Scheduling to Minimize Age of Synchronization in Wireless Broadcast Networks with Random Updates

Haoyue Tang, Jintao Wang, Zihan Tang, Jian Song
Beijing National Research Center for Information Science and Technology (BNRist),
Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China
{thy17@mails, wangjintao@, tangzh14@mails, jsong@}tsinghua.edu.cn

***Abstract*—In this work, a wireless broadcast network with a base station (BS) sending random time-sensitive information updates to multiple users with interference constraints is considered. The Age of Synchronization (AoS), namely the amount of time elapsed since the information stored at the network user becomes desynchronized, is adopted to measure data freshness from the perspective of network users. Compared with the more widely used metric—the Age of Information (AoI), AoS accounts for the freshness of the randomly changing content. We formulate the scheduling problem into a discrete time Markov decision process and approximate the optimal solution through finite state policy iteration. An index based heuristic scheduling policy based on restless multi-arm bandit (RMAB) is provided to reduce computational complexity. Numerical results are presented to demonstrate the performance of the proposed policies.**

I. Introduction

The design of next generation mobile and wireless communication networks are driven partly by the need of mission-critical services like real-time control and the Internet of Things (IoT). Moreover, the proliferation of mobile devices have boosted the needs to enhance the timeliness of services like instant chatting, mobile ads, social updates notifications, etc. In all these scenarios, new stringent requirements are imposed on the freshness of the received data.

To measure the data freshness from the perspective of users, the metric *Age of Information* (AoI) is proposed [1]. It describes the time elapsed since the time-stamp that the freshest information at the receiver has been generated. Another metric called *Age of Synchronization* is proposed [2] to measure the time difference between now and the time-stamp that the freshest data at the receiver becomes desynchronized compared with the source. The two metrics are essentially different. Briefly speaking, AoS only measures the time of status desynchronization and is more appropriate for the scenario in which status changes randomly, e.g., caching systems [2]; AoI also takes the inter packet generation time into account, hence is more appropriate for scenarios that the update source keeps changing all the time. Relative to the more widely used AoI metric, the AoS metric accounts for whether the process that is being tracked has actually changed when updates appear randomly.

Previous work on optimizing data freshness from the perspective of users focus mainly on minimizing AoI. Centralized scheduling algorithm to optimize AoI performance in a single-hop network is first studied in [3]. Distributed scheduling policy to optimize AoI performance under general interference constraint is considered in [4]. The above research are based on models that update packets can be generated at will. For networks with random update sources, scheduling to optimize data freshness performance are considered in various settings [5]–[7]. However, these work all assume error free transmissions and no transmission randomness is taken into account. Moreover, the AoS metric has not received so much attention and so it's less well understood.

In this work, we aim at designing scheduling policies to minimize the expected AoS of an unreliable wireless broadcast network with random information updates. Both update packet generation and transmission randomness is considered in our model. We formulate the problem into a Markov decision process (MDP) with countable infinite state space and approximate the optimum solution through finite state policy iteration. To overcome the computational load by the approximated MDP solution, we propose an index-based heuristic algorithm based on restless multi-arm bandit that can achieve compatible performance compared with MDP.

The remainder of this paper is organized as follows. The network model and the two metrics, age of information and age of synchronization are introduced and compared in Sec. II. In Sec. III, we reformulate the problem into a Markov decision process (MDP) and approximate the optimal MDP solution with truncated policy iteration. In Sec. IV, we propose an index-based heuristic algorithm based on restless multi-arm bandit. The scheduling strategies are evaluated through simulations in Sec. V. The paper is concluded in Sec. VI.

Notations: Vectors are written in boldface letters. The probability of event $\mathcal{A}$ given condition $\mathcal{B}$ is denoted as $\Pr(\mathcal{A}|\mathcal{B})$, the expectation with regard to random variable X is denoted as $\mathbb{E}_X[\cdot]$.

II. System Overview

A. Network Model

We consider a wireless broadcast network with a base station (BS) broadcasting time sensitive information of N random update sources to N users, as shown in Fig. 1. Each user is interested in the information updates from the corresponding source, i.e., user n is only interested in updates from source n, $n = 1, 2, \cdots, N$. Let the time be slotted, the update of source n appears independently and identically with probability λ_n in each time slot. Let the indicator function

$\Lambda_n(t) \in \{0,1\}$ denote whether an update of source n happens during slot t. If $\Lambda_n(t)=1$, then the corresponding information update packet arrives at BS by the end of time slot t. In this work, we focus on the scenario where the BS keeps one snapshot for each source, i.e., the old update information will be kept in BS until a new update arrives.

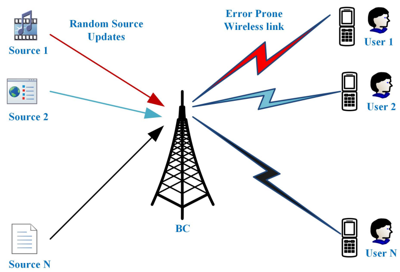

Fig. 1. A broadcast network with the BS sending time sensitive information updates to network users.

At the beginning of each time slot, the BS schedules to broadcast information updates over error-prone wireless link. Here we use indicator function $u_n(t) \in \{0,1\}$ to denote scheduling action. If $u_n(t)=0$, then the corresponding update message of source n is not selected for transmission during time t. If $u_n(t)=1$, then update from source n is scheduled to be transmitted to user n. Assume that packet erasure is a memoryless Bernoulli process and user n has a fixed channel characterized by the Bernoulli packet success probability p_n. Then update packet will be successfully received at the end of time slot t with probability p_n. Due to the limited communication resources and wireless interference constraint, the BS can only send one update to a single user at each time slot, which imposes the following constraint on scheduling decisions:

$$\sum_{n=1}^{N} u_n(t) \leq 1. \tag{1}$$

B. Age of Information and Age of Synchronization

In this part, first we briefly review and compare the two freshness metrics AoI and AoS. Then we introduce the dynamics of AoS. To demonstrate the concept of AoI and AoS, let us consider a single-source-discrete-time scenario as an example. In the following part, suppose the i^{th} packet is generated during slot g_i. If it is scheduled to be transmitted to the user at the beginning of slot r_i, then it will be received at the end of slot r_i if the transmission succeeds.

The AoI measures the time elapsed since the generation time-stamp of the newest update received by the user [1]. Let $q(t) = \max_{i\in\mathbb{N}^+}\{i | r_i \leq t\}$ be the index of the latest update at time t from the perspective of the receiver, the AoI at the beginning of slot t is defined as follows:

$$h(t) = t - g_{q(t)}. \tag{2}$$

The AoS describes the gap of current time and the earliest time that the remote source gets an update since the last refresh of the local copy has been received. Notice that $q(t)+1$ is the index of the earliest update information since the last refresh of the receiver, then the AoS at the beginning of slot t is defined to be:

$$s(t) = (t - g_{q(t)+1})^+, \tag{3}$$

where function $(\cdot)^+ = \max\{0,\cdot\}$. According to the definition, if no new update arrives after the generation time-stamp of the latest refresh of the user, i.e., $g_{q(t)+1} \geq t$, then $s(t)=0$. The sample paths of AoI and AoS of a source are depicted in Fig. 2. From the figure, we can see that AoS remains zero until a new fresh update arrives, i.e., the information at the receiver becomes desynchronized with the source, while AoI keeps accumulating if no update has been received. The difference between AoI and AoS is the reference object. The AoS measures data freshness compared to the source with random updates itself, while AoI takes the staleness of inter packet generation time into account. Since source updates appear randomly, and the source is assumed to remain unchanged between two updates, AoS accounts for the whether the process that is being tracked has actually changed.

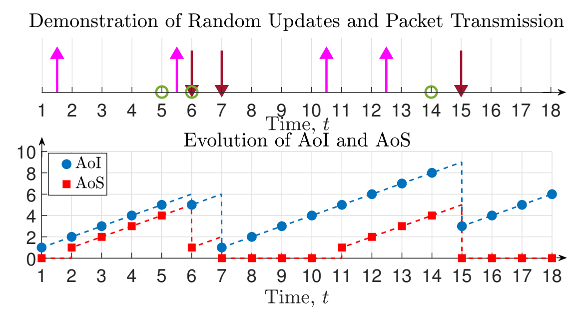

Fig. 2. On the top, sample sequences time-stamps of update arrivals (upward magenta arrows), update sending decisions (green circles) and update received (downward brown arrows). On the bottom, sample paths of AoI (blue) and AoS (red).

Now we return to the multiple-user scenario and introduce the dynamics of AoS considered in this paper. Let $s_n(t)$ be AoS of user n at the beginning of slot t before the scheduling decision and transmission. Assume that user n has no direct link to source n, and can only learn about source update through BS. If information at the user side has already been synchronized at the beginning of time t and no update is generated during slot t, i.e., $s_n(t)=0$ and $\Lambda_n(t)=0$, then the AoS of user n keeps to be zero at the beginning of next time slot $s_n(t+1)=0$, indicating that the update information will still be synchronized at the beginning of next time slot. If the information is synchronized at time t but an update packet arrives, i.e., $s_n(t)=0$ and $\Lambda_n(t)=1$, then information of user n will be desynchronized at the beginning of next time slot and $s_n(t+1)=1$. If $u_n(t)=1$ and the transmission succeeds with probability p_n, then the latest update by the end of slot $t-1$ about source n at the BS will be received by the end of slot t in this case. If no update packet arrives at time t, we will have $s_n(t+1)=0$, and if $\Lambda_n(t)=1$, an update during slot t arrives at BS by the end of slot t, then local information of user n will be out-of-date immediately at the beginning of slot

$t+1$, then $s_n(t+1)=1$. In other situations, the information at source n is desynchronized and the fresh update packet has not been received, the AoS increases linearly with time slots with $s_n(t+1)=s_n(t)+1$. Based on the above analysis, the dynamics of the AoS for user n is:

$$s_n(t+1)=\begin{cases}0, & s_n(t)=0,\Lambda_n(t)=0;\\ 1, & s_n(t)=0,\Lambda_n(t)=1;\\ 0, & \Lambda_n(t)=0,u_n(t)=1,\text{succeeds};\\ 1, & \Lambda_n(t)=1,u_n(t)=1,\text{succeeds};\\ s_n(t)+1, & \text{otherwise.}\end{cases} \tag{4}$$

C. Problem Formulation

Let $\mathbf{u}(t)=[u_1(t),u_2(t),\cdots,u_N(t)]$ be a scheduling decision at each time slot. We measure the data freshness over the entire network by the sum of time average expected AoS of all users,

$$\lim_{T\to\infty}\frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=1}^{T}\sum_{n=1}^{N}s_n(t)|\mathbf{s}(0)\right],$$

where $\mathbb{E}_\pi$ denotes the conditional expectation of AoS over the entire network when policy π is employed and the vector $\mathbf{s}(t)=[s_1(t),s_2(t),...,s_N(t)]\in\mathbb{N}^N$ denote the AoS of all users at time t. In this work, we assume that all the sources have been synchronized initially, i.e., $\mathbf{s}(0)=0$ and hence omit the item. We aim at designing a non-anticipated scheduling policy π that minimizes the above time-average expected AoS. The problem considered in this paper is organized as follows:

$$\text{minimize} \ \lim_{T\to\infty}\frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=1}^{T}\sum_{n=1}^{N}s_n(t)\right], \tag{5a}$$

$$\text{subject to} \ \sum_{n=1}^{N}u_n(t)\le 1, t=1,2,\cdots,. \tag{5b}$$

III. Markov Decision Process

In this section, we formulate our problem as a discrete time MDP aiming at minimizing the average cost over infinite horizon and then solve it by applying relative policy iteration through finite-state approximation.

A. MDP Formulation

The MDP problem consists of a quadruplet $(\mathbb{S},\mathbb{A},\text{Pr}(\cdot|\cdot,\cdot),C(\cdot,\cdot))$, where each item is explained as follows:

- **State space:** The state space $\mathbb{S}$ at time slot t is defined to be the AoS of all the users over the entire network $\mathbf{s}(t)$, which is countable but infinite because of possible transmission failures.
- **Action space:** We define the action $a(t)$ at time t to be the index of the selected user corresponding to a scheduling decision $\mathbf{u}(t)$, where $u_n(t)=\mathbb{1}_{\{n=a(t)\}}$ and $\mathbb{1}_{\{\cdot\}}$ is the indicator function. Denote $a(t)=0$ if the BS choose to be idle. The action space $\mathbb{A}=\{0,1,2,\cdots,N\}$ is hence countable and finite.
- **Transition probability:** Let $\text{Pr}(\mathbf{s}'|\mathbf{s},a)$ be the transition probability from state $\mathbf{s}(t)=\mathbf{s}=[s_1,s_2,\cdots,s_N]$ to state $\mathbf{s}(t+1)=\mathbf{s}'=[s'_1,s'_2,\cdots,s'_N]$ at the next slot by taking action a at slot t. Since the new update packet arrival and packet erasure are independent among the users, according to the AoS evolution dynamics (4), the transition probability can be decomposed into:

$$\text{Pr}(\mathbf{s}'|\mathbf{s},a)=\prod_{n=1}^{N}\text{Pr}(s'_n|s_n,a), \tag{6}$$

where $\text{Pr}(s'_n|s_n,a)$ denotes the one-step transition probability of user n given action a and has the following expression according to (4):

$$\text{Pr}(s'_n|s_n,a)=\begin{cases}1, & s'_n=s_n+1, s_n\neq 0, a\neq n;\\ 1-p_n, & s'_n=s_n+1, s_n\neq 0, a=n;\\ \lambda_n p_n, & s'_n=1, s_n\neq 0, a=n;\\ (1-\lambda_n)p_n, & s'_n=0, s_n\neq 0, a=n;\\ \lambda_n, & s'_n=1, s_n=0, \forall a\in\mathbb{A};\\ 1-\lambda_n, & s'_n=0, s_n=0, \forall a\in\mathbb{A}.\end{cases}$$

- **One-step cost:** Let $C(\mathbf{s}(t),a(t))$ be the one-step cost at state $\mathbf{s}(t)$ given action $a(t)$, representing the total AoS growth of the entire network:

$$C(\mathbf{s}(t),a(t))=\sum_{n=1}^{N}s_n(t).$$

The goal of the MDP is to minimize the AoS over the entire network by designing a policy $\pi: a(t)=\pi(\mathbf{s}(t))$. The optimal policy can approximated by minimizing the α-discounted cost over infinite horizon for $\alpha\to 1$. Denote $J_\alpha(\mathbf{s},\pi)=\limsup_{T\to\infty}\mathbb{E}_\pi\left[\sum_{t=1}^{T}\alpha^{t-1}C(\mathbf{s}(t),a(t))\right]$ to be the α-discounted cost over infinite horizon starting from state $\mathbf{s}(1)=\mathbf{s}$ by employing policy π and let $V_\alpha(\mathbf{s})=\min_\pi J_\alpha(\mathbf{s},\pi)$. Then $V_\alpha(\mathbf{s})$ satisfies the following Bellman equation:

$$V_\alpha(\mathbf{s})=\min_{a\in\mathbb{A}}\{C(\mathbf{s},a)+\alpha\sum_{\mathbf{s}'}V_\alpha(\mathbf{s}')\text{Pr}(\mathbf{s}'|\mathbf{s},a)\}. \tag{7}$$

B. Relative Policy Iteration through Finite-state Approximation

MDP problems with countable finite states can be solved by policy iteration or value iteration. To deal with the infinite state space in this problem, we set an upper bound of AoS for each user. Denote $x_n^{(m)}(t)$ be the truncated AoS of user n when the upper bound is m, the relationship with $x_n^{(m)}(t)$ and the actual AoS $s_n(t)$ is: $x_n^{(m)}(t)=\min\{s_n(t),m\}$. With such approximation, we can obtain a class of approximate MDP problems, where each problem differs from the primal problem with:

- **State space:** We substitute the AoS $\mathbf{s}(t)$ by truncated AoS $\mathbf{x}^{(m)}(t)=[x_1^{(m)}(t),x_2^{(m)}(t),\cdots,x_N^{(m)}(t)]$.
- **Transition probability:** The transition probability changes in accordance with the state space, let

$\Pr(\mathbf{x}^{(m)'}|\mathbf{x}^{(m)}, a)$ be the transition probability from state $\mathbf{x}(t) = \mathbf{x}^{(m)}$ to $\mathbf{x}(t+1) = \mathbf{x}^{(m)'}$ by applying action a:

$$\Pr(\mathbf{x}^{(m)'}|\mathbf{x}^{(m)}, a) = \prod_{n=1}^{N} \Pr(x_n^{(m)'}|x_n^{(m)}, a). \quad (8)$$

It should be noted that $\Pr(x_n^{(m)'}|x_n^{(m)}, a)$ is the same with $\Pr(s'_n|s_n, a)$ except:

$$\Pr(x_n^{(m)'}|x_n^{(m)}, a) = \begin{cases} 1, & x_n^{(m)'} = x_n^{(m)} = m, a \neq n; \\ 1-p_n, & x_n^{(m)'} = x_n^{(m)} = m, a = n. \end{cases} \quad (9)$$

For a given upper bound m and discount factor α, we can obtain a deterministic policy through relative policy iteration. The initial policy is set to be: select the user with largest AoS, i.e., $\pi^{(0)}(\mathbf{x}) = \arg_n x_n$. The policy $\pi^{(k+1)}(\mathbf{x})$ after the $(k+1)^{\text{th}}$ iteration can be obtained by the following steps:

Relative policy evaluation:

$$V_\alpha^{\text{tmp}}(\mathbf{x}) = C(\mathbf{x}, \pi^{(k)}(\mathbf{x})) + \alpha \sum_{\mathbf{x}'} \Pr(\mathbf{x}'|\mathbf{x}, \pi^{(k)}(\mathbf{x})) V_\alpha^{(k)}(\mathbf{x}'). \quad (10a)$$

$$V_\alpha^{(k+1)}(\mathbf{x}) = V_\alpha^{\text{tmp}}(\mathbf{x}) - V_\alpha^{\text{tmp}}(0). \quad (10b)$$

Policy selection:

$$\pi^{(k+1)}(\mathbf{x}) = \arg\min_{a \in \mathcal{A}} \{C(\mathbf{x}, a) + \alpha \sum_{\mathbf{x}'} \Pr(\mathbf{x}'|\mathbf{x}, a) V_\alpha^{(k)}(\mathbf{x}')\}. \quad (10c)$$

The iteration terminates until the policy $\pi(\mathbf{x})$ and value function $V_\alpha(\cdot)$ remains stable. The MDP scheduling policy is obtained as follows: at each slot with state $\mathbf{s}(t)$, compute the corresponding virtual AoS $\mathbf{x}^{(m)}(t)$ and choose action $a(t) = \pi(\mathbf{x}^{(m)}(t))$.

IV. Index-based Heuristic

MDP solution is computationally demanding for a large number of access users. To reduce computational complexity, we propose a simple index-based heuristic policy based on restless multi-arm bandit (RMAB) [8].

A. Decoupled sub-problem

First we relax the transmission constraints in each time slot into an average transmission constraint, $\frac{1}{T}\sum_{t=1}^{T}\sum_{n=1}^{N} \mathbb{E}[u_n(t)] \leq 1$. Then we use the Lagragre multiplier to formulate the relaxed multi-arm bandit problem with the relaxed constraint,

$$\text{minimize} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\pi \left[\sum_{n=1}^{N} s_n(t) + W\mathbb{E}[u_n(t)] - \frac{W}{N} \right], \quad (11a)$$

$$\text{subject to } W \geq 0. \quad (11b)$$

Each of the users can be viewed as a restless bandit and can be decoupled separately. The subscript is omitted henceforth. For each bandit, we try to minimize: $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[s(t) + Wu(t)]$ The multiplier $W \geq 0$ is positive, thus it can be seen that the bandit has an extra cost of being selected to send updates. Given W, the goal of each decoupled optimization problem is to derive an optimum update strategy of whether to transmit update or not, to achieve a balance between update cost and the cost incurred by AoS. The subproblem is then formulated into an MDP consists of a quadruplet $(\mathbb{S}, \mathbb{A}, \Pr(\cdot|\cdot,\cdot), C(\cdot,\cdot))$:

- **State space:** The state at time t is the current AoS $s(t) \in \mathbb{N}$, which is countable infinite.
- **Action space:** There are two possible actions at each time slot, either choose the bandit to send updates $a(t) = 1$ or remain idle $a(t) = 0$. It should be noted here that the action $a(t)$ here is different from the scheduling actions $\mathbf{u}(t)$, which has strict interference constraint.
- **Transition probability:** Let $\Pr(s'|s, a)$ be the transition probability from state $s(t) = s$ to $s(t+1) = s'$ by taking action a at time t, the state evolves with the action following Eqn. (4).
- **One-step cost:** According to the subproblem formation, for fixed W, the one step cost of state $s(t)$ with action $a(t)$ consists of the AoS in the current time slot and the extra cost of being active:

$$C(s(t), a(t)) = s(t) + Wa(t). \quad (12)$$

We would like to design a policy $\pi : a(t) = \pi(s(t))$ such that the average cost over infinite horizon can be minimized. The α-discounted cost of policy π over infinite horizon starting from initial state $s(1) = s$ is defined by $J_\alpha(s, \pi) = \lim_{T\to\infty} \sup \mathbb{E}_\pi \left[\sum_{t=1}^{T} \alpha^{t-1} C(s(t), \pi(s(t)))\right]$. Next we will investigate the structure of the optimum stationary deterministic policy by examining the value function $V_\alpha(s) = \min_\pi J_\alpha(s, \pi)$ and derive the Whittle's index based on the structure. The Bellman equation is provided as follows:

$$V_\alpha(s) = \min_{a \in \mathbb{A}} \{C(s, a) + \alpha \sum_{s'} V_\alpha(s') \Pr(s'|s, a)\}. \quad (13)$$

B. Proof of Indexability

By definition [8], a bandit is indexable if the passive set increases monotonically with extra cost W. We first prove the monotonic of the value function $V_\alpha(s)$ and show that the optimum policy to minimize the α-discounted cost has a threshold structure. The threshold structure to minimize the average cost can be obtained by $\alpha \to 1$.

Lemma 1. *For fixed $W > 0$, the value function $V_\alpha(s)$ increases monotonically with s.*

Proof: The main step of proving the lemma is by induction of the sum of discounted cost over finite time horizon. Starting from quite intuitive conclusion that the one-step cost function is increasing, by induction we can show that the sum discounted cost function under any finite steps by applying any policy is increasing, then by taking the minimum of all the policies we will obtain the increasing characteristic of the value function. The detailed proof is omitted due to space constraint. ■

Theorem 1. *The optimal policy has a threshold structure, i.e., if at state s, it is optimal to keep the bandit idle, then for the*

$$\tau_{opt} = \lfloor\left(\frac{5}{2}-\frac{1}{p}-\frac{1}{\lambda}\right)+\sqrt{\left(\frac{5}{2}-\frac{1}{p}-\frac{1}{\lambda}\right)^2+2(\frac{W}{p}+\frac{1-\lambda}{\lambda}\frac{1-p}{p})+2\frac{1-p}{p}}\rfloor. \quad (15)$$

all $s' < s$ it is optimal to keep the bandit idle; otherwise, if it is optimal to activate the bandit at state s, then for states $s+1, s+2, \cdots$, the optimal strategy is to keep the bandit active.

Proof: The decision $a(t)$ is chosen according to the Bellman equation (13). Suppose at state s, the optimum strategy to minimize the total cost is to choose the bandit to be active $a = 1$, then we have $\lambda V_\alpha(1) + (1-\lambda)V_\alpha(0) + \frac{1}{\alpha}W \leq V_\alpha(s)$. According to the monotonic characteristic, for all states $s' > s$, we have $V_\alpha(s) \leq V_\alpha(s')$. Hence for state $s' > s$ the choice will always be choosing the bandit to be active. By taking $\alpha \to 1$, this provides insight that the optimum policy to minimize the average cost possesses a threshold structure. ■

Corollary 1. *Denote $F_\tau(W)$ to be the average cost for given W if threshold policy τ is employed, i.e., the bandit will be active for state $s \geq \tau$ and passive for $s < \tau$. Then,*

$$F_\tau(W) = \frac{\tau(\tau-1)}{2}\xi_1^{(\tau)} + \frac{\xi_1^{(\tau)}}{p}(\frac{1}{p}-1) + \frac{\xi_1^{(\tau)}}{p}(\tau+W), \quad (14)$$

where $\xi_1^{(\tau)} = 1/\left(\frac{1-\lambda}{\lambda}+\tau+\frac{1}{p}-1\right)$ denotes the steady state distribution for $s = 1$ if threshold policy τ is employed.

For given W, the optimal threshold τ_{opt} should satisfy $F_{\tau_{opt}+1}(W) \geq F_{\tau_{opt}}(W)$ and $F_{\tau_{opt}-1}(W) \geq F_{\tau_{opt}}(W)$. As derived in (15), the threshold is a non-decreasing function of W. Thus, the passive set increases monotonically with W and indexability is proved.

C. Whittle's Index Policy

According to [8], the Whittle's index $I(s)$ is the extra cost that makes action $a = 1$ and $a = 0$ for states s equally desirable, which can be computed as follows:

$$I(s) = \frac{F_{\tau+1}(0) - F_\tau(0)}{(\xi_1^{(\tau)} - \xi_1^{(\tau+1)})/p}. \quad (16)$$

The scheduling algorithm is provided as follows: at the beginning of each time slot, the BS observes current AoS of each user $s_n(t)$ and compute the Whittle's index for each user $I_n(s_n(t))$. Then, broadcast the corresponding message of user n with the highest $I_n(s_n(t))$, with ties broke arbitrarily.

V. Simulations

In Fig. 3, we consider a three user broadcast network with arriving rate $\lambda = [0.3, 0.4, 0.3]\lambda_{total}$ and success transmission probability $p = [0.2, 0.55, 0.9]$. The threshold m for computing the truncated MDP solution is set to be $m = 20$. The AoS is by taking the average of $T = 10^6$ time slots. Our results show that the performance of the proposed index policy is comparable to the policy obtained by truncated MDP. Compared to the greedy policy of selecting the user with largest AoS, the proposed algorithm improve the performance significantly.

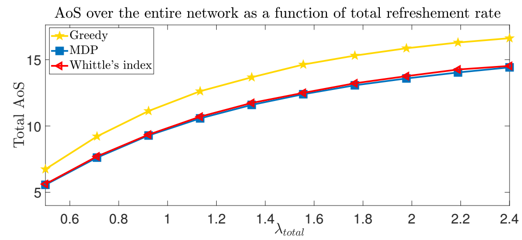

Fig. 3. Simulation results of expected AoS for a three user broadcast network with $\lambda = [0.3, 0.4, 0.3]\lambda_{total}$ and $p = [0.2, 0.55, 0.9]$.

VI. Conclusions

In this paper, we study the problem of *Age of Synchronization* optimization over wireless broadcast network. Scheduling policies based on MDP techniques and Whittle's index are proposed and analyzed. Our setting is based on the assumption that the statistics of random update arrival rate and channel states are known, in the future we will study scheduling with these coefficients unknown.

Acknowledgement

This work was supported in part by the National Key R&D Program of China (Grant No.2017YFE011230) and Tsinghua University Tutor Research Fund. The authors are grateful to the anonymous reviewers for their valuable suggestions.

References

[1] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*, March 2012, pp. 2731–2735.
[2] J. Zhong, R. D. Yates, and E. Soljanin, "Two Freshness Metrics for Local Cache Refresh," in *2018 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2018, pp. 1924–1928.
[3] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2016, pp. 844–851.
[4] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.
[5] Y. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 561–565.
[6] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc' 18)*, ACM, New York, NY, USA, pp. 191–200.
[7] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Niu, "Decentralized status update for age-of-information optimization in wireless multi-access channels," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2276–2280.
[8] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.