Age of Information based cache updating with popularity contents: Whittle's index based approach

Philippe Ciblat[1], Guiseppe Caire[2], Roy Yates[3]
(1) LTCI, Telecom Paris, Institut Polytechnique de Paris, Palaiseau, France
(2) TU Berlin, Berlin, Germany
(3) Rutgers University, New Brunswick, New Jersey, USA
philippe.ciblat@telecom-paris.fr

***Abstract*—We focus on the scheduling algorithm for updating files from a cloud server to a local server having cache. We consider that only K out of N files can be updated at each timeslot. Each file is time-sensitive and the content relevance is thus measured through the Age of Information. In addition, each file has its own popularity which is time-varying according to a Markovian model. In this paper, we offer two contributions: first, we exhibit Whittle's index for this scheduling problem when the popularity is known and fixed over time. Second, we propose a heuristic based on previous Whittle's index for the time-varying popularity case assuming that only the past popularity is available.**

I. Introduction

Some content such as newspaper websites, live TV programs, video streams, and data coming from sensors, are time-sensitive in the sense that they can become obsolete. Such data are less useful if obtained and/or viewed too late. A way to measure this obsolescence is the so-called *Age of Information* (AoI) proposed in [1].

We here consider the edge cache updating problem. We consider three levels in the network: i) users requiring time-sensitive contents; ii) an central distant server contains the entire fresh contents; iii) a cache, closed to the users and easy to access, containing the entire contents but at each timeslot, the cache may be updated only partially. Therefore the freshness of files will vary over time. In this context, the goal is to find a relevant cache updating policy that specifies which contents to update from the server to the cache at a given timeslot in order to minimize a criterion related to the AoI.

Several simple updating policies can be developed, for example: i) the famous Round-Robin (RR) scheduling where each content is updated according to a cycle; ii) the scheduling updating the content with the highest age is another well-known approach. However, these policies assume that the contents are equivalent and that the users' requests are uniformly distributed over the contents. In practice, some content items are more popular than others according to an empirical law, such as, the Zipf distribution [2]. When popularity is content-dependent, the updating policy should take this into account and should provide a relevant trade off between the AoI of the content and its popularity. Indeed, a popular content item should be updated more often than a less popular content but less popular context should also be updated (certainly with a smaller rate) in order to keep its AoI with a bounded value. When the popularity is known and fixed over time, the square-root law is advocated in [3].

The popularity can be also time-varying since some contents can be more or less popular over time such as a streamed video or some online newspapers' articles [4]. In that context, [4] proposed to learn the popularity online. In [5], the popularity, even if time-varying, was known in advance. In contrast, [6] has developed a policy for the case when the current popularity is not known in advance. Based on a Markovian model of the popularity rate, they proposed an approximation of the optimal policy using standard tools of Constrained Markov Decision Processes (CMDP). The main drawback of [6] is the un-interpretability of the obtained policy; they obtained a look-up table as a black box.

One way to avoid an un-interpretable policy in a CMDP is to focus on index-based policies. The most common index is the so-called Whittle's index (WI) [7]. Some WIs for AoI based scheduling policies have been proposed but never by taking into account the popularity [8]–[10]. Actually considering the popularity in addition to AoI is challenging when the popularity is time-varying since the underlying Markovian state per user is then two-dimensional. And expressing WI in closed-form is not always possible, as mentioned in [11]. Indeed, it was seen in [11] that some rewards have to be specifically ranked to ensure the existence of the WI. In our context, we have observed that such conditions were not fulfilled.

Therefore the contributions of this paper are twofold: first, we exhibit in closed-form the WI for cache update scheduling when the popularity is time-invariant Based on the obtained expression of the WI under time-invariant popularity, we propose a heuristic when the popularity is time-varying and not known in advance. We observe that the proposed heuristic is interpretable and close in performance to the policy offered in [6].

The paper is organized as follows: in Section II we introduce the mathematical model. In Section III, we derive analytically the WI when the popularity is fixed over time. In Section IV, we extend the WI as a heuristic to the time-varying popularity

This work was partly supported by ERC Grant "CTOCom". We thank Dr. Urtzi Ayesta for his fruitful advices.

case. Numerical results are provided in Section V. Concluding remarks are drawn in Section VI.

II. Problem definition

We consider a central distant server with N time-sensitive files. At each timeslot $t \in \{1, \ldots, T\}$, the cache can update only K current versions of files from the server due to a bandwidth constraint. For each file $n \in \{1, \ldots, N\}$, let $\{u_{n,t}\} \in \{0, 1\}$ be the download decision at time t, i.e., $u_{n,t} = 1$, if file n is downloaded at timeslot t, and $u_{n,t} = 0$ otherwise. The bandwidth constraint requires that

$$\sum_{n=1}^{N} u_{n,t} \leq K, \quad \forall t. \tag{1}$$

We denote by $X_{n,t}$ the AoI of file n at timeslot t, i.e., the number of timeslots that have passed since the cache has downloaded file n. For sake of simplicity, we assume that $X_{n,1} = 1, \forall n$. Afterwards for $t > 1$, the AoI is modified as follows:

$$X_{n,t+1} = \begin{cases} 1, & u_{n,t} = 1; \\ X_{n,t} + 1, & u_{n,t} = 0. \end{cases} \tag{2}$$

In each timeslot t, the number of requests for file n depends on its current popularity mode $m_{n,t} \in \mathcal{M} := \{1, \cdots, M\}$ as in [5], [6]. The expected number of requests of file n is determined by a function $p_n : \mathcal{M} \to \mathbb{R}^+$, so that the expected number of requests for file n at timeslot t is given by $p_n(m_{n,t})$. It is assumed that for each file n, the sequence $\{m_{n,t}\}_{t=1}^{T}$ evolves according to an M-state Markov chain with transition probabilities $T^n_{m,m'} := \Pr(M_{n,t+1} = m' | M_{n,t} = m)$ for $(m, m') \in \mathcal{M}^2$. In [3], $T^n_{m,m'} = \delta_{m,m'}$ with δ the Kronecker index, and $p_n(m_{n,t})$ does not depend on t anymore and corresponds to the popularity rate directly denoted by p_n.

Let Π be a set of cache updating policies, i.e., policies providing the decisions $\{u_{n,t}\}$ at timeslot t which only depends on the current and past popularity modes $\{m_{n,\tau}\}_{\tau \leq t}$ and AoIs $\{X_{n,\tau}\}_{\tau \leq t}$ as well as on the statistics $\{T^n_{r,r'}\}$. The future popularity modes $\{m_{n,\tau}\}_{\tau > t}$ cannot be used. The goal is to design such policies $\pi \in \Pi$ related to the minimization of the expected total AoI of all requested files averaged over an infinite-time horizon described by this optimization problem.

Problem 1:

$$\min_{\pi \in \Pi} \lim_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} X_{n,t} p_n(m_{n,t}) \right], \tag{3a}$$

$$\text{s.t.} \sum_{n=1}^{N} u_{n,t} \leq K, \quad \forall t. \tag{3b}$$

III. Time-invariant popularities

In this section, we consider that the popularity mode is not time-varying. As a consequence, $p_n(m_{n,t})$ simplifies into $p_n(m_n)$ and only depends on n and can be denoted by p_n. Then p_n refers to as the popularity rate. Typically, the values of p_n are related to those of a Zipf distribution [2].

Here, we propose to exhibit in closed-form the WI related to Problem 1 when the popularity is not time-varying. To define the WI, we first relax Problem 1 having bandwidth constraints (3b) per timeslot into Problem 2 having average bandwidth constraints.

Problem 2:

$$\min_{\pi \in \Pi} \lim_{T \to \infty} \mathbb{E}_\pi \left[\sum_{t=1}^{T} \sum_{n=1}^{N} X_{n,t} p_n \right], \tag{4a}$$

$$\text{s.t.} \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{n=1}^{N} u_{n,t} - K \right) \leq 0. \tag{4b}$$

Then Problem 2 can be transformed again into a new Problem 3 thanks to the introduction of a Lagrangian multiplier λ for the average bandwidth constraints.

Problem 3:

$$\min_{\pi \in \Pi} \lim_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} (X_{n,t} p_n + \lambda u_{n,t}) \right]. \tag{5}$$

One can easily observe that Problem 3 can be decoupled file by file and we obtain a per-file Problem. This Problem 4 just corresponds to a unconstrained Markov Decision Process problem. The optimal policy does not directly take into account the other files but only through the parameter λ. For the sake of simplicity, we now omit the index n corresponding to the file. So p is now the time-invariant popularity for any current file of interest.

Problem 4:

$$\min_{\pi \in \Pi} \lim_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \sum_{t=1}^{T} (X_t p + \lambda u_t) \right]. \tag{6}$$

Before going further, we want to solve Problem 4. Problem 4 is MDP with the following ingredients per file:

- Action: $u(t)$ is 1 if the file is updated at timeslot t, and 0 otherwise. When $u(t) = 1$, the file is said to be *active*. When $u(t) = 0$, the file is said to be *passive*.
- State: $X(t)$ the age of the file at timeslot t. We force $X(1) = 1$. The update law is as follows

$$\begin{cases} \Pr(X(t+1) = 1 | X(t) = x, u(t) = 1) & = & 1 \\ \Pr(X(t+1) = x+1 | X(t) = x, u(t) = 0) & = & 1 \end{cases}. \tag{7}$$

- Cost: the instantaneous reward in Eq. (6) at timeslot t with age x and action u can be written as

$$c(x, u) = (x + 1 - xu)p + \lambda u$$

  where λ plays the role of the weight for doing an update.

According to [12], we know the optimal value $\mathcal{C}$ of the criterion given in Eq. (6) satisfies the following Bellman's equation

$$\mathcal{C} + f(x) = \min \left((x+1)p + \int f(y) \Pr(y|x, 0) dy, \; p + \lambda + \int f(y) \Pr(y|x, 1) dy \right) \tag{8}$$

where f is the so-called value function and $\Pr(y|x,u)$ is provided by Eq. (7). In addition, Problem 4 admits an optimal stationary deterministic policy denoted by $\pi_\lambda^\star : \mathbb{N}_+ \to \{0,1\}$ and given by

$$\pi_\lambda^*(x) = \arg\min_u c(x,u) + \int f(y)\Pr(y|x,u)dy.$$

Let us now define an index-based policy. Let us consider a real-valued function depending on some information state s of an agent (in this paper, the agent corresponds to a content/file), denoted by $s \mapsto I(s)$ and entitled "index". When K agents out of N have to be selected to make their action, an index-based policy selects the K agents with the highest indices. Mathematically speaking, we have

$$\mathcal{K} = \arg_K \max_{n\in\{1,\dots,N\}} I(s_n)$$

where s_n is the current state of agent n, $\arg_K \max$ stands for the operator selecting the K agents with the highest values, and $\mathcal{K}$ is the set of selected agents. When s is based on the Age of the agent (the age may correspond to the AoI, the delay, $\cdots$), the index policy may be called the *Schedule Ordered by Age-based Priority* (SOAP) and is well described in [13]. Let us now focus on the WI. This index is the λ in Eq. (8) such that the first term in the minimum operator is equal to its second term, i.e., both actions are equivalent. Such a λ only depends on the parameter p and the AoI x. This index will be denoted by $x \mapsto I_W(x)$. To find this index in our specific problem, we will proceed as in [9] where two steps are necessary:

Step S1- proof of optimality of a threshold policy for Problem 4. To do that, we need to show that by choosing appropriately the threshold of the considered threshold policy, such a policy satisfies the Bellman's equation. As a byproduct, we obtain the optimal long-term age $\mathcal{C}$. The threshold is denoted by $\mu^\star(\lambda)$.

Step S2- exhibition of λ such that $x+1=\mu^*(\lambda)$, since the age $\mu^*(\lambda)$ corresponds to the age where the previous policy equilikely decides the file to be active or passive. Then we obtain λ with respect to (wrt) x and identify it to the WI denoted by I_W.

Let us consider a threshold policy on the age with the threshold μ. We have

$$\left\{\begin{array}{lcl} x & \geq & \mu \text{ leads to active file} \\ x & < & \mu \text{ leads to passive file}\end{array}\right. .$$

Let us begin with *Step S1*. The goal is to prove that a threshold policy with a well-tuned threshold is optimal. We assume that $f(1)=0$. When $x<\mu$, the Bellman's equation (8) simplifies to

$$\mathcal{C} + f(x) = (x+1)p + f(x+1) \tag{9}$$

and this occurs when

$$xp + f(x+1) \leq \lambda. \tag{10}$$

When $x \geq \mu$, the Bellman's equation becomes

$$\mathcal{C} + f(x) = p + \lambda$$

and this occurs when

$$xp + f(x+1) > \lambda. \tag{11}$$

Finally, by iterating Eq. (9) wrt x from 1 until $x<\mu$, we get

$$f(x) = \frac{(\mu-x)(\mu-x+1)p}{2} + (\mu-x)(xp-\mathcal{C}) + f(\mu). \tag{12}$$

In contrast, when $x \geq \mu$, we have

$$f(x) = p + \lambda - \mathcal{C}, \tag{13}$$

and we remark that this function f is constant as soon as $x \geq \mu$.

Our objective now is to prove that such a value function f with an appropriate μ satisfies the Bellman's equation. If so, we prove the threshold policy with the above-mentioned appropriate μ is optimal. To do that, we need to check that we can connect the function at point μ since $f(\mu)$ is used and defined in both intervals (when $x<\mu$ and when $x \geq \mu$).

By applying $x=1$ in Eq. (12), we obtain

$$f(\mu) = (\mathcal{C}-p)(\mu-1) - \frac{(\mu-1)\mu p}{2}. \tag{14}$$

In addition, based on Eq. (13), we also obtain

$$f(\mu) = p + \lambda - \mathcal{C}. \tag{15}$$

Consequently, if Bellman's equation works for this policy, we need at least gather both Eqs. (14)-(15) which means that

$$\lambda = (\mathcal{C}-p)\mu - \frac{(\mu-1)\mu p}{2}. \tag{16}$$

We also have constraints on λ according to Eqs. (10)-(11). Around the threshold, we need at least satisfy

$$(\mu-1)p + f(\mu) \leq \lambda < \mu p + f(\mu+1), \tag{17}$$

which leads to

$$\mu p + \lambda - \mathcal{C} \leq \lambda < (\mu+1)p + \lambda - \mathcal{C}. \tag{18}$$

Eq. (18) is always true if

$$\frac{\mathcal{C}}{p} - 1 < \mu \leq \frac{\mathcal{C}}{p}. \tag{19}$$

This holds when there exists $\varepsilon \in [0,1)$ such that

$$\mu + \varepsilon = \frac{\mathcal{C}}{p}. \tag{20}$$

Combining Eqs. (16) and (20) by removing the optimal cost $\mathcal{C}$ leads to

$$\frac{p}{2}\mu^2 + (p\varepsilon - \frac{p}{2})\mu - \lambda \quad = \quad 0$$

This leads to an unique positive threshold given by

$$\mu_\varepsilon = -(\varepsilon - 1/2) + \sqrt{(\varepsilon-1/2)^2 + 2\lambda/p}.$$

We have $\mu_0 = 1/2 + \sqrt{1/4 + 2\lambda/p}$ and $\mu_1 = -1/2 + \sqrt{1/4+2\lambda/p}$. So $\mu_0 = \mu_1 + 1$. It is easy to prove that $\varepsilon \mapsto \mu_\varepsilon$ is a non-increasing function. As a consequence, there exists an unique ε^* such that μ_{ε^*} is an integer equal to $\lfloor \mu_0 \rfloor$. So we obtain

$$\mu^*(\lambda) = \left\lfloor \frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{\lambda}{p}} \right\rfloor. \tag{21}$$

For completing step S1, we need to prove that the value function f described above satisfies Eqs. (10)-(11). According to Eq. (17), we know that

$$\lambda < \mu p + f(\mu + 1).$$

As $f(x) = f(\mu+1)$ for any $x > \mu$ and $x \mapsto xp$ is increasing, Eq. (11) holds. According to Eq. (17), we know that

$$(\mu - 1)p + f(\mu) \leq \lambda$$

and we want prove that $xp + f(x+1) \leq \lambda$ for any $x \in \mathbb{N}$ such that $x < \mu$. Let $g(y) = yp + f(y+1)$ on the interval $y \in [1, \mu - 1]$. Due to Eq. (12), we have

$$g(y) = (\mu - y - 1)\left[\frac{yp}{2} + \frac{\mu p}{2} + p - \mathcal{C}\right] + yp + f(\mu),$$

which has derivative

$$g'(y) \quad = \quad -yp + \frac{p}{2} + \mathcal{C}.$$

This derivative is positive iff

$$y \leq \frac{1}{2} + \frac{\mathcal{C}}{p},$$

which holds via Eq. (19). Consequently $g(y) \leq g(\mu - 1) \leq \lambda$, and Eq. (10) holds. Therefore the threshold policy with the threshold given by Eq. (21) is optimal.

Let us move on *Step S2*. The WI $I_W(x)$ is equal to the λ such that the following equation holds:

$$x + 1 = \mu^*(\lambda) = \frac{1}{2} + \sqrt{\frac{1}{4} + 2\frac{\lambda}{p}}. \tag{22}$$

Solving Eq. (22) for the index λ leads to

$$I_W(x) = \frac{1}{2}px\,(x+1)\,. \tag{23}$$

We advocate for a simplified index relying on Eq. (23). We propose to neglect the $+1$. In addition, as index policy is based on index comparison, the factor $1/2$ in Eq. (23) is useless and comparing the index or its square-root leads to the same algorithm. Therefore we obtain Proposition 1.

Proposition 1: Based on the WI given on Eq. (23) and associated with Problem 2, we propose the following index policy at timeslot t

$$\hat{\mathcal{K}}(t) = \arg_K \max_{n \in \{1,\cdots,N\}} \sqrt{p_n} X_n(t)$$

where $\hat{\mathcal{K}}(t)$ corresponds to the set of K users out of N offering the K minimum values of the corresponding index.

IV. Time-varying popularities

With time-varying popularities, the state of the related MDP is bidimensional and defined by $S(t) = [X(t), p(m_t)]$ (once again, the file index is omitted). Deriving the WI for this new MDP is often not feasible, as mentioned in [11]. We can prove that the conditions given in [11] for the existence of the WI are not satisfied here. Therefore we just propose a heuristic based on the WI obtained in the time-invariant popularity case. We propose to replace in Proposition 1 the fixed popularity p_n by the predicted popularity at time $t+1$ given the popularity at time t. According to Markov chain property of popularity model, the predicted popularity is actually the conditional mean at time $t+1$ given the value of the popularity at time t and can be expressed as follows:

$$\tilde{p}_n(t+1) = \sum_{m=1}^{M} T^n_{m_{n,t},m} p_n(m).$$

We recall that the update done at timeslot t, will be requested at timeslot $t+1$, therefore the expected popularity at time $(t+1)$ is a relevant information when updating at time t. This leads to Proposition 2

Proposition 2: Based on Proposition 1, we propose the following index policy at timeslot t

$$\hat{\mathcal{K}}(t) = \arg_K \max_{n \in \{1,\cdots,N\}} \sqrt{\tilde{p}_n(t+1)} X_n(t).$$

V. Numerical results

In this Section, we numerically evaluate the performance of the proposed policy. Except when otherwise stated, we consider two popularity modes $\mathcal{M} = \{1, 2\}$, the transition matrix between both popularity modes is independent of file n, and is given by

$$\mathbf{T}^n = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}, \quad \forall n \in \{1, \ldots, N\}, \tag{24}$$

with $q \in (0, 1)$. So q corresponds to the probability to stay in the same mode. We consider $N = 64$ files.

In Fig. 1, we consider the special case of no-time-varying popularity with $K = 1$. We consider for the popularity rate a Zipf distribution ($p_n \propto 1/n^\alpha$) with coefficient $\alpha = 1.5$. We plot the probability update rate per file when the policy given in Proposition 1 is applied and the square-root probability distribution (distribution suggested in [3] for updating the file). We observe that the WI based policy offers the same distribution of update as the square-root law and that it is a way for mimicking the algorithm proposed in [3], and will finally offer remarkable performance as in [3].

From now, we consider time-varying popularities. We constrain the system to update only $K = 8$ files at a time. We first consider the expected number of requests of file n in the two states is $p_n(1) = 0.2\overline{p}_n$ and $p_n(2) = 1.8\overline{p}_n$, where $\overline{p}_n \propto 1/n^\alpha$ follows a Zipf distribution with coefficient $\alpha = 1.5$. According to Eq. (24), the steady-state probability for both popularities are identical, so $\overline{p}_n$ corresponds to the average number of requests for file n. In Fig. 2, we plot the average AoI versus

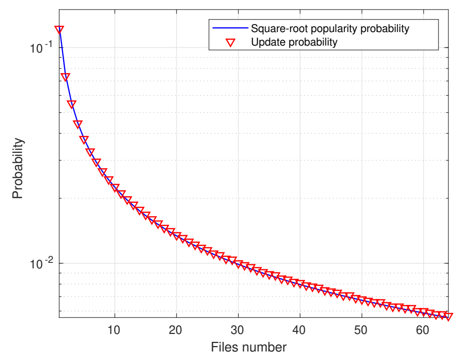

Fig. 1. Files update probability vs files numbering

q for four algorithms: i) the policy proposed in [6], ii) the proposed algorithm given in Proposition 1, iii) the WI of Proposition 1 with the current popularity $p_n m_{n,t}$, and iv) the WI of Proposition 1 with the average popularity $\overline{p}_n$. We remark that the proposed policy is close to the one given by [6] in terms of performance. But the proposed policy is much simpler to implement and interpretable while in [6], the policy is obtained after an optimization phase and the resulting policy has no interpretation since it has no closed-form expression. We also remark that the policy based on the current popularity fails even for middle value of q. In contrast, the policy based on the average popularity is insensitive to q and so offers good performance when q is one half, i.e., when the knowledge of the current popularity is not informative.

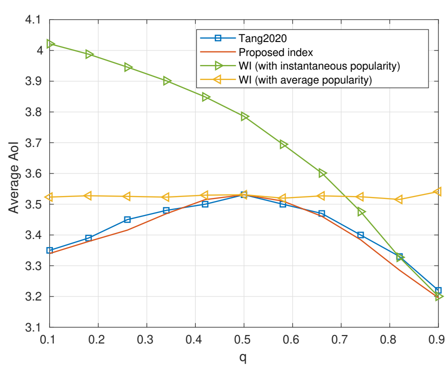

Fig. 2. Average AoI vs q for two specific popularity modes.

In Fig. 3, we plot the average AoI versus q but for different popularity modes: here, $p_n(1) = p_n$ where $_n$ follows a Zipf distribution with coefficient $\alpha = 1.5$, and $p_n(2) = 1/N$ for $n \in \{1, \cdots, N\}$. We consider the three last policies of Fig. 2. We can do the same remarks. The proposed policy offers the best performance and so is adapted to many types of modes. Actually, we also tested the following modes (not reported here): $p_n(1) = \overline{p}_n$ and $p_n(2) = \overline{p}_{N-n}$, where $\overline{p}_n \propto 1/n^\alpha$ follows a Zipf distribution with coefficient $\alpha = 1.5$. So the main advantage due to the closed-form expression given in Proposition 2 is that no additional computation has to be run for applying the proposed policy.

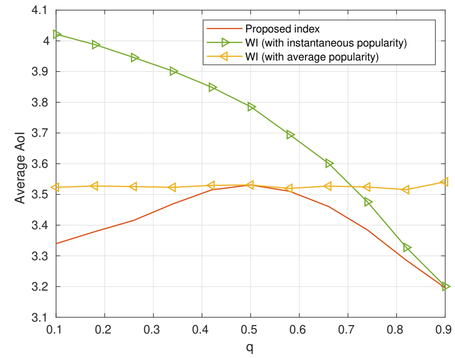

Fig. 3. Average AoI vs q for two other popularity modes.

VI. Conclusion

In this paper, we proposed an heuristic based on Whittle's index derivations for cache update scheduling when the Age of Information and the popularity of the content to update has to be taken into account. We have seen the proposed policy offers nice performance compared to simpler heuristic as well as a quasi-optimal solution provided in [6]. As a future work, such policies should be applied in a real context where the popularity variation does not follow the considered model.

References

[1] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," IEEE Infocom, 2012.

[2] S.A. Krashakov, A.B. Teslyuk, L.N. Shchur, "On the universality of rank distributions of website popularity," Computer Networks, August 2006.

[3] R. Yates, P. Ciblat, A. Yener, and M. Wigger, "Age-optimal constrained cache updating," IEEE International Symposium on Information Theory (ISIT), 2017.

[4] B.N. Bharath, K.G. Nagananda, D. Gündüz, and H.V. Poor, "Caching with time-varying popularity profiles: a learning-theoretic perspective," IEEE Transactions on Communications, September 2018.

[5] C. Tao, J. Luo, F.-C. Zheng, and L. Gao, "Freshness-aware cache update with time-varying popularities in edge networks," IEEE Infocom, 2024.

[6] H. Tang, P. Ciblat, M. Wigger, and R. Yates, "Cache Updating Strategy Minimizing the Age of Information with Time-Varying Files' Popularities," IEEE Information Theory Workshop (ITW), 2020.

[7] P. Whittle, "Scheduling and characterization problems for stochastic networks," Journal of the Royal Statistical Society, March 1985.

[8] Y.-P. Hsu, "Age of Information: Whittle index for scheduling stochastic arrivals," IEEE International Symposium on Information Theory (ISIT), 2018.

[9] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing Age of Information in broadcast wireless networks," IEEE/ACM Transactions on Networking, December 2018.

[10] S. Krioule, M. Assaad, and A. Maatouk, "On the optimality of the Whittle's index policy for minimizing the Age of Information," IEEE Transactions on Information Theory, January 2022.

[11] S. Aalto, P. Lassila, and P. Osti, "Whittle index approach to size-aware scheduling with time-varying channels," Sigmetrics, 2015.

[12] M.L. Puterman, "Markov Decision Processes: discrete stochastic dynamics," 1974.

[13] Z. Scully, M. Harchol-Balter, and A. Scheller-Wolf, "SOAP: one clean analysis of all age-based scheduling policies," ACM Measurement Analysis Computer Systems, March 2018.