEEE/ACM Trans. Netw.*, vol. 28, no. 1, pp. 15–28, 2019.
- [6] C. Joo and A. Eryilmaz, "Wireless scheduling for information freshness and synchrony: Drift-based design and heavy-traffic analysis," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2556–2568, 2018.
- [7] G. Stamatakis, N. Pappas, and A. Traganitis, "Control of status updates for energy harvesting devices that monitor processes with alarms," in *Proc. IEEE GLOBECOM*, 2019.
- [8] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. IEEE ISIT*, 2018.
- [9] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, 2018.
- [10] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "On the optimality of the whittle's index policy for minimizing the age of information," *IEEE Trans. Wireless Commun.*, vol. 26, no. 2, pp. 1263–1277, 2020.
- [11] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Towards an effective age of information: Remote estimation of a markov source," in *Proc. IEEE INFOCOM*, 2018.
- [12] Z. Jiang, S. Zhou, Z. Niu, and C. Yu, "A unified sampling and scheduling approach for status update in multiaccess wireless networks," in *Proc. IEEE INFOCOM*, 2019.
- [13] G. Chen, S. C. Liew, and Y. Shao, "Uncertainty-of-information scheduling: A restless multi-armed bandit framework," *arXiv preprint arXiv:2102.06384*, 2021.
- [14] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. IEEE ISIT*, 2017.
- [15] T. Z. Ornee and Y. Sun, "Sampling for remote estimation through queues: Age of information and beyond," in *Proc. IEEE WiOpt*, 2019.
- [16] X. Gao, E. Akyol, and T. Başar, "Optimal sensor scheduling and remote estimation over an additive noise channel," in *Proc. IEEE ACC*, 2015.

- [17] X. Gao, E. Akyol, and T. Başar, "On remote estimation with multiple communication channels," in *Proc. IEEE ACC*, 2016.
- [18] J. Yun, C. Joo, and A. Eryilmaz, "Optimal real-time monitoring of an information source under communication costs," in *Proc. IEEE CDC*, 2018.
- [19] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.
- [20] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queuing network control," *Math. Operations Research*, vol. 24, no. 2, pp. 293–305, 1999.
- [21] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probability*, vol. 25, no. A, pp. 287–298, 1988.
- [22] N. Akbarzadeh and A. Mahajan, "Restless bandits with controlled restarts: Indexability and computation of whittle index," in *Proc. IEEE CDC*, 2019.
- [23] R. R. Weber and G. Weiss, "On an index policy for restless bandits," *J. Appl. Probability*, vol. 27, no. 3, pp. 637–648, 1990.
- [24] K. Liu, R. Weber, and Q. Zhao, "Indexability and whittle index for restless bandit problems involving reset processes," in *Proc. IEEE CDE*, 2011.
- [25] P. Ansell, K. D. Glazebrook, J. Niño-Mora, and M. O'Keefe, "Whittle's index policy for a multi-class queueing system with convex holding costs," *Mathematical Methods of Operations Research*, vol. 57, no. 1, pp. 21–39, 2003.
- [26] M. Larranaga, U. Ayesta, and I. Verloop, "Stochastic and fluid index policies for resource allocation problems," in *Proc. IEEE INFOCOM*, 2015.
- [27] R. Meshram, D. Manjunath, and A. Gopalan, "On the whittle index for restless multiarmed hidden markov bandits," *IEEE Trans. Automat. Control*, vol. 63, no. 9, pp. 3046–3053, 2018.
- [28] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5547–5567, 2010.
- [29] T. He, A. Anandkumar, and D. Agrawal, "Index-based sampling policies for tracking dynamic networks under sampling constraints," in *Proc. IEEE INFOCOM*, 2011.
- [30] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena scientific Belmont, MA, 2005, vol. 1, no. 3.
- [31] E. J. Sondik, "The optimal control of partially observable markov processes over the infinite horizon: Discounted costs," *Operations research*, vol. 26, no. 2, pp. 282–304, 1978.
- [32] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. IEEE INFOCOM*, 2018.
- [33] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. II*, 3rd ed. Athena Scientific, 2007.
- [34] O. L. d. V. Costa and F. Dufour, "The vanishing discount approach for the average continuous control of piecewise deterministic markov processes," *J. Appl. Probability*, vol. 46, no. 4, pp. 1157–1183, 2009.
- [35] J. R. Norris, *Markov chains*. Cambridge university press, 1998, no. 2.
- [36] P. K. Dutta, "What do discounted optima converge to?: A theory of discount rate asymptotics in economic models," *J. Econ. Theory*, vol. 55, no. 1, pp. 64–94, 1991.



Changhee Joo received the Ph.D. degree from Seoul National University in 2005. He was with Purdue University and The Ohio State University, and then worked at Korea University of Technology and Education (KoreaTech), and Ulsan National Institute of Science and Technology (UNIST). Since 2019, he has been with Korea University. His research interests are in the broad areas of networking, learning, modeling, and optimization. He was a recipient of the IEEE INFOCOM 2008 Best Paper Award, the KICS Haedong Young Scholar Award (2014), the ICTC 2015 Best Paper Award, and the GAMENETS 2018 Best Paper Award. He is a Division Editor of the Journal of Communications and Networks and an editor of IEEE Transactions on Vehicular Technology. He served as an Associate Editor of the *IEEE/ACM Transactions on Networking*, and several primary conferences as a technical program committee member, including IEEE INFOCOM and ACM MOBIHOC.



Sunjung Kang received her M.S. degree from the school of ECE at Ulsan National Institute of Science and Technology (UNIST) in 2018. She is currently a Ph.D. student in the department of ECE at The Ohio State University. Her research interests include the age of information, remote estimation and multi-armed bandits.