Achieving Information Freshness with Selfish and Rational Users in Mobile Crowd-Learning

Bin Li, *Senior Member, IEEE,* and Jia Liu, *Senior Member, IEEE*

***Abstract*—The proliferation of smart mobile devices has spurred an explosive growth of mobile crowd-learning services, where service providers rely on the user community to voluntarily collect, report, and share real-time information for a collection of scattered points of interest (PoI). A critical factor affecting the future large-scale adoption of such mobile crowd-learning applications is the freshness of the crowd-learned information, which can be measured by a metric termed "age-of-information" (AoI). However, we show that the AoI of mobile crowd-learning could be arbitrarily bad under selfish and rational users' behaviors if the system is poorly designed. This motivates us to design efficient reward mechanisms to incentivize mobile users to report information in time, with the goal to keep the AoI and congestion level of each PoI low. Toward this end, we consider a simple linear AoI-based reward mechanism and analyze its AoI and congestion performances in terms of price of anarchy (PoA), which characterizes the degradation of the system efficiency due to selfish and rational behavior of users. In this paper, we consider both average maximum age and average weighted sum of age. Remarkably, we show that the proposed mechanism achieves the optimal AoI performance in terms of average maximum age asymptotically in a deterministic scenario, i.e., the corresponding PoA decreases to 0 asymptotically. Moreover, the PoA in terms of average total age under our proposed mechanism can be upper-bounded by $1/2$ asymptotically. Further, we prove that the proposed mechanism achieves a bounded PoA in general stochastic cases, and the bound only depends on system parameters. Particularly, when the service rates of PoIs are symmetric in stochastic cases, the achieved PoA is upper-bounded by $1/2$ asymptotically. Collectively, this work advances our understanding of information freshness in mobile crowd-learning systems.**

***Index Terms*—Age of Information, Congestion Control, Price of Anarchy, Mobile Crowd-Learning.**

I. Introduction

Fueled by the proliferation of smart mobile devices (e.g., smartphones, tablets, etc.), recent years have witnessed a rapid growth of information services and data analytics based on large-scale *crowd-learning*. A key defining feature of these crowd-learning applications is that they rely on the user community to voluntarily collect, report, and share real-time information for a set of distributed points of interest (PoI). Such crowd-learned information will in turn benefit the users themselves and attract more users to join the community (by reputation, word of mouth, etc.), which further enhances the accuracy, value, and significance of the crowd-learning applications. For example, the real-time traffic congestion and accident information on Google Waze [1] (a community-based GPS system) relies on the reports from mobile devices and the tracking of their locations, densities, and trajectories. As another example, by offering a variety of incentives, many data analytics services leverage their user communities to share real-time information of scattered commodities and resources, such as cheap gasoline stations (e.g., GasBuddy [2]), parking space availability (e.g., Pavemint [3]), free WiFi hotspots (e.g., WiFi Finder [4]), popular grocery deals information (e.g., Basket [5]), to name just a few. It can be foreseen that new crowd-learning applications will continue to emerge.

Although mobile crowd-learning holds a great potential to fundamentally change our modern society, a critical factor affecting its future large-scale adoption is the *freshness* of the crowd-learned information, which can be measured by a fundamental metric termed "Age-of-Information" (AoI). Guaranteeing information freshness in crowd-learning is critical because stale information discourages existing and new users from participating, which in turn degrades the information freshness and creates a vicious circle. Unfortunately, due to the special dynamics between the service provider and the users, there is an inherent lack of information freshness guarantee in mobile crowd-learning: First, to maintain information freshness, the service provider needs to incentivize the users to update the states of the PoIs. Second, the crowd-learning users are "selfish and rational" in the sense that their best interest is to maximize their own benefit from participating in crowd-learning, rather than minimizing the AoI for the service provider. Hence, a poorly designed incentive mechanism could result in two undesirable consequences: (i) too many users flock to an attractive PoI, which leads to redundant sampling and severe queueing congestion; and (ii) all other PoIs suffer from large AoI because of under-sampling. In light of these unique characteristics of mobile crowd-learning, several fundamental open questions naturally arise:

1) Is it possible to guarantee information freshness by incentivizing selfish and rational users in mobile crowd-learning?
2) If the answer to 1) is "yes," what is the fundamental relationship between reward and AoI in crowd-learning?
3) How to design reward mechanisms to avoid large queueing congestion while guaranteeing AoI in crowd-learning?

Manuscript received July 8, 2020, revised November 28, 2020, and accepted February 13, 2021. This work is supported in part by NSF grants CNS-1717108, CNS-1815563, CNS-1942383, ECCS-1818791, CCF-2110252, CNS-2102233, CNS-2110259, ONR grant N00014-17-1-2417, and a Google Faculty Research Award. This paper was presented in part in the 17th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), Paris, France, June 2019. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of ONR.

Bin Li is with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, 02881, USA (e-mail: binli@uri.edu).

Jia Liu is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH, 43210, USA (e-mail: liu@ece.osu.edu).

Authors are listed in alphabetical order. Both authors are primary authors.

However, answering the above questions is non-trivial because joint AoI and congestion analysis in mobile crowd-learning faces the following challenges: First, there is a lack of analytical model that characterizes the essential features of mobile crowd-learning in the literature. Most of the existing work on crowd-sensing are based on static models that hardly capture the dynamic and stochastic nature of participating users in mobile crowd-learning. Second, as shown by recent studies (see, e.g., [6]–[9]), AoI dynamics are fundamentally different from the traditional queueing evolution, which necessitates new theoretical tools. Third, as will be shown later, there is a strong *coupling* between the AoI and queue-length processes in crowd-learning, where changing the design of either one would significantly affect that of the other.

In this paper, we overcome the above challenges and propose a new analytical model coupled with the *Price of Anarchy* (PoA) metric, which characterizes the degradation of a system due to selfish and rational behavior of users[1]. This enables us to analyze and understand the relationships between AoI, queueing congestion, and reward mechanism design under users' selfishness and rationality. The most significant and perhaps somewhat surprising insight is that, even under selfish users' behaviors, it remains possible to design simple (linear) reward mechanisms for mobile crowd-learning systems that enjoy good (or even optimal in the deterministic case) AoI and queueing congestion performances in terms of PoA (price of anarchy) metrics. The main results and contributions of this paper are summarized as follows:

- First, we develop a new analytical model for mobile crowd-learning, which takes into account the strong couplings between the stochastic arrivals of participating users, PoIs' information evolutions, and reward mechanisms. As will be discussed next, this new analytical model enables us to reveal the fundamental scaling law between AoI, queueing congestion, and the reward rate set by the service provider.
- Next, as a starting point, we analyze the AoI performance under a linear AoI-based reward mechanism in a deterministic setting, where there is exactly one arriving user in each time slot, and each PoI serves exactly one user (if any) in each time slot (and hence no queueing effect in this setting). We show that given an AoI reward rate β, the PoA in terms of average maximum age and average total age are upper-bounded by $O(1/\beta)$ and $O(1/2+1/\beta)$, respectively, which implies that the system achieves the *optimal* average maximum age and *guaranteed* average total age as β increases asymptotically.
- Finally, based on our results for the deterministic case, we characterize the joint AoI-congestion performance of mobile crowd-learning for stochastic settings. Although the reward policy design for joint AoI and queueing congestion optimization remains an open problem in stochastic settings, surprisingly, we show that the above linear AoI-based reward mechanism yields a bounded PoA, which only depends on the arrival and service parameters of the system. In the case of symmetric services, the PoA is upper-bounded by $1/2$ as the reward rate β increases asymptotically.

Collectively, our results in this paper advance the understanding of achieving information freshness in mobile crowd-learning with selfish and rational users. The remainder of this paper is organized as follows: Section II reviews related work. Section III introduces system model and problem statement. Section IV introduces a linear reward mechanism, and Sections V–VI study its PoAs in the deterministic and stochastic cases, respectively. Section VII presents numerical results and Section VIII concludes this paper.

II. Related Work

To put our work in comparative perspectives, in this section, we provide an overview on the related work in the areas of crowd-sensing and age-of-information, respectively.

a) Crowd-Sensing: In the literature, crowd-sensing refers to the sensing model where a group of individuals collectively measure some *common* phenomena, e.g., environmental quality monitoring [10], noise pollution assessment [11], [12], and traffic monitoring [13]. Although crowd-sensing bears some similarity to mobile crowd-learning, the main focuses of the crowd-sensing research community are on network resource management, system infrastructure, incentive mechanism designs, etc. (see [14] for a comprehensive survey). In contrast, the overarching theme of this paper is to guarantee *information freshness* in learning *scattered* objects by a *selfish and rational* crowd. Moreover, most of the existing crowd-sensing research adopts either a static model, where the set of sensing individuals is fixed (see, e.g., [15] and references therein); or based on a static game-theoretic model, where a fixed set of sensing individuals are incentivized/contracted by a fixed set of employers (see, e.g., [16] and references therein). These are fundamentally different from our *dynamic* model described in Section III. Hence, our work fills a critical gap in understanding large-scale mobile crowd-learning.

b) Age-of-Information (AoI): Originated from sensing systems, AoI has attracted increasing attention from the information theory, signal processing, and communications communities in recent years. Besides being a useful performance metric, AoI also possesses several key features that distinguish itself from the traditional notion of queueing delay. Most notably, in many sensing systems, it has been found that while queueing delay benefits from lower sampling rates (implying less data traffic), AoI is non-monotone with respect to sampling rates. This key difference has sparked AoI research in several aspects, e.g., real-time sampling and remote estimation trade-off [17], [18], joint source-channel coding exploitation [19], [20], caching [21], [22], optimization algorithms for AoI minimization [23], [24], age-based scheduling [25]–[27], just to name a few. We note that the key differences between our research and the existing AoI research are: i) the tight coupling and dependence between *multi-user arrival dynamics* and *multi-source information time series* on a *network level*; and ii) the complex interactions between AoI, fresh/outdated information, and queueing, all of which are governed by the service provider's *reward mechanism designs*. These key

[1]The value of PoA is always between 0 and 1, and the larger the PoA, the less efficient the system. See Sections IV–VI for more in-depth discussions.

differences introduce new dimensions of challenges in guaranteeing stochastic network information freshness that is unseen in existing AoI research.

III. Network Model and Problem Statement

As shown in Fig. 1, we consider a mobile crowd-learning system consisting of N nodes that represent N points of interest (PoI) associated with the same application. Depending on the application, the PoIs could be road intersections, parking garages, potential WiFi hotspots, or gas stations. We consider a time-slotted system. This mobile crowd-learning system follows the operating procedure: we suppose that the system has a set of points of interest (PoI) (which could represent gas stations, or a set of routes between Point A and Point B on Google Waze). Each PoI has a real-time state information denoted as $p_n[t]$, which could represent gas price, real-time congestion state, etc. For each PoI n, the service provider maintains a record, denoted as $r_n[t]$, which is sampled by the users, and may or may not be up-to-date. Meanwhile, a stream of users randomly arrive at the service provider, which models the fact that users at different places pull out their phones to check the recorded information $r_n[t]$. Based on the information they see, they would pick one PoI to go, e.g., picking the cheapest gas station, or the least congested route, etc. For some applications, it could happen that when they arrive at the chosen PoI, they need to stay in line waiting to get service. We assume that once they reach the head of line and get their services, the real-time state $p_n[t]$ is revealed. Upon observing the real-time information, the users will be able to earn some reward by feeding back the information to the service provider and the $r_n[t]$ will get updated.

In what follows, we model the above operating procedure analytically. In each time slot t, each PoI n has some state information $p_n[t]$ (e.g., congestion level, parking rate and space, gas price, etc.) that is time-varying and to be sampled by their users. We assume that $p_n[t] \in [p_{\min}, p_{\max}], \forall t$, for some positive constants $p_{\min}$ and $p_{\max}$. A service provider (i.e., a crowd-learning-based information/data analytics platform) relies on randomly arriving users to sample and report the states of the PoIs. The service provider maintains a record for each PoI, whose value in time slot t is denoted as $r_n[t]$, $n = 1, \ldots, N$. For ease of exposition, we will refer to $p_n[t]$ and $r_n[t]$ as "*price*" and "*recorded price*" in the rest of this paper, respectively. Let $u_n[t]$ be the most recent update time up to time slot t for PoI n's record. Hence, the *age* (freshness) of record $r_n[t]$ in time slot t can be represented as $\Delta_n[t] = t - u_n[t]$.

Let $A[t]$ be the random number of users arriving at the system in time slot t. We assume that $A[t], t \geq 0$, are independently and identically distributed (i.i.d.) across time with mean $\lambda \triangleq \mathbb{E}[A[t]] > 0$ and a bounded second moment $\mathbb{E}[A^2[t]] < \infty$. The arrivals model the scenario that users at different locations use their mobile apps in each time slot to acquire information of the PoIs before making decisions. Each arriving user will first observe the current records of all PoIs and choose one that maximizes his/her benefits, e.g., choosing the least congested route, the lowest gas price, or the cheapest and nearest parking space, depending on the specific application. However, due to the *random* updating time in crowd-learning, the information of some PoI n's record could be old and hence $r_n[t]$ may be outdated and inaccurate.

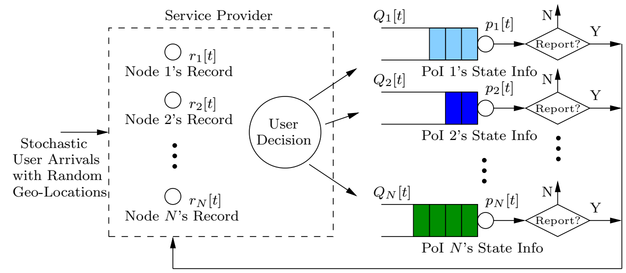


Fig. 1: A system model for mobile crowd-learning.

On the other hand, upon the arrival at his/her chosen PoI, say n in time slot t, the user will report the PoI's real-time state (e.g., real-time price, congestion level), i.e., $p_n[t]$. Let $R_n[t]$ denote the number of users that can be served by PoI n in time slot t. We assume that $R_n[t]$, $t \geq 0$, are i.i.d. across time and independently distributed across PoIs with mean $\mu_n \triangleq \mathbb{E}[R_n[t]] > 0, \forall n$, and $R_n[t] \leq R_{\max}, \forall n, t$, for some $R_{\max} < \infty$. We use $Q_n[t]$ to denote the number of users awaiting for service at PoI n in time slot t.

The service provider's goal is to achieve minimum time-average AoI while keeping queueing congestion at each PoI low. The rationale behind this goal is that low AoI (i.e., fresh information) implies multiple benefits, e.g., high information accuracy, which attracts more users; hence more advertising revenues due to large user volume, etc. However, the following toy example shows that, without an appropriately designed incentive mechanism, the natural greedy behavior of selfish and rational users could yield *AoI instability* in a mobile crowd-learning environment:

A Motivating Example (AoI Instability due to Selfishness): Consider a two-PoI example as shown in Fig. 2. The most straightforward incentive mechanism is that the service provider offers the *same* reward for whatever PoI a user samples. Since sampling different PoIs makes no difference in terms of reward, the users would naturally choose the following *price-greedy* policy: In time slot t, each arriving user compares the recorded prices $r_1[t]$ and $r_2[t]$ and chooses the cheaper PoI, i.e., choosing $n^*[t] \in \arg\min_{n \in \{1,2\}} \{r_n[t]\}$. Suppose that $p_n[t] \in [0, p_{\max}]$, $n = 1, 2$. Assume that the probability $\Pr\{p_n[t] = p_{\max}\} = \epsilon$, $n = 1, 2$, where $\epsilon > 0$ is some small value. Now, suppose also that in the initial state, $p_1[0] = p_{\max}$ and $p_2[0] = \delta < p_{\max}$. Thus, at $t = 0$, *all* users choose PoI 2 and the record $r_2[t]$ will be updated, in which case the age of PoI 2 in time slot 1 becomes zero, i.e., $\Delta_2[1] = 0$. However, due to the high initial price $p_1[0]$, no user chooses PoI 1. Also, due to the low probability of $p_2[t]$ reaching $p_{\max}$, it would take a long time (could be unbounded if ϵ is arbitrarily small) for PoI 1 to receive *any* user to update $r_1[\cdot]$, although $p_1[t]$ may be lower than $p_2[t]$. For example, in Fig. 3, $p_1[t]$ and $p_2[t]$ are uniformly distributed in $[0, 1]$. We let $p_1[0] = 0.999$ (large initial value) and $p_2[0] = 0.1$. Clearly, we can see that PoI 1's AoI is large and grows linearly with respect to time. □

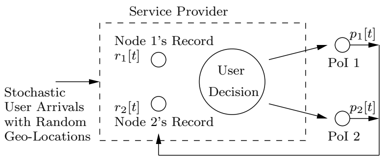


Fig. 2: A two-PoI motivating example with $p_1[0] = 0.999$ and $p_2[0] = 0.1$.

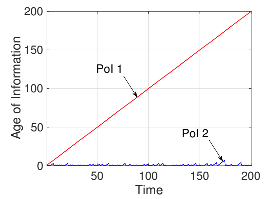


Fig. 3: Large and unstable AoI of PoI 1 in Fig. 2.

The above observation of AoI instability due to users' selfishness and rationality highlights the importance of reward mechanism design and motivates us to develop crowd-learning reward mechanisms to ensure information freshness.

IV. A Linear AoI-Based Reward Mechanism

To keep the AoI being bounded, the service provider would like users to go to and sample a PoI with the most outdated information. However, unlike traditional scheduling problems, the crowd-learning service provider cannot enforce each arriving selfish and rational user to go to a certain PoI. Rather, the service provider can only offer incentives/rewards to influence the users to choose certain PoIs. So far, however, the problem of optimal reward mechanism design for mobile crowd-learning with selfish and rational users has not been addressed in the literature. As a starting point, in this paper, we consider a simple *linear* reward mechanism for mobile crowd-learning.

Specifically, we let $\beta > 0$ represent the "*reward per unit of age*" offered by the service provider. Note that each user prefers to select a PoI with both low price and congestion level. We use a parameter $\gamma > 0$ to denote users' sensitivity to queueing congestion, which depends on specific mobile crowd-learning application[2]. Hence, in each time slot t, an arriving user's *presumed benefit* for choosing PoI n and reporting its state is: $\beta\Delta_n[t] - \gamma Q_n[t] - r_n[t]$. In this work, we assume that all arriving users are selfish and rational in the sense that they would select a PoI $n^*[t]$ to maximize their presumed benefit, i.e.,

$$n^*[t] \in \underset{n\in\{1,2,\ldots,N\}}{\arg\max} \left(\beta\Delta_n[t] - \gamma Q_n[t] - r_n[t]\right), \ \forall t. \quad (1)$$

Although the above linear reward mechanism is simple, it captures the following *first-order dynamics* of age-based incentives: On one hand, for any fixed γ, when the reward rate diminishes, i.e., $\beta \downarrow 0$, each user essentially follows the "greedy" scheme to select a PoI with the smallest value of $\gamma Q_n[t] + r_n[t]$. On the other hand, when the reward rate approaches infinity, i.e., $\beta \uparrow \infty$, the component of age-based reward dominates and the effects of $Q_n[t]$ and $r_n[t]$ become negligible asymptotically in users' presumed benefit. Thus, it encourages users to help the service provider maintain information freshness.

To facilitate our subsequent analysis, we use $S_n[t]$ to denote whether there is at least one user selecting PoI n in time slot t. In particular, $S_n[t] = 1$ if at least one user selects PoI n in time slot t, and $S_n[t] = 0$ otherwise. Under the assumption of users' selfishness and rationality, the dynamics of queue-length and age of PoI n can be described as follows:

$$Q_n[t+1] = \max\{Q_n[t] + A[t]S_n^*[t] - R_n[t], 0\}, \forall n. \quad (2)$$

$$\text{and } \Delta_n[t+1] = \begin{cases} \Delta_n[t]+1, & \text{if } S_n^*[t]\mathbb{1}_{\{A[t]>0\}} = 0, \\ 0, & \text{otherwise}, \end{cases} \quad (3)$$

where $S_n^*[t] = 1$ if $n = n^*[t]$ and $S_n^*[t] = 0$ otherwise, and $\mathbb{1}_{\{\cdot\}}$ is an indicator function. Let $\mathbf{S}^*[t] \triangleq (S_n^*[t])_{n=1}^N$.

To understand the impact of users' selfishness and rationality on AoI and queueing congestion, in this paper, we adopt the so-called *Price of Anarchy* (PoA) metric from the game theory literature, which characterizes the degradation of the system efficiency due to the selfish and rational behavior of users compared to the optimum. Roughly speaking, the notion of PoA ρ is defined as:

$$\rho = 1 - \frac{\text{Minimum cost}}{\text{Cost under selfish and rational behavior}}. \quad (4)$$

Here, the definition of cost in (4) depends on specific systems scenarios that will be further defined in subsequent sections. Note that $\rho \in [0, 1]$ and the smaller the PoA, the more efficient the system under selfish and rational user behaviors.[3] In what follows, we will analyze the PoA performance of the linear reward scheme.

V. Price of Anarchy: A Deterministic Case

In our proposed crowd-learning model, there exists a strong coupling between AoI and queueing congestion. Indeed, when a user joins a PoI, the AoI of that PoI decreases while its queue-length increases. Such a coupling poses significant challenges on the performance analysis of the linear reward mechanism. As a starting point, in this section, we first consider a simple deterministic case, where, in each time slot, there is exactly one arriving user and each PoI serves exactly one user if there is any. This deterministic case not only provides interesting insights, its results and proof strategies will also serve as a foundation for analyzing general cases with stochastic arrivals and services later in Section VI.

Note that due to the special arrival and service patterns in this deterministic case, there is *no queueing effect* at each PoI. Hence, user's selfish and rational selection (cf. (1)) becomes:

$$n^*[t] \in \underset{n\in\{1,2,\ldots,N\}}{\arg\max} \left(\beta\Delta_n[t] - r_n[t]\right). \quad (5)$$

In addition, the evolution of age of PoI n in (3) becomes:

$$\Delta_n[t+1] = (\Delta_n[t] + 1)(1 - S_n^*[t]). \quad (6)$$

Next, we study the information freshness performance under the selfish and rational behavior of users (cf. (5)) based on the notion of PoA. Since queueing does not play a role,

[2]Here, we assume that all users are homogeneous and have the same γ-value. The impact of users' heterogeneity in congestion sensitivity will be studied in our future work.

[3]It is insightful to point out that the notion of PoA in this paper has the following subtle difference compared to the traditional PoA in the game theory literature. As will be seen later, the cost in (4) will be based on AoI and queueing congestion, which are more from the system provider's perspective. By contrast, the traditional PoA in the game theory literature is often measured by the users' social costs (or welfare).

the cost in PoA in the deterministic case is only related to age. To this end, we consider two different cost functions characterizing the AoI performance: average maximum age $\overline{\Delta}^{(\beta)}_{\max} \triangleq \mathbb{E}\left[\max_n \Delta_n[t]\right]$ and total average age $\overline{\Delta}^{(\beta)}_{\Sigma} \triangleq \sum_{n=1}^{N} \mathbb{E}\left[\Delta_n[t]\right]$ under some reward rate β. Correspondingly, their associate PoAs are defined as follows.

$$\rho_m(\beta) \triangleq 1 - \frac{\overline{\Delta}^{(\text{OPT})}_{\max}}{\overline{\Delta}^{(\beta)}_{\max}}, \quad \text{and} \quad \rho_s(\beta) \triangleq 1 - \frac{\overline{\Delta}^{(\text{OPT})}_{\Sigma}}{\overline{\Delta}^{(\beta)}_{\Sigma}},$$

where $\overline{\Delta}^{(\text{OPT})}_{\max}$ and $\overline{\Delta}^{(\text{OPT})}_{\Sigma}$ are the average maximum age and the total average age under an optimal policy, respectively. The first main result of this paper is stated as follows:

Theorem 1 (AoI-Based PoA for the Deterministic Case)**.** *If there is exactly one user arriving in each time slot and each PoI serves exactly one user per time-slot if there is any, the users' selfishness and rationality yields the following PoA performance:*

$$\rho_m(\beta) \leq \frac{p_{\max}}{(N-1)\beta + p_{\max}} = O\left(\frac{1}{\beta}\right), \tag{7}$$

$$\textit{and } \rho_s(\beta) \leq \frac{(N-1)\beta + 2p_{\max}}{2(N-1)\beta + 2p_{\max}} = O\left(\frac{1}{2} + \frac{1}{\beta}\right). \tag{8}$$

Proof. The proof consists of two main steps: (i) Finding an upper bound on the average maximum age $\overline{\Delta}^{(\beta)}_{\max}$ and the total average age $\overline{\Delta}^{(\beta)}_{\Sigma}$ due to users' selfishness and rationality, respectively; and (ii) deriving a lower bound on the average maximum age $\overline{\Delta}^{(\text{OPT})}_{\max}$ and the total average age $\overline{\Delta}^{(\text{OPT})}_{\Sigma}$ achieved by their corresponding optimal policies, respectively. Please see Appendix A for details. □

Remark 1. From Theorem 1, we can observe that if β increases asymptotically (i.e., $\beta \uparrow \infty$), we have $\rho_m(\beta) \downarrow 0$. This implies that the system is optimal in terms of average maximum age and mimicking Round-Robin when the service provider increases the incentive asymptotically. On the other hand, if β reduces to zero (i.e., $\beta \downarrow 0$), we can see from (5) that each user just follows a price-greedy strategy. In this case, Theorem 1 suggests that the upper bound of $\rho_m(\beta)$ approaches 1, which is consistent with our observation (cf. motivating example in Section III) that the system suffers a poor AoI performance and potentially AoI instability (i.e., $\overline{\Delta}^{(\beta)}_{\max} \uparrow \infty$). iii) $\rho_s(\beta) \downarrow 1/2$ as β increases to infinity. However, we will observe from the numerical results in Section VII that $\rho_s(\beta)$ indeed converges to zero as β increases asymptotically. This indicates the looseness of our upper bound analysis for the average total age. This is due to the intrinsic deficiency of the Lyapunov analysis methodology, which only captures the drift among neighboring slots in temporal domain and is not able to characterize the Round-Robin behavior in spatial domain.

VI. Price of Anarchy: Stochastic Cases

Based on the results for the deterministic case, we are now in a position to analyze the AoI and congestion performances under users' selfishness and rationality in cases with stochastic arrivals and services. To facilitate analysis, we define a parameter $q \triangleq \Pr\{A[t] > 0\}$ for the arrivals, which is strictly positive for $\lambda \triangleq \mathbb{E}[A[t]] > 0$. Let $\mu_\Sigma \triangleq \sum_{n=1}^{N} \mu_n$. Here, we adopt the cost function with the weighted sum of average age and average queue-length, i.e.,

$$J(\beta,\gamma) \triangleq \frac{\gamma\epsilon}{N} \sum_{n=1}^{N} \overline{Q}_n + \beta \sum_{n=1}^{N} \frac{\mu_n}{\mu_\Sigma} \overline{\Delta}_n, \tag{9}$$

where $\epsilon > 0$ satisfies $\mu_n/\lambda \geq \mu_n/\mu_\Sigma + \epsilon/N, \forall n = 1, 2, \ldots, N$ due to the fact that $\lambda < \mu_\Sigma$ (necessary for guaranteeing the system's queueing stability[4]), and $\overline{Q}_n$ and $\overline{\Delta}_n$ are the average queue-length and average age of PoI n under the user's selfishness, respectively. We note that in $J(\beta,\gamma)$, ϵ is used as a scaling parameter to reduce the cost's sensitivity to average queue-length $\frac{1}{N}\sum_{n=1}^{N} \overline{Q}_n$ under different arrival rates λ. Also, γ and β are used to emphasize the relative importance between queueing and AoI costs, as in the presumed benefit for users' selfish decisions (cf. (1)). Also, note that $J(\beta,\gamma)$ is based on weighted average age, where the weight $\frac{\mu_n}{\mu_\Sigma}$ is used to "equalize" the different AoI scales caused by the heterogeneity of the PoIs.[5] As a result, the PoA is specialized to

$$\rho(\beta,\gamma) \triangleq 1 - \frac{J^{(\text{OPT})}(\beta,\gamma)}{J(\beta,\gamma)}.$$

Here, we would like to point out that we could not provide a tight analysis for cost function involving average maximum age as we did in the deterministic case. This is due to the intricate coupling among queue-lengths and age. Nevertheless, the PoA performance under the considered cost function can still provide an upper bound on that with a cost function involving average maximum age, despite its looseness. In Section VII, we provide numerical results on the PoA performance with cost functions involving both average maximum age and average total age.

Next, we state our second key result for the stochastic cases as follows:

Theorem 2 (Joint AoI-Congestion PoA of Stochastic Cases)**.** *If $\lambda < \mu_\Sigma$, then there exists an $\epsilon > 0$ satisfying $\mu_n/\lambda \geq \mu_n/\mu_\Sigma + \epsilon/N, \forall n = 1, 2, \ldots, N$. In such a case, the users' selfishness and rationality yields the following PoA performance:*

$$\rho(\beta,\gamma) \leq \frac{B(\gamma) - \gamma M + p_{\max}}{B(\gamma) + \beta\left(\frac{N}{q} - 1\right) + p_{\max}} + \frac{\beta\left(\frac{N}{q} - \frac{1}{2q\mu_{\max}}\sum_{n=1}^{N}\mu_n - \frac{1}{2}\right)}{B(\gamma) + \beta\left(\frac{N}{q} - 1\right) + p_{\max}}, \tag{10}$$

where $B(\gamma) \triangleq \frac{\gamma}{2\lambda}\left(\mathbb{E}[A^2[t]] + \sum_{n=1}^{N}\mathbb{E}[R_n^2[t]]\right)$, $M \triangleq \frac{\epsilon}{2N(\mu_\Sigma - \lambda)}\left(Var(A[t]) + \sum_{n=1}^{N} Var(R_n[t]) + (\mu_\Sigma - \lambda)^2\right) - \frac{1}{2}\epsilon R_{\max}$, and $\mu_{\max} \triangleq \max_n \mu_n$.

Proof. Similar to the proof of Theorem 1, we first find an upper bound on $J(\beta,\gamma)$ by using the Lyapunov drift analysis and

[4] In this paper, we say that a queue n is *stable* if its average queue-length is finite, i.e., $\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[Q_n[t]] < \infty$. A system is *stable* if all its queues are stable.

[5] This is also motivated by the fact that the service provider prefers a better AoI for the PoI with a faster service rate.

then determine a fundamental lower bound on $J^{(\mathrm{OPT})}(\beta, \gamma)$. The detailed proof is available in Appendix B. □

Remark 2. From Theorem 2, we can see that for any fixed γ value, we have

$$\lim_{\beta \to \infty} \rho(\beta, \gamma) \leq 1 - \frac{1}{2} \frac{\frac{1}{q\mu_{\max}} \sum_{n=1}^{N} \mu_n - 1}{\frac{N}{q} - 1},$$

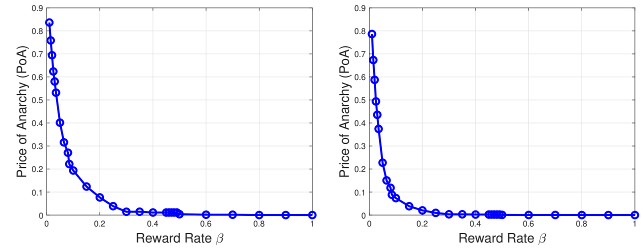
whose upper bound is equal to $1/2$ in the case with symmetric services, i.e., $\mu_1 = \mu_2 = \cdots = \mu_N$. However, we shall see from the numerical results presented in Section VII that for any fixed γ value, as β increases, the PoA actually converges to zero in the case with symmetric services. The looseness of the upper bound analysis is due to the intrinsic nature of the Lyapunov analysis methodology, which only captures the drift among neighboring slots in temporal domain and does not characterize the Round-Robin behavior in spatial domain.

VII. Numerical Results

In this section, we conduct simulations to study the PoA performance under users' selfishness and rationality (cf. (1)) in a mobile crowed-learning system. We use a 10-PoI system and assume that each PoI n's state information $p_n[t]$ belongs to the finite set $\{0.25, 0.5, 0.75, 1\}$, and $p_n[t]$ changes to a different value uniformly at random every 100 time slots. We consider both deterministic and stochastic cases. For the deterministic case, we assume that there is exactly one arriving user in each time slot and each PoI can serve one user in one time slot if any. For the stochastic case, we assume that users arrive at the system according to the Bernoulli distribution with mean $\lambda = 0.9$ and service provided by each PoI n follows an i.i.d. Bernoulli distribution with mean μ_n, $n = 1, 2, \ldots, 10$. We consider both symmetric and asymmetric services: For symmetric services, we let $\mu_n = 0.1, \forall n$; For asymmetric services, we let $\mu_n = 0.11, n \in \{1, 2, \ldots, 5\}$ and $\mu_n = 0.09, n \in \{6, 7, \ldots, 10\}$.

1) Deterministic Scenario: Fig. 4 illustrates the PoA performance in terms of both average maximum age and total average age in the deterministic case. In this case, there is no queueing effect and the PoA performance reflects the information freshness due to users' selfish and rational behavior compared to the optimal AoI performance. We can observe from Fig. 4a and Fig. 4b that PoA in terms of both average maximum age and total average age decrease as the reward rate β increases and roughly follows the $O(1/\beta)$ law, meaning that the AoI performance improves. This corroborates our analytical result on the PoA performance in Theorem 1. This simulation result also indicates that although the PoA analysis in Theorem 1 in terms of total average age is not asymptotically tight, the key message is true in the sense that the PoA improves as the reward rate β increases. Moreover, PoA in terms of both average maximum age and total average age decrease to zero for $\beta \geq 0.5$. This means that the AoI performance is optimal even with selfish and rational users.

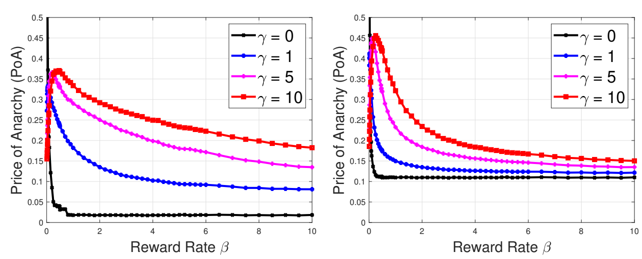
2) Stochastic Scenario: Next, we study the PoA performance with cost function involving average weighted sum of age (as defined in (9)) as well as average maximum



(a) (b)

Fig. 4: PoA with respect to reward rate β in the deterministic case: (a) PoA in terms of average maximum age; (b) PoA in terms of average total age.

age in stochastic cases. We consider both symmetric and asymmetric services. Here, PoA reflects the gap between joint AoI-congestion performance under users' selfishness and rationality compared to the optimal performance. We note that, even without incorporating AoI, it remains an open problem to find an optimal policy to minimize the total mean queue-length. In deriving the upper bound on PoA, we use the fundamental lower bound on total mean queue-length (cf. [28, Lemma 5]), which may not be tight. In this simulation, we adopt the Join-the-Shortest-Queue (JSQ) policy (e.g., [28]) and use its mean queue-length to serve as a lower bound for the queueing component in PoA. This is because JSQ minimizes the total mean queue-length (see [29, Proposition 3]) in the case with Bernoulli arrival and symmetric Bernoulli services, and it is optimal (see [28]) in the case with general arrival and service processes in the heavy-traffic regime (i.e., arrival rate approaches the total service rate asymptotically). Since we also evaluate the PoA performance with cost function involving average maximum age, we need to know the fundamental lower bound on the average maximum weighted age. This can be easily derived from equation (25) that $\mathbb{E}\left[\max_n \mu_n \widehat{\Delta}_n^{(\mathrm{OPT})}\right] \geq \mu_\Sigma / q - \mu_{\max}$.



(a) (b)

Fig. 5: PoA with respect to β and γ in the stochastic case with asymmetric services: (a) PoA in terms of average maximum age; (b) PoA in terms of average total age.

Fig. 5 shows the PoA performance with cost function involving both average maximum age and total average age in the case with asymmetric services under different values of γ. From Fig. 5b, we can see that, for any fixed γ value, PoA converges to 0.1 instead of 0 as β increases. The main reason is that we adopt the weighted sum of mean age as the metric for information freshness, and the policy that achieves optimal information freshness is unknown. Thus, we use our derived fundamental lower bound on the weighted mean sum-age to replace the optimal value for the information freshness, which may render a loose bound on PoA. However, we point out that

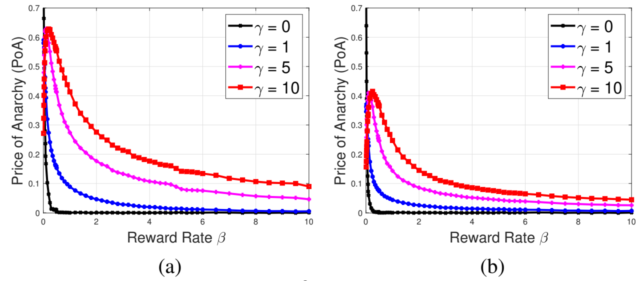


Fig. 6: PoA with respect to β and γ in the stochastic case with symmetric services: (a) PoA in terms of average maximum age; (b) PoA in terms of average total age.

our derived bound on PoA is tight in the symmetric service case as the reward rate β increases asymptotically, even though the derived upper bound of PoA is $1/2$ (cf. Remark 2). Indeed, we can observe from Fig. 6b that PoA actually converges to zero as β increases in the case with symmetric services. Despite that we did not provide any analytical results on the PoA with cost function involving average maximum age, we can see similar phenomenon from both Fig. 5a and Fig. 6a.

VIII. Conclusion

In this paper, we have strived to understand whether or not we can achieve information freshness guarantee with selfish and rational users in mobile crowd-learning. To answer this question, we first developed a new analytical model that takes into account the essential features of mobile crowd-learning. Then, based on this model, we showed that the natural greedy behavior of selfish and rational users could lead to AoI instability, which necessitates the design of reward mechanisms to induce information freshness guarantee. Toward this end, we proposed a linear AoI-based reward mechanism, under which we analyzed the impacts of users' selfishness and rationality on AoI based on the notion of Price of Anarchy (PoA). We showed that the proposed reward mechanism achieves bounded AoI and congestion performances in terms of PoA, and can even achieves optimal AoI asymptotically in a deterministic scenario. Collectively, these results serve as an exciting first step toward optimizing information freshness in mobile crowd-learning systems.

Appendix A
Proof of Theorem 1

Step 1): To find upper bounds on both average maximum age and average total age due to users' selfish and rational behavior, we perform Lyapunov drift analysis through an age-based Lyapunov function defined as follows:

$$V[t] \triangleq \sum_{n=1}^{N} \Delta_n[t]. \tag{11}$$

Let $\mathbf{M}[t] \triangleq (\{\Delta_n[t]\}_{n=1}^N, \{r_n[t]\}_{n=1}^N)$, and consider the one-step conditional expected drift of $V[t]$ as follows:

$$\begin{aligned}
\Delta V[t] \triangleq & \mathbb{E}\left[V[t+1]-V[t]|\mathbf{M}[t]\right] \\
=& \sum_{n=1}^{N} \mathbb{E}\left[\Delta_n[t+1]-\Delta_n[t]|\mathbf{M}[t]\right] \\
\overset{(a)}{=}& \sum_{n=1}^{N} \mathbb{E}\left[1-(\Delta_n[t]+1)S_n^*[t]|\mathbf{M}[t]\right] \\
\overset{(b)}{=}& N-1-\sum_{n=1}^{N} \mathbb{E}\left[\Delta_n[t]S_n^*[t]|\mathbf{M}[t]\right] \qquad (12) \\
\leq& N-1-\sum_{n=1}^{N} \mathbb{E}\left[\left(\Delta_n[t]-\frac{1}{\beta}r_n[t]\right)S_n^*[t]\middle|\mathbf{M}[t]\right], \qquad (13)
\end{aligned}$$

where (a) uses dynamics of $\Delta_n[t]$ in (6); (b) follows from the fact that each user joins one of the PoIs in each time slot, i.e., $\sum_{n=1}^{N} S_n^*[t]=1$.

Next, we first develop an upper bound on the average maximum age due to users' selfishness and rationality. As such, we consider the following irrational policy

$$\widetilde{\mathbf{S}}[t] \triangleq (\widetilde{S}_n[t])_{n=1}^N \in \arg\max_{\mathbf{S}} \sum_{n=1}^{N} \Delta_n[t] S_n[t],$$

i.e., users select the PoI with the largest age. Hence, (13) becomes

$$\begin{aligned}
\Delta V[t] \overset{(a)}{\leq}& N-1-\sum_{n=1}^{N} \mathbb{E}\left[\left(\Delta_n[t]-\frac{1}{\beta}r_n[t]\right)\widetilde{S}_n[t]\middle|\mathbf{M}[t]\right] \\
\overset{(b)}{\leq}& N-1-\Delta_{\max}[t]+\frac{1}{\beta}p_{\max}, \qquad (14)
\end{aligned}$$

where (a) follows from the definition of $S_n^*[t]$; and (b) uses the fact that $r_n[t] \leq p_{\max}, \forall n, t \geq 0$, the definition of $\widetilde{\mathbf{S}}[t]$, and the fact that exactly one $\widetilde{S}_n[t]$ is non-zero. It then follows from (13) that:

$$\mathbb{E}\left[V[t+1]-V[t]\right] \leq N-1-\mathbb{E}[\Delta_{\max}[t]]+\frac{1}{\beta}p_{\max}. \tag{15}$$

Summing (15) for $t=0,1,2,\ldots,T-1$, we obtain:

$$\mathbb{E}[V[T]-V[0]] \leq -\sum_{t=0}^{T-1} \mathbb{E}[\Delta_{\max}[t]]+(N-1)T+\frac{T}{\beta}p_{\max},$$

which implies that

$$\overline{\Delta}_{\max}^{(\beta)} \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}[\Delta_{\max}[t]] \leq N-1+\frac{1}{\beta}p_{\max}. \tag{16}$$

Finally, we develop an upper bound on the total average age due to users' selfishness and rationality. To that end, we consider a uniformly randomized policy $\check{\mathbf{S}}[t]$ with $\Pr\{\check{S}_n[t]=1\}=1/N$. As such, (13) becomes

$$\begin{aligned}
\Delta V[t] \overset{(a)}{\leq}& N-1-\sum_{n=1}^{N} \mathbb{E}\left[\left(\Delta_n[t]-\frac{1}{\beta}r_n[t]\right)\check{S}_n[t]\middle|\mathbf{M}[t]\right] \\
=& N-1-\sum_{n=1}^{N} \mathbb{E}\left[\left(\Delta_n[t]-\frac{1}{\beta}r_n[t]\right)\frac{1}{N}\middle|\mathbf{M}[t]\right] \\
\overset{(b)}{\leq}& N-1-\frac{1}{N}\sum_{n=1}^{N} \Delta_n[t]+\frac{1}{\beta}p_{\max}, \qquad (17)
\end{aligned}$$

Taking the expectation on both sides of the above inequality and using the similar telescoping technique in deriving the upper bound on the average maximum age, we can derive the following upper bound on the total average age.

$$\overline{\Delta}_{\Sigma}^{(\beta)} \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\Delta_n[t]] \leq N(N-1) + \frac{N}{\beta}p_{\max}.$$

Step 2): We first derive a fundamental lower bound on the average maximum age that can be achieved by the optimal policy. By using the same Lyapunov function in (11) to compute the conditional expected one-step drift under the optimal policy $\{S_n^{(\text{OPT})}[t]\}$ and following similar steps, we have

$$\Delta V[t] = N - 1 - \sum_{n=1}^{N}\mathbb{E}[\Delta_n^{(\text{OPT})}[t]S_n^{(\text{OPT})}[t]|\mathbf{M}[t]],$$

where $\Delta_n^{(\text{OPT})}[t]$ is the age of PoI n in time slot t under the optimal policy. In Step 1, we have already shown that the average maximum age is finite under the selfish and rational policy. This readily implies that the average maximum age is also finite under the optimal policy. Therefore, $\mathbb{E}[\Delta V[t]]$ will be equal to zero in steady-state and thus we have

$$\sum_{n=1}^{N}\mathbb{E}[\widehat{\Delta}_n^{(\text{OPT})}\widehat{S}_n^{(\text{OPT})}] = N-1, \tag{18}$$

where $\widehat{\Delta}_n^{(\text{OPT})}$ and $\widehat{S}_n^{(\text{OPT})}$ are random variables with the same distribution as $\Delta_n^{(\text{OPT})}[t]$ and $S_n^{(\text{OPT})}[t]$ in steady-state under the optimal policy, respectively. Hence, we have

$$\begin{aligned}\overline{\Delta}_{\max}^{(\text{OPT})} &\overset{(a)}{=} \mathbb{E}[\widehat{\Delta}_{\max}^{(\text{OPT})}] \overset{(b)}{=} \mathbb{E}[\max_n \widehat{\Delta}_n^{(\text{OPT})}]\\ &\overset{(c)}{\geq} \sum_{n=1}^{N}\mathbb{E}[\widehat{\Delta}_n^{(\text{OPT})}\widehat{S}_n^{(\text{OPT})}] \overset{(d)}{=} N-1, \end{aligned}\tag{19}$$

where step (a) follows from the boundedness of the average maximum age under the optimal policy; (b) is true for $\widehat{\Delta}_{\max}^{(\text{OPT})} \triangleq \max_n \widehat{\Delta}_n^{(\text{OPT})}$; (c) follows from the fact that each arriving user joins exactly one of the PoIs, i.e., $\sum_{n=1}^{N}\widehat{S}_n^{(\text{OPT})} = 1$; and (d) uses (18).

The fundamental lower bound on the total average age that can be achieved by the optimal policy is expressed as follows:

$$\overline{\Delta}_{\Sigma}^{(\text{OPT})} = \sum_{n=1}^{N}\mathbb{E}\left[\widehat{\Delta}_n^{(\text{OPT})}\right] \geq \frac{1}{2}N(N-1). \tag{20}$$

Its derivation is a special case of that in Step 2 in the proof of Theorem 2 and thus is omitted here for brevity.

Note that the lower bounds of both $\overline{\Delta}_{\max}^{(\text{OPT})}$ and $\overline{\Delta}_{\Sigma}^{(\text{OPT})}$ are *tight* and can be achieved by the Round-Robin policy, i.e., the system guides each arriving user to the PoIs in a Round-Robin fashion. Indeed, under Round-Robin, the ages of PoIs are a permutation of $\{0,1,2,\ldots,N-1\}$ in each time slot, and hence the maximum age and total age under Round-Robin are $\Delta_{\max}^{(\text{RR})}[t] = N-1$ and $\Delta_{\Sigma}^{(\text{RR})}[t] = 0+1+2+\ldots+(N-1) = N(N-1)/2$, $\forall t \geq 0$, respectively, which imply that $\overline{\Delta}_{\max}^{(\text{RR})} = N-1 = \overline{\Delta}_{\max}^{(\text{OPT})}$ and $\overline{\Delta}_{\Sigma}^{(\text{RR})} = N(N-1)/2 = \overline{\Delta}_{\Sigma}^{(\text{OPT})}$.

Lastly, by dividing the upper bounds in Step 1 by the lower bounds in Step 2, the desired PoA results in Theorem 1 follow and the proof is complete.

APPENDIX B
PROOF OF THEOREM 2

Step 1): Consider the following Lyapunov function:

$$L[t] \triangleq \frac{\gamma}{2\lambda\beta}\sum_{n=1}^{N}Q_n^2[t] + \frac{1}{q}\sum_{n=1}^{N}\Delta_n[t].$$

Let $\mathbf{Z}[t] \triangleq ((Q_n[t])_{n=1}^N, (\Delta_n[t])_{n=1}^N, (r_n[t])_{n=1}^N)$. Then, the one-step conditional expected drift can be computed as:

$$\begin{aligned}\Delta L[t] \triangleq& \mathbb{E}\left[L[t+1]-L[t]|\mathbf{Z}[t]\right]\\ =&\mathbb{E}\Bigg[\frac{\gamma}{2\lambda\beta}\sum_{n=1}^{N}\left(Q_n^2[t+1]-Q_n^2[t]\right)\\ &+\frac{1}{q}\sum_{n=1}^{N}(\Delta_n[t+1]-\Delta_n[t])\Bigg|\mathbf{Z}[t]\Bigg]\\ \overset{(a)}{\leq}& \frac{B(\gamma)}{\beta} + \mathbb{E}\Bigg[\frac{\gamma}{\lambda\beta}\sum_{n=1}^{N}Q_n[t](A[t]S_n^*[t]-R_n[t])\\ &+\frac{1}{q}\sum_{n=1}^{N}\left(1-(\Delta_n[t]+1)S_n^*[t]\mathbb{1}_{\{A[t]>0\}}\right)\Bigg|\mathbf{Z}[t]\Bigg],\end{aligned}$$

where (a) is true for $B(\gamma) = \frac{\gamma}{2\lambda}(\mathbb{E}[A^2[t]] + NR_{\max}^2) < \infty$, and uses dynamics of $Q_n[t]$ (cf. (2)) and $\Delta_n[t]$ (cf. (3)), and the fact that $(\max\{x,0\})^2 \leq x^2, \forall x$.

Next, we let $\mathbf{Z}'[t] \triangleq (\mathbf{Z}[t], \mathbb{1}_{\{A[t]>0\}})$. Then, for any function $f(\mathbf{Z}[t])$, the following sequence of equalities holds:

$$\begin{aligned}\mathbb{E}\{f(\mathbf{Z}[t])A[t]|\mathbf{Z}[t]\} =& q\mathbb{E}\{f(\mathbf{Z}[t])A[t]|\mathbf{Z}'[t]\}\\ =& q\mathbb{E}\{A[t]|A[t]>0\}\mathbb{E}\{f(\mathbf{Z}[t])|\mathbf{Z}'[t]\}\\ =&\mathbb{E}\{A[t]\}\mathbb{E}\{f(\mathbf{Z}[t])|\mathbf{Z}'[t]\}.\end{aligned}\tag{21}$$

Note that each arriving user joins one of the PoIs in each time slot, i.e., $\sum_{n=1}^{N}S_n^*[t] = 1$. Also, the users' decisions $\mathbf{S}^*[t]$ only depend on $\mathbf{Z}[t]$. Hence, we have that:

$$\begin{aligned}\Delta L[t] \overset{(a)}{\leq}& \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 - \frac{\gamma}{\lambda\beta}\sum_{n=1}^{N}\mu_n Q_n[t]\\ &+\mathbb{E}\left[\frac{\gamma}{\beta}\sum_{n=1}^{N}Q_n[t]S_n^*[t] - \sum_{n=1}^{N}\Delta_n[t]S_n^*[t]\middle|\mathbf{Z}'[t]\right]\\ \leq& \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 - \frac{\gamma}{\lambda\beta}\sum_{n=1}^{N}\mu_n Q_n[t]\\ &-\mathbb{E}\left[\sum_{n=1}^{N}\left(\Delta_n[t]-\frac{\gamma}{\beta}Q_n[t]-\frac{1}{\beta}r_n[t]\right)S_n^*[t]\middle|\mathbf{Z}'[t]\right],\end{aligned}\tag{22}$$

where (a) follows from (21) and the fact that $q \in (0,1)$. Next, consider an unselfish stationary randomized policy with $\mathbb{E}[\widetilde{S}_n[t]] = \mu_n/\mu_\Sigma$, $\forall n$, if $A[t] > 0$, and $\mu_\Sigma \triangleq \sum_{n=1}^{N}\mu_n$. Clearly, from the definition of $\mathbf{S}^*[t]$, we have:

$$\begin{aligned}(22) \leq& \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 - \frac{\gamma}{\beta}\sum_{n=1}^{N}\frac{\mu_n}{\lambda}Q_n[t]\\ &-\mathbb{E}\left[\sum_{n=1}^{N}\left(\Delta_n[t]-\frac{\gamma}{\beta}Q_n[t]-\frac{1}{\beta}r_n[t]\right)\widetilde{S}_n[t]\middle|\mathbf{Z}'[t]\right],\end{aligned}$$

Noting that $\mu_n/\lambda \geq \mu_n/\mu_\Sigma + \epsilon/N, \forall n$, we have

$$\Delta L[t] \overset{(a)}{\leq} \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 + \frac{p_{\max}}{\beta} - \frac{\gamma\epsilon}{N\beta}\sum_{n=1}^{N} Q_n[t]$$
$$- \frac{\gamma}{\beta}\sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}Q_n[t] - \sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\mathbb{E}\left[\left(\Delta_n[t] - \frac{\gamma}{\beta}Q_n[t]\right)\middle|\mathbf{Z}'[t]\right]$$
$$= -\frac{\gamma\epsilon}{N\beta}\sum_{n=1}^{N} Q_n[t] - \sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\Delta_n[t] + \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 + \frac{p_{\max}}{\beta},$$

where (a) follows from $r_n[t] \leq p_{\max}, \forall n, t$, and the definition of the stationary randomized policy $\{\widetilde{\mathbf{S}}[t]\}_{t\geq 0}$. This implies

$$\mathbb{E}\left[L[t+1] - L[t]\right] \leq -\frac{\gamma\epsilon}{N\beta}\sum_{n=1}^{N}\mathbb{E}[Q_n[t]] - \sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\mathbb{E}[\Delta_n[t]]$$
$$+ \frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 + \frac{p_{\max}}{\beta}. \tag{23}$$

Summing (23) for $t = 0, 1, 2, \ldots, T-1$, we obtain

$$\mathbb{E}[L[T] - L[0]] \leq -\frac{\gamma\epsilon}{N\beta}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}[Q_n[t]]$$
$$-\sum_{t=0}^{T-1}\sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\mathbb{E}[\Delta_n[t]] + \left(\frac{B(\gamma)}{\beta} + \frac{N}{q} - 1 + \frac{p_{\max}}{\beta}\right)T,$$

which further implies the following upper bound on $J(\beta, \gamma)$:

$$J(\beta,\gamma)$$
$$\triangleq \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\left[\frac{\gamma\epsilon}{N}\sum_{n=1}^{N}\mathbb{E}[Q_n[t]] + \beta\sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\mathbb{E}[\Delta_n[t]]\right]$$
$$\leq B(\gamma) + \beta\left(\frac{N}{q} - 1\right) + p_{\max}. \tag{24}$$

Step 2): Next, we derive a fundamental lower bound on $J^{(\mathrm{OPT})}(\beta, \gamma)$. Since we have shown that $J(\beta, \gamma)$ is upper-bounded under the selfish and rational policy in Step 1, $J^{(\mathrm{OPT})}(\beta, \gamma)$ is also bounded under the optimal policy. Therefore, we have

$$J^{(\mathrm{OPT})}(\beta,\gamma) = \frac{\gamma\epsilon}{N}\sum_{n=1}^{N}\mathbb{E}[\widehat{Q}_n^{(\mathrm{OPT})}] + \beta\sum_{n=1}^{N}\frac{\mu_n}{\mu_\Sigma}\mathbb{E}[\widehat{\Delta}_n^{(\mathrm{OPT})}],$$

where $\widehat{Q}_n^{(\mathrm{OPT})}$ and $\widehat{\Delta}_n^{(\mathrm{OPT})}$ are random variables with the same distribution as $Q_n[t]$ and $\Delta_n[t]$ in steady-state under the optimal policy, respectively. Next, we lower-bound $\sum_{n=1}^{N}\mathbb{E}[\widehat{Q}_n^{(\mathrm{OPT})}]$ and $\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n^{(\mathrm{OPT})}]$ individually. In the rest of the proof, we omit the signifier "(OPT)" for notational convenience and better readability.

We first consider $\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n]$. By choosing the Lyapunov function $V_1[t] \triangleq \sum_{n=1}^{N}\mu_n\Delta_n[t]$ and following similar steps as in the derivation of (12), we have

$$\Delta V_1[t] \triangleq \mathbb{E}\left[V_1[t+1] - V_1[t]|\mathbf{Z}'[t]\right]$$
$$= \mu_\Sigma - q\sum_{n=1}^{N}\mu_n\mathbb{E}\{S_n[t]|\mathbf{Z}'[t]\}$$
$$- q\sum_{n=1}^{N}\mu_n\mathbb{E}\{\Delta_n[t]S_n[t]|\mathbf{Z}'[t]\}.$$

Since $J^{(\mathrm{OPT})}(\beta, \gamma)$ is bounded under the optimal policy, the weighted sum of average age must also be finite under the optimal policy. Therefore, one can conclude that $\mathbb{E}[\Delta V_1[t]] = 0$ in steady-state. It then follows that:

$$\sum_{n=1}^{N}\mu_n\mathbb{E}\left[\widehat{\Delta}_n\widehat{S}_n\right] = \frac{1}{q}\mu_\Sigma - \sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{S}_n], \tag{25}$$

where $\widehat{S}_n$ is the random variable with the same distribution as $S_n[t]$ in the steady-state under the optimal policy.

Similarly, using Lyapunov function $V_2[t] \triangleq \sum_{n=1}^{N}\mu_n\Delta_n^2[t]$ and setting its drift to zero in steady-state yields:

$$2\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n] = q\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n^2\widehat{S}_n] + q\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n\widehat{S}_n]. \tag{26}$$

For any sample path, by Cauchy-Schwarz's Inequality, we have

$$\left(\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n\widehat{S}_n\right)^2 = \left(\sum_{n=1}^{N}\sqrt{\mu_n\widehat{S}_n}\cdot\sqrt{\mu_n\widehat{S}_n}\widehat{\Delta}_n\right)^2$$
$$\leq \left(\sum_{n=1}^{N}\mu_n\widehat{S}_n\right)\left(\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n^2\widehat{S}_n\right), \tag{27}$$

which implies $\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n^2\widehat{S}_n \geq \frac{(\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n\widehat{S}_n)^2}{\sum_{n=1}^{N}\mu_n\widehat{S}_n}$, and hence

$$\mathbb{E}\left[\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n^2\widehat{S}_n\right] \geq \mathbb{E}\left[\frac{\left(\sum_{n=1}^{N}\mu_n\widehat{\Delta}_n\widehat{S}_n\right)^2}{\sum_{n=1}^{N}\mu_n\widehat{S}_n}\right]. \tag{28}$$

Since $f(X, Y) = X^2/Y$ is convex for all $X \geq 0$ and $Y > 0$, by using Jensen's Inequality, we have $\mathbb{E}[\frac{X^2}{Y}] \geq \frac{(\mathbb{E}[X])^2}{\mathbb{E}[Y]}$. Thus, setting $X = \sum_{n=1}^{N}\mu_n\widehat{\Delta}_n\widehat{S}_n$ and $Y = \sum_{n=1}^{N}\mu_n\widehat{S}_n$, inequality (28) becomes:

$$\sum_{n=1}^{N}\mu_n\mathbb{E}\left[\widehat{\Delta}_n^2\widehat{S}_n\right] \geq \frac{\left(\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n\widehat{S}_n]\right)^2}{\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{S}_n]}. \tag{29}$$

By combining (25), (26) and (29), we have:

$$\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{\Delta}_n] \geq \frac{\mu_\Sigma}{2}\left[\frac{\mu_\Sigma}{q\sum_{n=1}^{N}\mu_n\mathbb{E}[\widehat{S}_n]} - 1\right]$$
$$\geq \frac{\mu_\Sigma}{2}\left[\frac{\mu_\Sigma}{q\mu_{\max}} - 1\right], \tag{30}$$

where the last step is true for $\mu_{\max} \triangleq \max_n \mu_n$.

In order to lower-bound $\sum_{n=1}^{N}\mathbb{E}[\widehat{Q}_n]$, we construct a hypothetical single-server queue $\{\Phi[t]\}$ with the same arrival process $\{A[t]\}_{t\geq 0}$ and an aggregated service process $\{R_\Sigma[t]\}_{t\geq 0}$, where $R_\Sigma[t] \triangleq \sum_{n=1}^{N}R_n[t]$. The queue-length evolution of this single-server queue can be written as: $\Phi[t+1] = \max\{\Phi[t] + A[t] - R_\Sigma[t], 0\}$. Due to resource pooling, the constructed hypothetical single-server's queue-length $\{\Phi[t]\}_{t\geq 0}$ is stochastically smaller than $\{\sum_{n=1}^{N}Q_n[t]\}_{t\geq 0}$ under any feasible policy. Hence, by [28, Lemma 5], we immediately have the following lower bound:

$$\sum_{n=1}^{N}\mathbb{E}[\widehat{Q}_n] \geq \frac{MN}{\epsilon}, \tag{31}$$

where $M \triangleq \frac{\epsilon}{2N(\mu_\Sigma - \lambda)}\big(\mathrm{Var}(A[t]) + \sum_{n=1}^{N}\mathrm{Var}(R_n[t]) + (\mu_\Sigma - \lambda)^2\big) - \frac{1}{2}\epsilon R_{\max}$. Lastly, combining (30), (31), and (24) yields the desired result in Theorem 2 and the proof is complete.

References

[1] Waze Mobile App. [Online]. Available: https://www.waze.com/
[2] GasBuddy Mobile App. [Online]. Available: https://www.gasbuddy.com/
[3] Pavemint Mobile App. [Online]. Available: https://www.pavemint.com/
[4] WiFi Finder Connect Internet, Mobile App. [Online]. Available: https://itunes.apple.com/us/app/wifi-finder-connect-internet/id1011519183?mt=8
[5] Basket – Grocery Shopping Mobile App. [Online]. Available: https://itunes.apple.com/us/app/basket-grocery-shopping/id1060139875?mt=8
[6] S. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. CISS*, Princeton, NJ, USA, March 2012, pp. 1–6.
[7] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE ISIT*, Cambridge, MA, USA, July 2012, pp. 2666–2670.
[8] S. Kaul, R. D. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Orlando, FL, USA, March 2012, pp. 2731–2735.
[9] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. IEEE SECON*, Salt Lake City, UT, USA, June 2011, pp. 350–358.
[10] "Creek watch," 2010. [Online]. Available: http://creekwatch.researchlabs.ibm.com/
[11] N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels, "Noisetube: Measuring and mapping noise pollution with mobile phones." Springer, 2009, pp. 215–228.
[12] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, and W. Hu, "Ear-phone: an end-to-end participatory urban noise mapping system," in *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*. ACM, 2010, pp. 105–116.
[13] D. Zhang, H. Xiong, L. Wang, and G. Chen, "Crowdrecruiter: selecting participants for piggyback crowdsensing under probabilistic coverage constraint," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 703–714.
[14] J. Liu, H. Shen, and X. Zhang, "A survey of mobile crowdsensing techniques: A critical component for the internet of things," in *Proc. IEEE ICCCN*, 2016.
[15] Y. Chen, J. Zhou, and M. Guo, "A context-aware search system for internet of things based on hierarchical context model." *Telecommunication Systems*, vol. 62, no. 1, pp. 77–91, 2016.
[16] K. Han, H. Huang, and J. Luo, "Posted pricing for robust crowdsensing," in *Proc. ACM MobiHoc*, 2016, pp. 261 – 270.
[17] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. IEEE ISIT*, June 2017, pp. 321–325.
[18] X. Gao, E. Akyol, and T. Basar, "Optimal estimation with limited measurements and noisy communication," in *Proc. IEEE CDC*, Osaka, Japan, December 2015, pp. 1775–1780.
[19] E. T. Ceran, D. Gunduz, and A. Gyorgy. (2017, October) Average age of information with hybrid arq under a resource constraint. [Online]. Available: https://arxiv.org/abs/1710.04971v1
[20] R. D. Yates, E. Najm, E. Soljanin, and J. Zhong, "Timely updates over an erasure channel," in *Proc. IEEE ISIT*, 2017.
[21] C. Kam, S. Kompella, G. Nguyen, J. Wieselthier, and A. Ephremides, "Information freshness and popularity in mobile caching," in *Proc. IEEE ISIT*, June 2017, pp. 136–140.
[22] R. D. Yates, P. Ciblat, A. Yener, and M. Wigger, "Age-optimal constrained cache updating," in *Proc. IEEE ISIT*, June 2017, pp. 141–145.
[23] 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 Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, November 2017.
[24] S. K. Kaul and R. D. Yates, "Status updates over unreliable multiaccess channels," in *Proc. IEEE ISIT*, 2017, pp. 331–335.
[25] B. Li, A. Eryilmaz, and R. Srikant, "On the universality of age-based scheduling in wireless networks," in *Proc. IEEE INFOCOM*, Kowloon, Hong Kong, April 2015, pp. 1302–1310.
[26] 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*, 2018, pp. 191–200.
[27] N. B. Lakshminarayana, J. Lee, and H. Kim, "Age based scheduling for asymmetric multiprocessors," in *Conference on High Performance Computing Networking, Storage and Analysis*, Portland, OR, USA, November 2009.
[28] A. Eryilmaz and R. Srikant, "Asymptotically tight steady-state queue length bounds implied by drift conditions," *Queueing Systems*, vol. 72, no. 3-4, pp. 311–359, 2012.
[29] B. Li, A. Eryilmaz, R. Srikant, and L. Tassiulas, "On optimal routing in overloaded parallel queues," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE, 2013, pp. 262–267.

Bin Li (S'12–M'16–SM'20) received his B.S. degree in Electronic and Information Engineering in 2005, M.S. degree in Communication and Information Engineering in 2008, both from Xiamen University, China, and Ph.D. degree in Electrical and Computer Engineering from The Ohio State University in 2014. Between 2014 and 2016, he worked as a Postdoctoral Researcher in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. He is currently an Assistant Professor in the Department of Electrical, Computer, and Biomedical Engineering at the University of Rhode Island. His research focuses on the intersection of networking, machine learning, and system developments, and their applications in networking for virtual/augmented reality, mobile edge computing, mobile crowd-learning, and Internet-of-Things. He is a senior member of the IEEE and a member of the ACM. He received both the National Science Foundation (NSF) CAREER Award and Google Faculty Research Award in 2020, and ACM MobiHoc 2018 Best Poster Award.

Jia Liu (S'03–M'10–SM'16) received his Ph.D. degree in the Bradley Department of Electrical and Computer Engineering at Virginia Tech, Blacksburg, VA in 2010. He is currently an Assistant Professor in the Dept. of Electrical and Computer Engineering at The Ohio State University, where he joined in Aug. 2020. From Aug. 2017 to Aug. 2020, he was an Assistant Professor in the Dept. of Computer Science at Iowa State University. His research areas include theoretical machine learning, control and optimization for stochastic networks, and optimization for data analytics infrastructure and cyber-physical systems. Dr. Liu is a senior member of IEEE and a member of ACM. He has received numerous awards at top venues, including IEEE INFOCOM'19 Best Paper Award, IEEE INFOCOM'16 Best Paper Award, IEEE INFOCOM'13 Best Paper Runner-up Award, IEEE INFOCOM'11 Best Paper Runner-up Award, and IEEE ICC'08 Best Paper Award. Dr. Liu is a recipient of the NSF CAREER Award in 2020. He is a recipient of the Google Faculty Research Award in 2020. He is also a winner of the LAS Award for Early Achievement in Research from the College of Liberal Arts and Sciences at Iowa State University in 2020, and the Bell Labs President Gold Award in 2001. His research is supported by NSF, AFOSR, AFRL, and ONR.