Average Age-of-Information with a Backup Information Source

Elvina Gindullina[†], Leonardo Badia[†] and Deniz Gündüz[‡]

[†] University of Padova, Dept. of Information Engineering, via Gradenigo 6B, 35131 Padova, Italy
[‡] Imperial College London, Dept. of Electrical and Electronic Engineering, London SW7 2AZ, U.K.
email: {elvina.gindullina, leonardo.badia}@dei.unipd.it, d.gunduz@imperial.ac.uk

Abstract—Data collected and transmitted by Internet of things (IoT) devices are typically used for control and monitoring purposes; and hence, their timely delivery is of utmost importance for the underlying applications. However, IoT devices operate with very limited energy sources, severely reducing their ability for timely collection and processing of status updates. IoT systems make up for these limitations by employing multiple low-power low-complexity devices that can monitor the same signal, possibly with different quality observations and different energy costs, to create diversity against the limitations of individual nodes. We investigate policies to minimize the average age of information (AoI) in a monitoring system that collects data from two sources of information denoted as *primary* and *backup* sources, respectively. We assume that each source offers a different trade-off between the AoI and the energy cost. The monitoring node is equipped with a finite size battery and harvests ambient energy. For this setup, we formulate the scheduling of status updates from the two sources as a Markov decision process (MDP), and obtain a policy that decides on the optimal action to take (i.e., which source to query or remain idle) depending on the current energy level and AoI. The performance of the obtained policy is compared with an aggressive policy for different system parameters. We identify few types of optimal solution structures and discuss the benefits of having a backup source of information in the system.

Index Terms—Energy harvesting; age of information; Internet of things; Markov decision process.

I. Introduction

Internet of things (IoT) is increasingly being deployed for critical operations such as factory and process automation, intelligent transportation and smart cities [1]. Differently from other networks that are generally characterized in terms of throughput and delay, a key performance indicator for such applications is the *age of information* (AoI), which quantifies the freshness of the destination's knowledge about the status of the system being monitored [2], [3]. Another distinct aspect of these systems is that IoT devices are typically limited in available energy. One way to increase the sustainability of such a system is to exploit energy harvesting capabilities from ambient sources [4]. However, harvested energy is characterized by irregular energy arrivals randomly distributed over time; requiring a rechargeable battery as an energy buffer, and some form of intelligent control to avoid outages at critical instances [5]. In particular, if AoI is to be taken into account, status updates must be acquired sparingly depending on the level of energy available in the battery.

A growing number of papers investigate the evolution and control of AoI in energy-harvesting systems [6]–[10]. The scenario of reference involves a device making optimal decisions about acquiring status updates depending on the energy cost and the available battery level. In this paper, we consider instead an IoT system that can exploit *multiple* sources of information, each providing a different energy-age trade-off. For example, an IoT device may exploit multiple sensors with different reliabilities and costs. Alternatively, we may think of a terminal that can update the system status through either a cellular technology, which guarantees reliability and high coverage, but is very expensive in terms of energy, or a low-range energy-aware technology.

For the purpose of the analysis we refer to the following model. We consider two *sources* providing information with different costs and qualities (freshness and/or reliability) to a *monitoring node* that tries to optimize the resulting AoI over time within a constrained energy budget. The information source with higher cost and quality is called the *primary* source of information, while the other is referred to as the *backup* source. These sources provide the monitoring node with the most fresh status update in their buffer. Therefore, the monitoring node does not know with certainty the AoI of the packet that will be delivered from a source node. The only assumption the monitoring node can make is the reliability of a source, i.e., the probability of receiving a fresh data packet from that source.

Minimization of AoI in a multi-source system is also considered in [11], where multiple sensors communicate with the monitoring system via orthogonal channels. However, in [11] each sensor monitors a different process, and the objective is a function of the ages of all the processes. The proposed policy converts the scheduling

problem into a bipartite matching problem between the sets of channels and sensors. Similarly, in [12], [13] multiple sources provide status updates to a monitoring node about different processes, which results in a multi-objective problem, and the goal is to schedule transitions in order to balance the AoI of these different processes. In contrast with the aforementioned contributions, we consider a system with energy harvesting capabilities, where the monitoring node selectively requests status updates of the same underlying process from different nodes. We justify having more than one information source in the system to obtain diversity by balancing across different combinations of energy cost and AoI. This contribution can be considered as a first step towards developing policies for a multi-source IoT-system that can achieve efficient management without any prior knowledge of the cost or quality of the information sources, which have to be learned over time.

The rest of this paper is organized as follows. In Section II, the system model description, problem formulation and solution approaches are introduced. Numerical results are presented in Section III, providing a performance comparison between the proposed solution and a reference approach, namely, the aggressive policy. The paper is concluded in Section IV, where possible further developments are also outlined.

II. SYSTEM OVERVIEW

We consider a system consisting of a single energy-harvesting monitoring node and two sources of information, where each source takes measurements of the same underlying process that is of interest to the monitoring node. Time is discretized into time slots with a unit slot length of arbitrary duration. At each time slot, the monitoring node can receive a status update from only one of the sources. The status update from the chosen source becomes available to the monitoring node at the beginning of the time slot.

The monitoring node consumes different amounts of energy to receive a status update from the two sources. We assume that the status updates provided by the two sensors are of age either α or β, referred to as *fresh* and *stale*, respectively, with $\alpha < \beta$. For the sake simplicity, we consider only two possible age values α and β, in this paper, which can model, for example, useful and useless data packets. We assume that source i can provide a fresh status update at each time slot with probability γ_i, $i = 1, 2$, and a stale packet with probability $1 - \gamma_i$, such that $\gamma_1 > \gamma_2$ for *primary* and *backup* sources, respectively. The AoI at the monitoring node increases by 1 if no new update is received.

The energy costs of requesting a status update from source i is denoted by c_i, $i = 1, 2$, where we assume $c_1 > c_2$. Here $c_1, c_2 \in \mathbb{Z}^+$ correspond to integer multiples of a unit of energy.

Battery level $b(t)$ is updated at each time slot depending on the energy harvested in the previous time slot and the energy cost of receiving a data packet from one of the sources:

$$b(t) = \min\{b(t-1) - \sum_{i=1}^{2} c_i \cdot \mathbb{1}(a(t) = a_i) + e(t), B\}, \quad (1)$$

where $e(t) \in \{0, \bar{e}\}$ denotes the harvested energy available to be used in time slot t, B is the battery capacity, and $\mathbb{1}(x)$ is an indicator function: $\mathbb{1}(x) = 1$ when x holds, and $\mathbb{1}(x) = 0$ otherwise. We assume $\{e(t)\}_{t=1}^{\infty}$ is an independent and identically distributed (i.i.d.) binary random process with $P(e(t) = \bar{e}) = \lambda$.

The monitoring node makes a decision at the beginning of each time slot whether to request a new status update or not, and if so, which source to request it from. We seek the policy that minimizes the average AoI at the monitoring node by optimally choosing the action to take at each time slot, accounting for the battery level and the current age of information. We first formulate the problem as an MDP.

A. MDP formulation

An MDP consists of a tuple $<S, A, P, R>$ of state space S, action space A, probability transition function P, and a reward or cost function R. In our problem, finite space of actions A includes requesting an update from either of the two sources (primary/backup) and remaining idle. We set $A = \{a_0, a_1, a_2\}$, where a_0 corresponds to remaining idle, a_1 updating from the primary source, and a_2 updating from the backup source.

Action a_i is not allowed if $b(t) < c_i$. This can be incorporated into the framework with the same action space by imposing very high energy costs for action a_i when $b(t) < c_i$, $i = 1, 2$.

Let $\delta(t) \in \{1, 2, ..., \delta_{max}\}$ denote the AoI at the monitoring node at time slot t, where δ_{max} is the maximum age in the system. Equivalently, we assume that having a status information of age δ_{max}, or any $\delta > \delta_{max}$ have the same utility. Depending on action $a(t)$, $\delta(t)$ can take one of the following values $\{\delta(t-1) + 1, \alpha, \beta\}$. The system state is described by the pair of variables $s(t) = (b(t), \delta(t))$. Note that we have a finite state space of dimension $(\delta_{max} + 1)(B + 1)$. In this paper, we set β as the maximum AoI, i.e., $\beta = \delta_{max}$, beyond which increase in age becomes irrelevant. Accordingly, receiving a stale status update is equivalent to not receiving a useful update.

P denotes the transition probabilities of the MDP, where $P(s'|s, a) = Pr(s(t + 1) = s'|s(t) = s, a(t) = a)$; that is, the probability that taking action a at state s will lead to a transition to state s' in the following time slot. The transition probabilities for our problem are given as follows for $a_i \in \{a_1, a_2\}$:

$$\begin{cases} P((\min\{b+\bar{e}-c_i, B\}, \min\{\alpha, \delta+1\})|(b, \delta), a_i) = \lambda\gamma_i \\ P((\min\{b+\bar{e}-c_i, B\}, \min\{\beta, \delta+1\})|(b, \delta), a_i) = \lambda(1-\gamma_i) \\ P((b-c_i, \min\{\alpha, \delta+1\})|(b, \delta), a_i) = (1-\lambda)\gamma_i \\ P((b-c_i, \min\{\beta, \delta+1\})|(b, \delta), a_i) = (1-\lambda)(1-\gamma_i) \end{cases}$$
$$(2)$$

Note that, if the received status update is older than the currently available one, then the monitoring node drops the new packet and keeps the previous status update. We can conclude that if $\delta_t < \alpha$, then the optimal action is to remain idle, i.e., $a_t = a_0$.

When the node remains idle, i.e., $a_t = a_0$, the transition probabilities are given as follows:

$$\begin{cases} P((b, \delta+1)|(b, \delta), a_0) = 1 - \lambda & b<B \\ P((\min\{b + \bar{e}, B\}, \delta+1)|(b, \delta), a_0) = \lambda & b<B \\ P((B, \delta+1)|(B, \delta), a_0) = 1 \end{cases}$$
$$(3)$$

The policy π defines an action $a(t)$ at each time slot depending on the current state. The infinite-horizon time average AoI, when policy π is employed, starting from initial state s_0, is defined as [14]:

$$V^\pi(s_0) = \lim_{T \to \infty} \sup \frac{1}{T} \mathbb{E}\left[\sum_{t=0}^{T} \delta^\pi(t)|s(0) = s_0\right]. \quad (4)$$

A policy is optimal if it minimizes the average AoI - $V^\pi(s_0)$. The optimal infinite-horizon average AoI for a starting state s_0 is found by solving:

$$V(s_0) = \min_\pi V^\pi(s_0). \quad (5)$$

To solve this optimization problem, we can use the offline dynamic programming approach adopting the relative value iteration (RVI) algorithm described in [15]. In the offline approach we model the state transition function based on the statistical prior knowledge of the information sources' reliability and environmental characteristics. The RVI differs from VI by the value function of some state $V(s^*)$ in each update. In this case, the Bellman equation is defined as:

$$V^n(s) = \min_{a \in A} \left(\delta(s, a) - V^{n-1}(s^*) + \sum_{s' \in S} P(s'|s, a)V^{n-1}(s')\right),$$
$$(6)$$

where V^n is the value function, and s^* is a fixed state chosen arbitrarily.

The optimal stationary deterministic policy, obtained by Algorithm 1, specifies the decision rule that maps the current energy level and AoI to actions taken with probability one. In Algorithm 1, $sp(V^n - V^{n-1}) < \epsilon$ stands for the stopping criteria, where $sp(V) = \max_{s \in S} V(s) - \min_{s \in S} V(s)$. We run the RVI algorithm until the stopping criteria holds. At that moment the policy π achieves an average-cost AoI that is within $\epsilon \cdot 100\%$ of optimal.

III. NUMERICAL RESULTS

In this section, we analyze the optimal policies for different settings, in particular, we consider the cost ratios between the primary and backup information sources, reliability of the sources, and the parameters of the energy harvesting process $(\lambda, \bar{e})$. We study the structure

---

Algorithm 1 Relative Value Iteration Algorithm

set $v^0(s) = 0, \forall s \in S$
set n = 1, $\epsilon > 0$
repeat
  $n \leftarrow n + 1$
  for all $s \in S$ **do**
$$v^n(s) = \min_{a \in A} \sum_{s' \in S} P(s'|s, a)\left[\delta(s'|s, a) + V^{n-1}(s')\right]$$
$$V^n(s) = v^n(s) - v^n(s_0)$$
    where s_0 is a fixed state chosen arbitrary
  end for
until $sp(V^n - V^{n-1}) < \epsilon$
return $\arg\min V(s)$

---

of the optimal policy, and try to identify the added value in average AoI from employing an extra information source in the system.

A. Simulation parameters

System parameters that remain constant for all the numerical simulations are presented in Table I. The efficiency of the optimal policy is verified via simulations run over $T = 5000$ time slots, and compared with a so-called *aggressive policy*. The aggressive policy (Algorithm 2) tries to always receive a status update whenever it has sufficient energy in its battery, and goes for the expensive source whenever it can afford it.

Algorithm 2 Aggressive Policy

set $b(0) = 0$, $\delta(0) = 0$
for $t = 1 : T$ **do**
  if $b(t) \geq c_1$ **then**
    $b(t) = \min\{b(t - 1) + e(t) - c_1, B\}$
    if $p \leq \gamma_1$ **then**
      $\delta(t) = \alpha$
    else
      $\delta(t) = \{\beta:\delta(t-1) \geq \beta; \delta(t-1)+1:\delta(t-1) < \beta\}$
    end if
  else if $c_2 \leq b(t) < c_1$ **then**
    $b(t) = \min\{b(t - 1) + e(t) - c_2, B\}$
    if $p \leq \gamma_2$ **then**
      $\delta(t) = \alpha$
    else
      $\delta(t) = \{\beta:\delta(t-1) \geq \beta; \delta(t-1)+1:\delta(t-1) < \beta\}$
    end if
  else
    $\delta(t) = \delta(t-1)+1,$
    $b(t) = \min\{b(t-1)+e(t), B\}$
  end if
end for

Denoting by $\bar{\delta}_T^m$ the time-average AoI over T time slots at the m-th run of the simulations, we consider the mean AoI $\bar{\delta}_T = \frac{1}{M} \sum_{m=1}^{M} \bar{\delta}_T^m$ and its standard deviation:

$$st \triangleq \sqrt{\frac{\sum_{m=1}^{M} (\bar{\delta}_T^m - \bar{\delta}_T)^2}{M - 1}}, \quad (7)$$

over $M = 1000$ runs of the simulations for each settings.

TABLE I – Default parameters.

Parameters	Values
Battery capacity, B	20
Maximum age in the system, δ_{max}	30
AoI states, $[\alpha, \beta]$	$[1, 20]$
Amount of harvested energy per time slot, $\{0, \bar{e}\}$	$\{0, 3\}$
Reliability of the primary source, γ_1	0.9

(a) Cost ratio = 0.0, $\lambda = 0.2$, $\gamma_2 = 0.2$

(b) Cost ratio = 0.4, $\lambda = 0.2$, $\gamma_2 = 0.2$

(c) Cost ratio = 0.8, $\lambda = 0.2$, $\gamma_2 = 0.2$

(d) Cost ratio = 0.0, $\lambda = 0.2$, $\gamma_2 = 0.8$

(e) Cost ratio = 0.4, $\lambda = 0.2$, $\gamma_2 = 0.8$

(f) Cost ratio = 0.8, $\lambda = 0.2$, $\gamma_2 = 0.8$

(g) Cost ratio = 0.0, $\lambda = 0.8$, $\gamma_2 = 0.2$

(h) Cost ratio = 0.4, $\lambda = 0.8$, $\gamma_2 = 0.2$

(i) Cost ratio = 0.8, $\lambda = 0.8$, $\gamma_2 = 0.2$

(j) Cost ratio = 0.0, $\lambda = 0.8$, $\gamma_2 = 0.8$

(k) Cost ratio = 0.4, $\lambda = 0.8$, $\gamma_2 = 0.8$

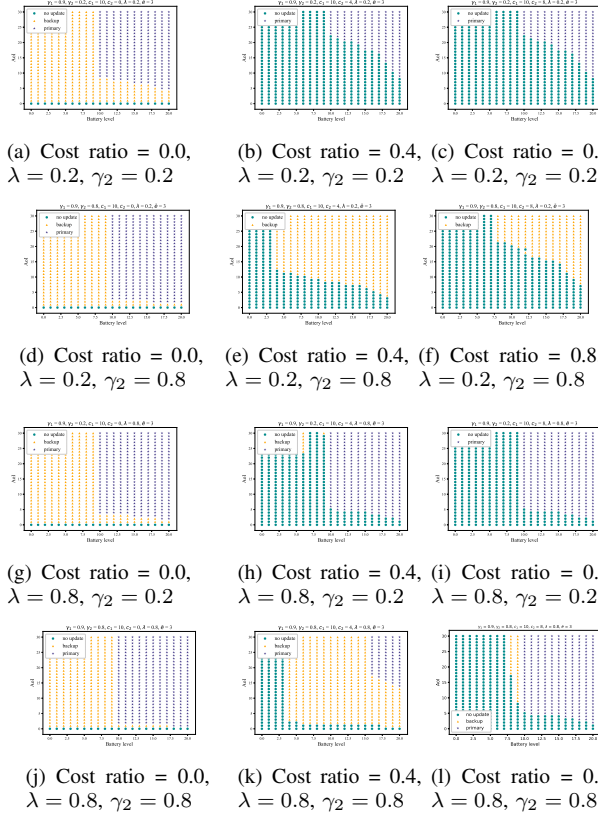(l) Cost ratio = 0.8, $\lambda = 0.8$, $\gamma_2 = 0.8$

Fig. 1 – Illustration of the optimal policy for different energy cost ratios c_2/c_1.

B. Cost ratio

The relative value of an information source can be measured by the portion of the states, in which the monitoring node chooses to exploit this source. To demonstrate this, we vary the cost ratio among the sources, c_2/c_1, and study the optimal policy obtained through RVI. We see in Fig. 1 that, when the cost ratio increases, the number of states at which the backup source is utilized shrinks, and the monitoring node opts to remain idle in most of the states. The disappearance of the backup source from the optimal action set is more rapid, if it is characterized by low reliability, γ_2 (see Figs. 1(a) - 1(c), 1(g) - 1(i)).

The relation between the average AoI and cost ratio is shown in Figs. 2 - 3. Predictably, the optimal average AoI grows when the cost ratio increases, but it saturates at a certain value, beyond which the backup source is not utilized at all. On the other hand, the average AoI increases quite rapidly at low values of the cost
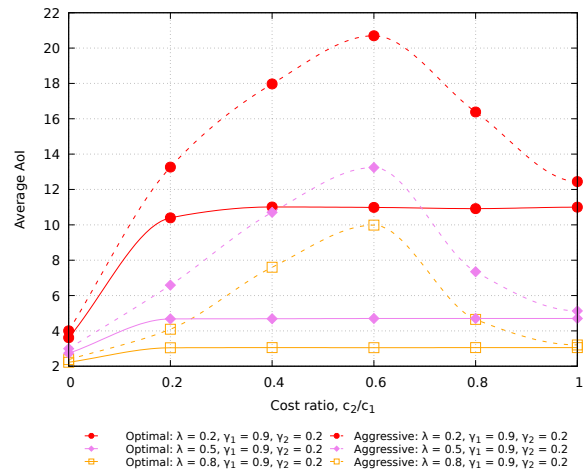

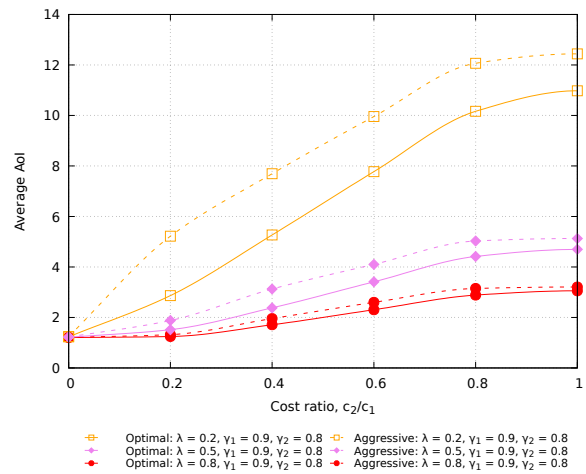Fig. 2 – Dependency of average AoI and energy cost ratio for $\gamma_2 = 0.2$.

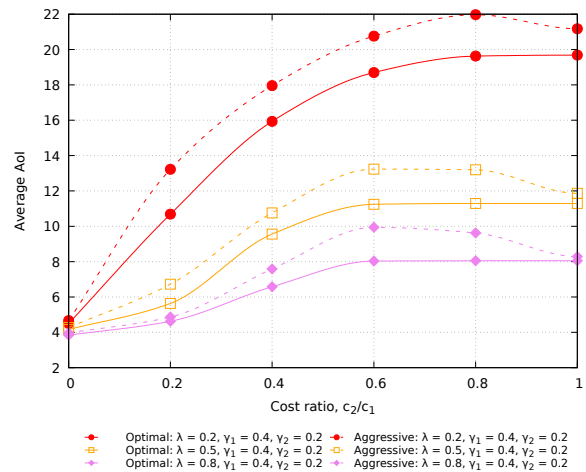Fig. 3 – Dependency of average AoI and energy cost ratio for $\gamma_2 = 0.8$.

Fig. 4 – Dependency of average AoI and energy cost ratio for $\gamma_1 = 0.4$.

ratio. Moreover, for low values of λ, i.e., low energy generation rate, the saturation of the optimal average AoI happens at lower values of the cost ratio (see Fig. 2). At lower values of source reliability, γ_2, the
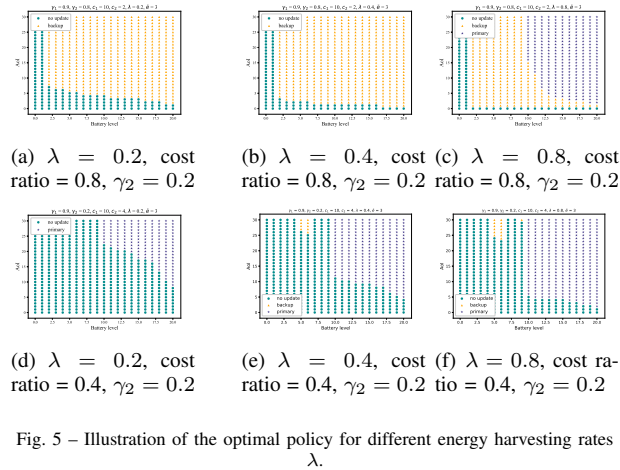
(a) $\lambda = 0.2$, cost ratio = 0.8, $\gamma_2 = 0.2$ (b) $\lambda = 0.4$, cost ratio = 0.8, $\gamma_2 = 0.2$ (c) $\lambda = 0.8$, cost ratio = 0.8, $\gamma_2 = 0.2$

(d) $\lambda = 0.2$, cost ratio = 0.4, $\gamma_2 = 0.2$ (e) $\lambda = 0.4$, cost ratio = 0.4, $\gamma_2 = 0.2$ (f) $\lambda = 0.8$, cost ratio = 0.4, $\gamma_2 = 0.2$

Fig. 5 – Illustration of the optimal policy for different energy harvesting rates λ.

Fig. 6 – Average AoI as a function of the energy harvesting rate, λ.

average AoI achieved by the aggressive policy does not have an intuitive behavior (see Fig. 2). Up to a certain point (when $c_1 - c_2 > \bar{e}$), increasing usage of the backup source causes the average AoI to grow. After some point (when $c_1 - c_2 \leq \bar{e}$), the system starts to be more energy conserving, i.e., starts to reserve energy for getting updates from the more reliable primary source, and the average AoI starts decreasing. If we set $\gamma_1 = 0.4$, a similar behaviour in average AoI is observed; however, the average AoI saturates at higher values of the cost ratio compared to the case when $\gamma_1 = 0.9$ (Fig. 2).

C. Energy harvesting

Another important parameter that impacts the optimal solution is the energy harvesting rate, λ. With increasing λ the monitoring node has tendency to request an update rather than staying idle (see Fig. 5). Furthermore, increasing energy harvesting capabilities enables the monitoring node to request updates more often from the primary information source, and reduces the utility of the backup source, which gradually disappears from the optimal solution.

Some system configurations are characterized by having a 'pocket' region, e.g., see Fig. 5(e) and 5(f). This situation is observed when the reliability of the backup source is quite low and the energy harvesting rate is sufficiently high. In this case, the energy buffer can recover in a short amount of time, which enables the monitoring node to request an update from a primary source, instead of an extremely unreliable backup source.

The dependence of the average AoI on λ is demonstrated in Fig. 6. As expected, the increase in the energy harvesting rate leads to a decrease in the achievable AoI.

D. Reliability of information sources

In Fig. 1 we can also observe the evolution of the optimal solution as the reliability of the backup source, γ_2, increases. The increase in γ_2 leads to an increase in th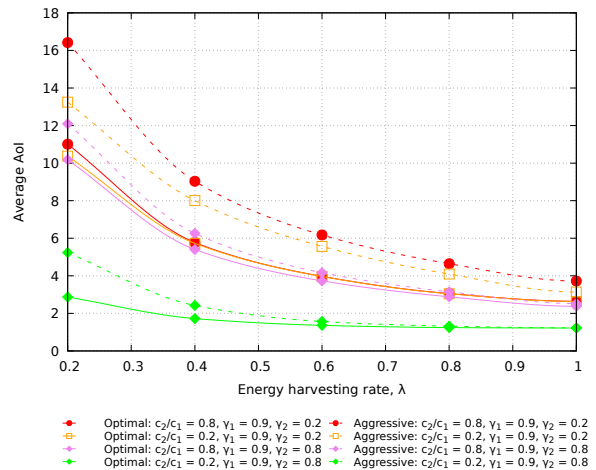e number of states in which the backup source is queried. In other words, the utility of the backup source also increases.

The dependence of average AoI on γ_2 is shown in Fig. 7. As expected, the increase in the energy harvesting rate decreases the achievable average AoI. However, if c_2/c_1 is high, then the increase in γ_2 does not severely affect the average AoI. As the backup source has a high cost, then the primary source prevails in the optimal solution, and the reliability of the backup source does not affect the average AoI significantly. If both the cost ratio and the energy harvesting rate are low, then the backup source becomes more preferable as its reliability increases. Therefore, in this case we observe a significant drop in average AoI (see Fig. 7).

E. Efficiency evaluation

Finally, in Fig. 8 we compare the performance of the optimal and aggressive policies in terms of the average AoI. The convergence time for both policies are similar, and does not exceed 200 time slots.

We observe that the gap between the average AoI achieved by the aggressive and optimal policies gets higher as the energy harvesting rate increases (Fig. 6), i.e., if the energy arrivals to the system are relatively stable, then the aggressive policy can be as effective as the optimal one. Similarly, there is no gain in average AoI if $c_2/c_1 = 0$. c_2/c_1 does not significantly influence the relative performance of the optimal policy over the aggressive one, since the gap remains relatively constant as a function of c_2/c_1 (Fig. 2-3). Generally speaking, since the backup source is less expensive but also less reliable than the primary one, the optimal policy tends to preserve energy when convenient in order to use the primary source, while the aggressive policy would always use the backup source whenever possible. Thus, the gap between the two policies shrinks as the backup source improves its reliability. However, if c_2/c_1 increases, the gap remains larger.
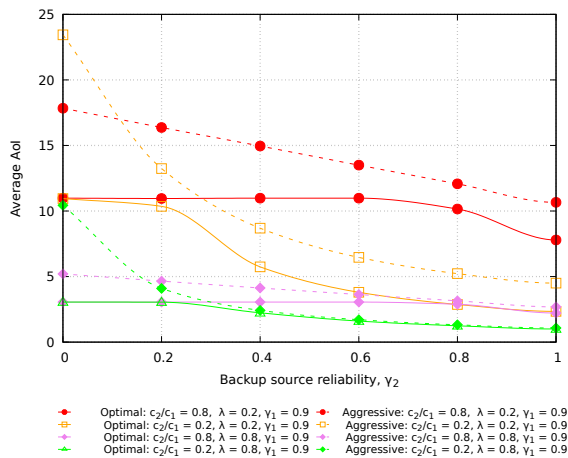
Fig. 7 – Average AoI as a function of the backup source reliability, γ_2.

(a) $\lambda = 1.0$, cost ratio = 0.2, $\gamma_2 = 1.0$

(b) $\lambda = 0.2$, cost ratio = 0.8, $\gamma_2 = 0.2$

Fig. 8 – Average AoI vs. time for the aggressive and optimal policies.

F. Discussion

We observe that the structure of the optimal solution varies depending on the characteristics of the environment and system parameters. In particular, we consider the energy harvesting rate as an environmental characteristic; the reliability of the information sources and the associated costs as system parameters. Two types of solution structure (behavior) can be distinguished: *pocket region*, or appearance of the buffer (or accumulating) region in the optimal action set, where the monitoring node chooses to stay idle in order to gain extra energy, and *monotonic disappearance* of a source from the optimal action set.

Results reported above answer the question when employing a backup source is beneficial in reducing average AoI. Low values of cost ratio, as well as high reliability of the backup source are key requirements to be met in order to integrate a backup source to the system. Improving environmental characteristics can reduce the need for the backup source, but also the benefits from employing the optimal policy. Sometimes improving the environmental characteristics (for instance, device relocation) can be a solution, instead of increasing the complexity of the system by adding extra backup devices.

IV. CONCLUSIONS

We have investigated a monitoring node that can query two distinct sources of information, a primary and a backup source, to receive status updates of an underlying process of interest. We formulated this problem as an MDP, and derived the optimal policy that minimizes the average AoI. We compared the performance of the optimal policy with that of the aggressive policy, which tries to query the most expensive source it can afford, and demonstrated that the gain from the optimal policy increases as the energy harvesting rate decreases or the backup source characteristics become worse (i.e., decreasing reliability or increasing cost).
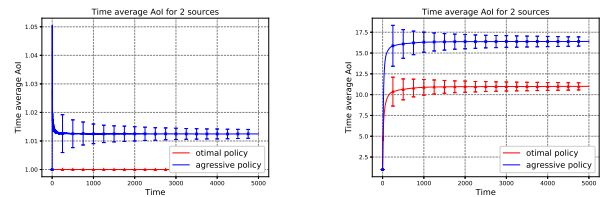
We have also shown that employing an alternative source of information is justified when the energy cost of requesting from the backup source is relatively low and its reliability is high.

As a future work, the case with more than two sources of information will be considered with more than two possible AoI states that can be received from each of the sources.

REFERENCES

[1] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel *et al.*, "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, 2017.

[2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, 2012, pp. 2731–2735.

[3] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *IEEE Conf. on Sensor, Mesh and Ad Hoc Comm. and Nets. (SECON)*, 2011, pp. 350–358.

[4] S. Basagni, M. Y. Naderi, C. Petrioli, and D. Spenza, "Wireless sensor networks with energy harvesting," *Mobile Ad Hoc Networking: Cutting Edge Directions*, pp. 701–736, 2013.

[5] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 210–216, 2014.

[6] B. T. Bacinoglu and E. Uysal-Biyikoglu, "Scheduling status updates to minimize age of information with an energy harvesting sensor," in *IEEE Int'l Symp. on Inf. Theory (ISIT)*, 2017, pp. 1122–1126.

[7] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal online policies for energy harvesting sensors with incremental battery recharges," *arXiv preprint arXiv:1802.02129*, 2018.

[8] ——, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," *arXiv preprint arXiv:1806.07271*, 2018.

[9] S. Feng and J. Yang, "Age of information minimization for an energy harvesting source with updating erasures: With and without feedback," *arXiv preprint arXiv:1808.05141*, 2018.

[10] E. T. Ceran, D. Gündüz, and A. György, "Reinforcement learning to minimize age of information with an energy harvesting sensor with HARQ and sensing cost," *IEEE INFOCOM*, 2019.

[11] V. Tripathi and S. Moharir, "Age of information in multi-source systems," in *IEEE GLOBECOM*, 2017, pp. 1–6.

[12] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. on Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, 2019.

[13] N. Pappas, J. Gunnarsson, L. Kratz, M. Kountouris, and V. Angelakis, "Age of information of multiple sources with queue management," in *IEEE Int'l Conf. on Comms. (ICC)*, 2015, pp. 5935–5940.

[14] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid ARQ under a resource constraint," *IEEE Trans. on Wireless Comms.*, vol. 18, no. 3, pp. 1900–1913, 2019.

[15] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena scientific Belmont, MA, 1995.