A Markov Game of Age of Information From Strategic Sources With Full Online Information

Matteo Pagin, Leonardo Badia, and Michele Zorzi
Department of Information Engineering (DEI), University of Padova, Italy
{paginmatte,badia,zorzi}@dei.unipd.it

***Abstract*—We investigate the performance of concurrent remote sensing from independent strategic sources, whose goal is to minimize a linear combination of the freshness of information and the updating cost. In the literature, this is often investigated from a static perspective of setting the update rate of the sources a priori, either in a centralized optimal way or with a distributed game-theoretic approach. However, we argue that truly rational sources would better make such a decision with full awareness of the current age of information, resulting in a more efficient implementation of the updating policies. To this end, we investigate the scenario where sources independently perform a stateful optimization of their objective. Their strategic character leads to the formalization of this problem as a Markov game, for which we find the resulting Nash equilibrium. This can be translated into practical smooth threshold policies for their update. The results are eventually tested in a sample scenario, comparing a centralized optimal approach with two distributed approaches with different objectives for the players.**

***Index Terms*—Remote sensing; Age of Information; Game theory; Stateful optimization.**

I. Introduction

Pervasive real-time sensing is a key component of several use cases for future generation wireless networks, such as multisensory communication for digital twins, augmented/virtual reality (AR/VR), robots, eHealth, Industry 4.0 and so on [1], [2]. Sensing capabilities are often not delegated to a single terminal but distributed across the network among multiple devices and even more numerous logical entities.

In such a context, freshness of information becomes very relevant and can be characterized by the key performance indicator of the Age of Information (AoI) [3], [4]. For these aforementioned use cases, having up-to-date information is possibly more interesting than optimizing raw throughput or average latency, and can be often precisely quantified through closed-form expressions, which makes it appealing for analytical investigations.

At the same time, to obtain scalability and ease of implementation across pervasive multi-terminal scenarios such as the Internet of Things, a distributed management with low complexity algorithms is required [5], [6]. This motivates a game theoretic approach to the evaluation of AoI coming from multiple sources of information [7]–[9].

This work was supported by the RESTART Program, financed by the Italian government with the resources of the PNRR – Mission 4, Component 2, Investment 1.3, theme 14 "Telecommunications of the future."

In light of these motivations, in the present paper we consider a scenario consisting of independent sources, tracking the same process of interest for an information sink, to which they may independently send periodic updates at the price of incurring an individual cost [10], [11]. We also assume that all sources can identically send a valuable information update that refreshes the AoI at the sink node. At the same time, they are also aware that sending redundant updates, when another source is already doing so, causes extra costs and does not further lower the AoI beyond the level already reached.

If we frame the strategic decision-making mechanism of the devices through game theory, we obtain that they are facing a classic dilemma. On the one hand, sending an update implies a cost that each source would prefer not to pay, as the AoI can be identically decreased if another source sends an update (but there is no cost to pay in this case). However, if all nodes decide not to update, they end up causing the AoI to soar, which is not efficient either.

To solve this conundrum, we frame the problem as a Markov game, on which we perform a distributed optimization from the individual perspective resulting in a Nash equilibrium (NE) [12], [13]. We can actually consider the strategic objective for a single node to be a selfish optimization of the AoI under individual cost, or the overall minimization involving the update cost for the entire network, but still from a local perspective. Finally, we can take the global network optimization achieved by a centralized approach as a benchmark.

Our analysis is original as only a handful of references study a distributed strategic optimization of AoI, and they all assume either *competing* terminals as in [9] or a stateless optimization [11], while we are the first to consider a stateful minimization of a common AoI objective.

Our results indicate that, compared with an optimal centralized approach, distributing the choice of updating causes a smoothing of the individual transmission probabilities. Some threshold effects are still present: the nodes do not update when the information is fresh, but also may fail to update with certainty when the information is stale. This is due to the lack of coordination, i.e., nodes may be strategically transferring the burden of updating to one another, to avoid paying the cost, resulting in a high price of anarchy (PoA) [14].

As a consequence, a distributed selfish system achieves worse performance with full online information than when the sources are not aware of the system AoI, in which case a close-to-optimal performance is achieved [8]. This can be seen as a consequence of providing strategic players with more information, which can be harmful in selfish setups [15], [16]. Nevertheless, our results show practical consequences that ought to be taken into account in the network design of remote sensing protocols, possibly designing effective rewarding mechanisms to benefit the strategic users [17].

The rest of this paper is organized as follows. In Section II, we characterize the system model to develop our game theoretic analysis. Section III proves some theoretical findings that are further quantitatively visualized in Section IV through numerical results. Finally, Section V gives the conclusions.

II. System model

We consider a system where N sources $S_1, \dots, S_N$ transmit updates related to a common underlying process to a receiving gateway R. We measure the value of such updates as the "freshness" of its information, formally defined as the AoI, i.e., the difference between the current time and the instant of the last update [3]. In our statement, these sources are not competing with one another nor actively collaborating. Instead, they independently send updates to R, being aware that receiving an update at the same time instant from multiple sources results in wasting resources. To account for this, we consider that each transmission incurs a cost c [18]. Indeed, without such a term, the sources would be transmitting at every possible instant, making the problem at hand trivial. Moreover, the model would not be consistent with the physical world, where sensors exhibit energy and processing limitations, especially in Internet of things (IoT) scenarios [10].

In line with similar analyses, we sample the system at periodic instants, obtaining a discrete-time axis of possible update epochs. Accordingly, AoI δ at one such instant t reads:

$$\delta(t) = t - \max\left(\left(\{\tau \leq t\} \right) \cap \left(\bigcup_{i=0}^{N} \boldsymbol{\tau}_i \right) \right), \tag{1}$$

where the set $\boldsymbol{\tau}_i$ collects the time instants of the i-th source updates, i.e., $\boldsymbol{\tau}_i \doteq \{\dots, \tau_i^{(1)}, \tau_i^{(2)}, \dots, \tau_i^{(n)}, \dots\}$. In each time slot, the sources independently decide whether to transmit an update with probability p_i that depends on the current value of the AoI, i.e., $p_i \doteq \mathbb{P}\left[\text{transmitting} \,|\, \delta = i\right]$. The knowledge of the AoI at the sources can be justified by assuming either that R sends a broadcast acknowledgment whenever an update is received, or that the sources eavesdrop the channel in a carrier sensing multiple access (CSMA) fashion [19]. Then, this information is used to keep track of the evolution of the AoI over time[1]. Finally, we assume that updates are always correctly received at R, since accounting for the presence of errors is already done in many papers [21], and including it here would distract the reader from the main focus of this analysis, i.e., quantifying the loss in efficiency caused by a lack of coordination.

We model the problem of finding the optimal transmission probabilities as a static game of complete information $\mathcal{G} = (\mathcal{S}, \mathcal{A}, \mathcal{R})$. The set of players $\mathcal{S} = \{S_1, \dots, S_N\}$ includes the sources, as the gateway is only a passive receiver of updates, incapable of making any move, hence it is not part of $\mathcal{S}$. The action set $\mathcal{A}$ contains the possible update probabilities p_i's, equal for each source. In $\mathcal{G}$, the players choose such transmission probabilities in one shot, based on the expected discounted reward over an infinite time horizon. Thus, we model the evolution of the AoI as a discrete-time, countably infinite Markov chain (MC) where the Markov property is satisfied by definition of p_i and the state-space includes all possible values of the AoI, i.e., $\mathbb{Z}_+$. From a generic state i, only states $i+1$ and 0 are accessible in one step, as the AoI either increases by one if no transmission occurs, or goes to zero whenever at least one of the sources transmits. The transition probabilities of the MC are

$$p_{i,j} = \begin{cases} (1-p_i)^N \;\; \forall\, i \in \mathbb{Z}_+ & j = i+1 \\ 1-(1-p_i)^N \;\; \forall\, i \in \mathbb{Z}_+ & j = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Finally, we define the reward $\mathcal{R}$ considering the *unilateral* payoff of each source instead of the system overall reward [6]. We first define the overall expenditure K of a source as the sum of the *system* AoI and the *individual* transmission cost. Then, we define the expected *distributed selfish* reward as $\mathbb{E}\left[\mathcal{R}^s(i, p_i)\right]$ from state i and given action p_i as:

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}^s(i, p_i)\right] &= \mathbb{E}\left[\mathcal{R}^s_{t+1} \,|\, \mathcal{S}_t = i, \mathcal{A}_t = p_i\right] \\ &= -\mathbb{E}\left[K(i, p_i)\right] \\ &= -(i+1)(1-p_i)^N - cp_i. \end{aligned} \tag{3}$$

We also consider a *centralized* policy, which implies that the nodes coordinate, with the goal of avoiding multiple concurrent transmissions. This corresponds to an equivalent system with just one source, whose expected reward thus reads

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}^c(i, p_i)\right] &= \mathbb{E}\left[\mathcal{R}^c_{t+1} \,|\, \mathcal{S}_t = i, \mathcal{A}_t = p_i\right] \\ &= -(i+1)(1-p_i) - cp_i \,\forall\, N \in \mathbb{Z}_+. \end{aligned} \tag{4}$$

Finally, we define a *distributed global* policy that is somehow intermediate, in that the costs incurred by the system are locally computed and all the nodes share the objective of minimizing the network AoI. In this way, we take the existence of multiple nodes tracking the same process into account, while still considering a distributed approach. In this case, the expected reward reads:

$$\begin{aligned} \mathbb{E}\left[\mathcal{R}^g(i, p_i)\right] &= \mathbb{E}\left[\mathcal{R}^g_{t+1} \,|\, \mathcal{S}_t = i, \mathcal{A}_t = p_i\right] \\ &= -(i+1)(1-p_i)^N - N \cdot cp_i. \end{aligned} \tag{5}$$

The two distributed policies consider different approaches to decentralized management; in *distributed selfish*, all the nodes are anarchical [14], whereas in *distributed global* they have the same goal but act without coordination, which may still decrease the efficiency from a centralized control [11].

We define a Markov decision process (MDP) $(\hat{\mathcal{S}}, \hat{\mathcal{A}}, \hat{\mathcal{P}}, \hat{\mathcal{R}})$ on top of the MC, with the aim of computing the set(s) of transmission probabilities leading to NEs. The set of transition probabilities $\hat{\mathcal{P}}$ and the state-space $\hat{\mathcal{S}}$ coincide

[1] The value of N can also be estimated in a distributed manner, for instance using the approach of [20].

with their counterparts of the MC. The action space $\hat{\mathcal{A}}$ is represented by the p_i's, equal for each source for symmetry reasons, and stationary. For practical purposes, we discretize the possible transmission probabilities into k values as $\hat{\mathcal{A}} \doteq \{0, \Delta, \ldots, (k-1)\Delta\}$; $\Delta = 1/(k-1)$. For similar reasons, we bound the state space as $\hat{\mathcal{S}} = \{n \in \mathbb{Z}_+ \,|\, n < \delta_{max}\}$.

From the MDP parameters, we find the optimal policy by using value iteration (VI) [22]. We introduce a discount factor $\gamma \in [0,1[$ to obtain a finite expected return despite the infinite horizon. Then, we estimate the discounted value-state function $v^n(s,\pi) \doteq \mathbb{E}\left[\sum_{k=0}^{+\infty} \gamma^k \mathcal{R}^n_{t+k+1} \,|\, \mathcal{S}_t = s,\, \pi = \{p_1^\pi, \ldots, p^\pi_{\delta max}\}\right] \forall\, s \in \hat{\mathcal{S}}, n \in \{s, g, c\}$ by repeating the following update for K iterations:

$$v^n_{k+1}(s,\pi) \doteq \max_{p_i} \mathbb{E}\left[\mathcal{R}^n_{t+1} + \gamma v^n_k(\mathcal{S}_{t+1},\pi) |\, \mathcal{S}_t = i, \mathcal{A}_t = p_i^\pi\right]$$
$$= \max_{p_i} \left[\mathcal{R}^n(i, p_i^\pi) + \sum_j \gamma\, p(j\,|\,i, p_i^\pi) v^n_k(j,\pi)\right]$$

and obtaining the optimal policy π^*, where the transmission probability for the i-th state $p_i^{\pi^*}$ satisfies a pseudo-steady-state condition $\underset{p_i}{\operatorname{argmax}}\, v^n_K(i,\pi) \approx \underset{p_i}{\operatorname{argmax}}\, v^n(i,\pi)$.

III. Theoretical Analysis

We now prove some structural results on the activation threshold depending on the state of the AoI [18].

Define $\bar{K}(i,\pi) \doteq -v^s(i,\pi)$ as the expected long-term total discounted cost starting from state i, choosing policy π. Let also $p_i^* \doteq p_i^{\pi^*}$, for the sake of readability. We prove Theorems 1 and 2 for the case of a single terminal sending data updates, and then generalizing to multiple transmitters.

Theorem 1. *For $\gamma \to 1$, the optimal policy π^* satisfies either $p_i^* = 1$ or $p_i^* = 0$, depending on the specific state i and following a threshold behavior, i.e., $p_i^* = 1$ if and only if i is greater than or equal to a threshold value, which depends on the cost (the larger the cost, the higher the threshold).*

Proof: At state i, the expected long-term discounted cost is the sum of three terms, namely: the transmission cost, which is equal to cp_i since the node transmits with probability p_i; the expected discounted cost from state 0, which is considered if the transmission is performed and therefore is equal to $\gamma p_i \bar{K}(0,\pi)$, and finally the expected discounted cost from state $i+1$, which is considered if the transmission is not performed instead, i.e., the related term is $\gamma(1-p_i)\bar{K}(i+1+(i+1,\pi))$. Thus,

$$\begin{aligned}\bar{K}(i,\pi) = {} & cp_i + \gamma p_i \bar{K}(0,\pi) + \\ & \gamma(1-p_i)(i+1+\bar{K}(i+1,\pi)),\end{aligned} \tag{6}$$

that, in turn, can be re-arranged to obtain

$$p_i = \frac{\gamma\left[i+1+\bar{K}(i+1,\pi)\right] - \bar{K}(i,\pi)}{\gamma\left[i+1+\bar{K}(i+1,\pi) - \bar{K}(0,\pi)\right] - c}, \tag{7}$$

from which it is easy to conclude that: (a) the p_i^* minimizing the average cost $\bar{K}(i,\pi^*)$ is non-decreasing in i, due to the only coefficient of p_i in (6) that depends on i being $-(i+1)$; and (b) the limit value of p_i^* for $i \to +\infty$ is 1.

Conversely, (6) can also be exploited to say that, for a fixed value of i, the choice of p_i results in a monotonic behavior depending on the coefficients in front of the p_i terms. Those with a negative sign, i.e., decreasing the cost as p_i increases are $i+1+\bar{K}(i+1,\pi)$, whereas those with a positive sign are $\bar{K}(0,\pi)$ and the cost c. Thus, it follows that if the cost is 0, the highest possible value of p_i will be chosen, i.e., $p_i^* = 1$. The same will happen if c is small, i.e., not able to make the sum of the positive coefficients greater than that of the negative ones. When this happens, the cost-minimizing value of p_i^* will be 0. For a different i, the trend is still the same but according to the previous reasoning a larger c is required to activate the transmissions. ∎

Theorem 2. *For $\gamma \to 1$, states s_1 and $s_2 > s_1$:*

$$\bar{K}(s_2,\pi) - \bar{K}(s_1,\pi) \le s_2 - s_1. \tag{8}$$

Proof: Let $\mathcal{A}$ be the class of the MDP underlying MC states such that their transmission probability p_i is strictly less than 1. It follows from Theorem 1 that this class is of the form $\mathcal{A} = \{n \in \mathbb{Z}_+ \,|\, n < \vartheta \le \delta_{max}\}$. Hence, it is irreducible, since all of its states communicate with each other:

$$p^{s_2-s_1}_{s_1,s_2} \ge p_{s_1,s_1+1} \cdot \ldots \cdot p_{s_2-1,s_2} > 0 \tag{9}$$

and
$$p^{s_1+1}_{s_2,s_1} \ge p_{s_2,0} \cdot p_{0,1} \cdot \ldots \cdot p_{s_1-1,s_1} > 0, \tag{10}$$

where $p^n_{s_1,s_2}$ is the probability of reaching state s_2 starting from state s_1 in exactly n steps. Thus, all states belonging to $\mathcal{A}$ are eventually reached in a finite number of steps and the MDP will collect, during its transition from s_1 to s_2, a finite reward. From that state onward, the evolution of the MC is statistically equivalent to that obtained when starting from s_2 due to the Markov property. It follows that

$$\lim_{\gamma\to 1} \bar{K}(s_2,\pi) - \bar{K}(s_1,\pi) = 0. \tag{11}$$

Then, the result follows directly from the definition of limit in the (ϵ,δ)-sense, taking $\epsilon = s_2 - s_1$. ∎

These findings state that the single transmitter never updates the information until the AoI is large enough, then it will transmit with probability 1 once a threshold state $i = \vartheta$ is reached. This means that all states $i > \vartheta$ will never be reached and the transmitter will count from 0 to ϑ. Thus, the long-term average AoI will be $\vartheta/2$ and the long-term average transmission cost will be $c/(\vartheta+1)$ since updates will happen periodically every $\vartheta+1$ slots. Such ϑ must be equal to $\lceil\sqrt{2c}-1\rceil$, which means that a cost coefficient c below 0.5 will be ineffective in limiting the transmissions [11].

For multiple sources, a similar reasoning to Theorems 1 and 2 can be applied. It is convenient to use $\sigma_i \doteq 1 - p_i$ as the probability of a terminal being silent in state i, and to define the auxiliary variable $y_i \doteq \sigma_i^N$.

If we proceed along the same lines of (6), we can write

$$\begin{aligned}\bar{K}(i,\pi) = {} & c(1-\sigma_i) + \gamma\bar{K}(0,\pi)(1-y_i) \\ & + \gamma y_i\big(i+1+\bar{K}(i+1,\pi)\big),\end{aligned} \tag{12}$$

which can be manipulated into

$$y_i = \frac{\bar{K}(i,\pi) - \gamma\bar{K}(0,\pi) - (1-\sigma_i)c}{\gamma\left[\bar{K}(i+1,\pi) + i + 1 - \bar{K}(0,\pi)\right]} \tag{13}$$

and σ_i is simultaneously satisfying $y_i = \sigma_i^N$ and (13).

While these equations contain more involved terms than (6), they can be used to derive similar, albeit less clear-cut, conclusions. From (12), we can remark that the expected long-term discounted cost at state i is increasing in σ_i and similar conclusions to the previous case can be drawn from (13) as (a) under the optimal policy, the transmission probability increases (i.e., σ_i decreases) in i, and (b) it ultimately tends to 1 when i goes to infinity.

However, due to the exponent N that was not present in the single-terminal case, the analogies are limited to the monotonic character, whereas the binary behavior of the terminal being always active or inactive no longer applies. On the contrary, instead of the terminal becoming active once a sufficiently high value of the AoI is reached, (12) implies a smoother increasing behavior of the transmission probability versus the AoI of the system that goes to 1 only asymptotically.

Still, if c is very large and the value of i is not high enough, the minimization of the total cost would require a negative value for p_i, which is not admissible. Thus, we can prove that the behavior of the transmission probability p_i for increasing i in a multi-terminal case is non-decreasing, possibly starting from zero for low values of i, then gradually increases, and only tends to 1 in the limit for very high values of i.

IV. Numerical Evaluations

We numerically evaluate the policies that arise from the different system configurations, with the goal of validating the theoretical results presented in the previous section. First, we consider the update probabilities u_i, defined as the probability that at least one source is transmitting in a given state i, that is, $u_i = 1 - \sigma_i^N = 1 - (1 - p_i^*)^N$. Then, we compute the stationary probabilities π_i of the underlying MC, given the optimal policy, and we inspect the average behavior of the system in terms of resulting average update probability $\bar{U}$, AoI $\bar{\Delta}$, system cost $\bar{C}$ and reward $\bar{\mathcal{R}}$, defined as

$$\bar{U} = \sum_i \pi_i u_i \qquad \bar{\Delta} = \sum_i \pi_i i$$
$$\bar{C} = \sum_i -\pi_i \left[1 + (N-1)\cdot\mathbb{D}\right]\cdot c p_i^*$$
$$\bar{\mathcal{R}} = \sum_i -\pi_i (i+1)(1-p_i^*)^{1+(N-1)\cdot\mathbb{D}} - \bar{C},$$

where $\mathbb{D}$ is the indicator function which maps to 1 the distributed policies and to 0 the centralized one. Unless specified otherwise, we focus on the *distributed selfish* policy.

Fig. 1 shows the update probabilities u_i versus the AoI of the system i. We consider $N = 10$ sources and the values of the transmission cost are $c \in \{50, 200\}$. In general, a higher state, i.e., a higher AoI, corresponds to an increased transmission probability, regardless of the specific policy considered, in accordance to the results proven in the previous section.

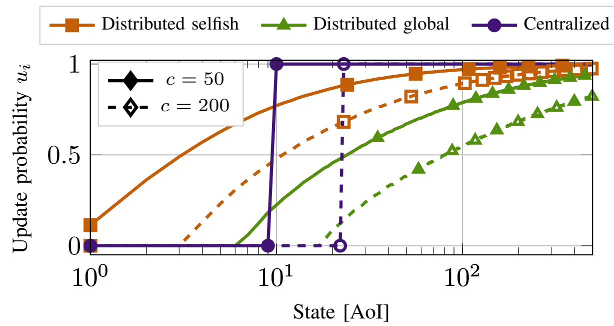
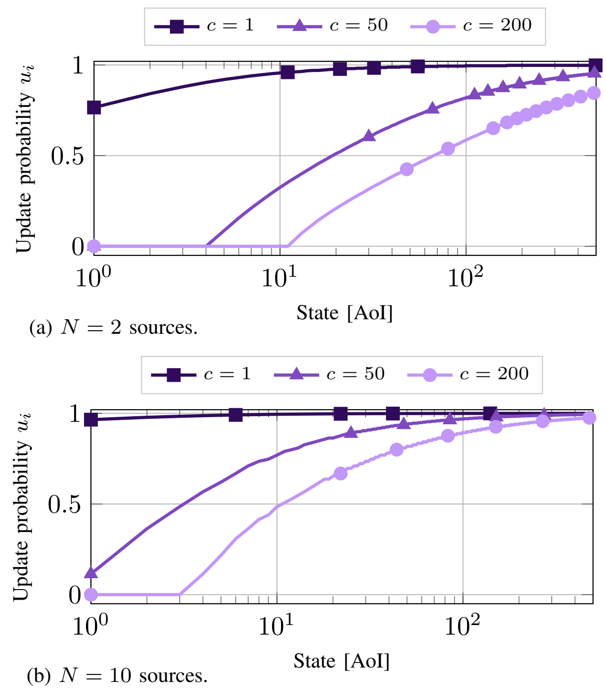


Fig. 1: Update probability u_i versus AoI for the different reward functions, for $N = 10$ and $c = \{50, 200\}$.



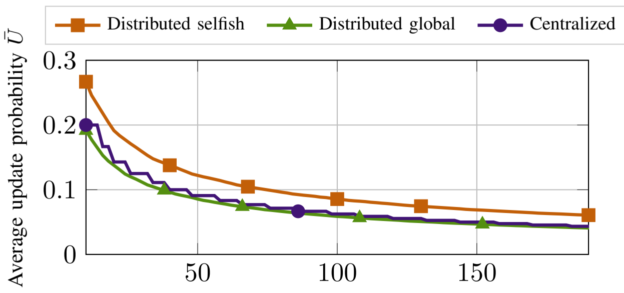
(a) $N = 2$ sources.

(b) $N = 10$ sources.

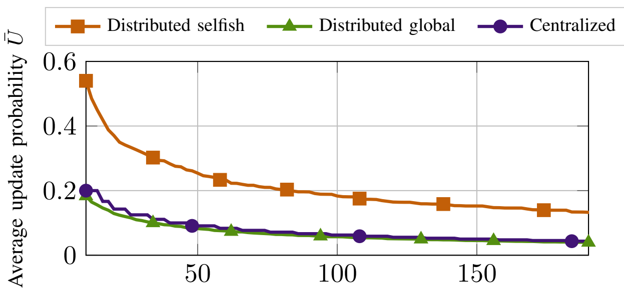
Fig. 2: Update probability u_i versus AoI for different values of the transmission cost c, $N = \{2, 10\}$ sources, *distributed selfish* system.

Note that, given a fixed transmission probability, the expected rewards are monotonically decreasing in the AoI. Therefore, it is convenient to transmit with a higher probability. Notice that u_i is also the probability that the state is reset to 0.

If the cost is increased, all curves shift to the right, which corresponds to a lower update probability. However, while this and all qualitative behaviors of the three policies follow the same monotonic trend, the shapes are different. This agrees with the theoretical results found, since the *centralized* policy has a sharp transition once a threshold value is reached. Distributed approaches have instead a smooth increase around the same threshold. It can also be noticed that the *distributed selfish* policy leads to higher update probabilities compared to the *distributed global* policy, since the cost incurred by other sources is neglected, thus a more aggressive update is obtained. This is not necessarily more efficient, since a higher transmission cost is paid (and notably, the nodes are not coordinated so redundant transmissions may happen).



(a) $N = 2$ sources.



(b) $N = 10$ sources.

Fig. 3: Average update probability $\bar{U}$ versus cost c for different reward functions, considering $N = \{2, 10\}$ sources.

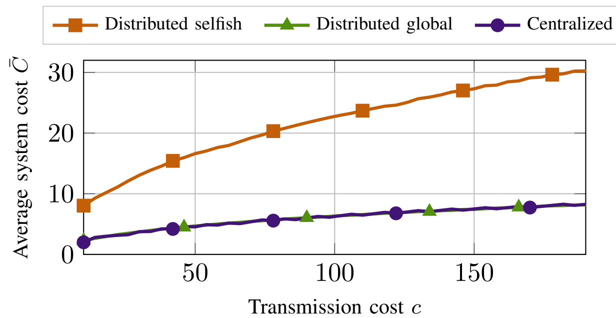


Fig. 4: Expected system cost $\bar{C}$ versus transmission cost c for different reward functions, considering $N = 10$ sources.

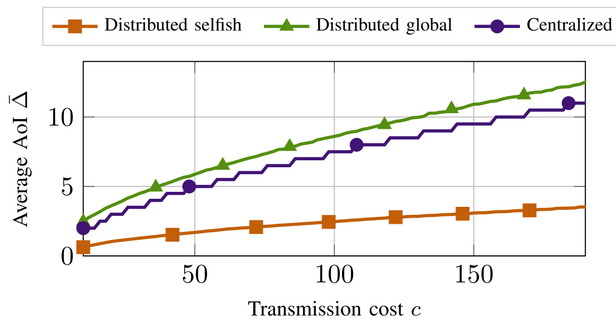


Fig. 5: Average AoI $\bar{\Delta}$ versus cost c for different reward functions, considering $N = 10$ sources.

Figs. 2a and 2b depict the update probabilities as a function of the transmission cost c within a *distributed selfish* system. For sufficiently low transmission costs, the optimal update probability is non-zero even in the initial state. On the other hand, as the cost increases $\frac{\partial \bar{K}(i,\pi)}{p_i}$ becomes negative for all states below an age threshold. Thus, the optimal strategy is to never transmit an update until such a state is visited, thus corroborating the theoretical results of Sec. III. Finally, it can also be seen how the update probability is directly proportional to the number of sources, which is to be expected as the *distributed selfish* reward function exhibits a myopic cost term which takes into account the individual transmission cost only.

Figs. 3a and 3b report the average update probability of the system $\bar{U}$ for $N = 2$ and $N = 10$ sources, respectively. Throughout the whole range of considered costs c, the *distributed global* and *centralized* policies lead to similar update probabilities, with the latter transmitting slightly more often. Conversely, the *distributed selfish* policy results in a more aggressive source behavior. The gap in terms of update probability with respect to the competitors is proportional to N, validating our choice of comparing the different policies only for low values of N. Furthermore, the *distributed global* (*centralized*) policy leads to solutions which show little (no) variability with respect to N, which suggests that these approaches are robust to estimation errors of N.

Fig. 4 presents the average cost $\bar{C}$ incurred by the whole system. It can be noticed that, as expected, the *distributed selfish* policy leads to the highest cost. Indeed, in this case the sources consider their individual cost only, thus overestimating the optimal transmission probability. Conversely, the *distributed global* and *centralized* policies result in similarly lower costs, despite the different reward functions. In fact, even though in (5) the cost term is proportional to N, the probability of transitioning to a higher AoI decreases exponentially with respect to the number of sources as well. The interplay between these two phenomena leads to an update probability which is similar to the *centralized* case.

The myopic behavior of the *distributed selfish* policy can be seen also in Fig. 5, which reports the average AoI of the system $\bar{\Delta}$. In fact, the *distributed selfish* policy achieves an AoI which is approximately 20% lower than the competitors. However, this decrease in AoI is obtained inefficiently and thus leads to the worst system performance.

Overall, these phenomena result in the average system reward $\bar{\mathcal{R}}$ depicted in Fig. 6. The *centralized* policy achieves the best performance in terms of total reward, although followed quite closely by the *distributed global* policy. The latter leads to similar update probabilities and costs, but pays for its inefficiencies related to a distributed management. Finally, the *distributed selfish* policy ends up performing quite poorly, due to its myopic overestimation of the transmission probabilities. This is motivated by the fact that the lack of coordination eventually leads to multiple concurrent updates by the sources, which effectively represents a waste of resources.

This poor performance of the distributed selfish policy may be regarded as surprising, at least in part. Even though [11] argued that the PoA of multiple uncoordinated sources is non-negligible even in the absence of an explicit competition (here, all sources try to update the same process of interest), the reason for the inefficiency seems to be more related to the distributed management rather than to the lack of awareness of the system state. Indeed, a stateful optimization

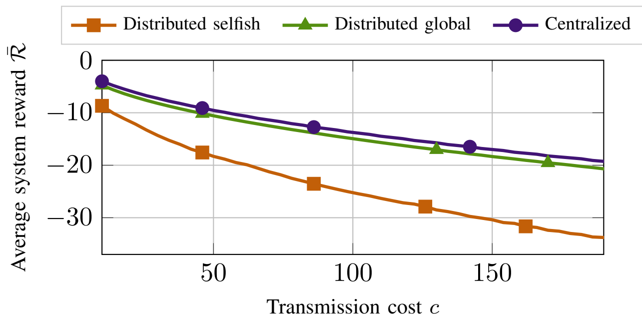


Fig. 6: Expected system reward $\bar{\mathcal{R}}$ versus cost c for different reward functions, considering $N = 10$ sources.

does not improve the situation and possibly may even make it worse. This is actually a known counterintuitive conclusion of many game theoretic approaches, most notably those involving Stackelberg games [16], where the increase of information gained by a selfish player does not necessarily improve the system performance, as selfish players try to use this additional knowledge to their own advantage and not to the system's. This may actually be a problem in massive IoT networks [23] and will possibly prompt more theoretical investigations on how to remove this inefficiency and compensate the PoA, so as to obtain a better AoI-aware management of multiple sources in pervasive scenarios. On a positive note, the *distributed global* policy incurs a negligible performance degradation compared to the *centralized* one, thus showing that a distributed system where the sources are incentivized to cooperate can achieve near-optimal efficiency despite the lack of coordination.

V. Conclusions and Future Work

We considered a scenario with multiple sensing nodes trying to update a single value of AoI at the receiver's side. Even when medium access intricacies are blurred and an idealized collision-free distributed scheme is approached, the lack of coordination of terminals may be harmful for the system AoI. This result was proven even under a stateful optimization policy, in which all nodes are aware of the value of this AoI, and desire to minimize it. However, if they incur a cost in sending updates and act in a distributed fashion without explicit competition, but just individually following a selfish objective, the overall result is a high PoA [14].

Extensions of this analysis include, from an engineering standpoint, the investigation of more specific access protocols [4], [24] to see whether the presence of collisions (and schemes to recover from collisions) confirm, mitigate, or even amplify this problem identified for an idealized scenario where multiple transmissions never collide [11].

Further game theoretic investigations can explore how to improve the network management by establishing explicit collaborations among the players. An often exploited game theoretic narrative seeks to establish collaboration through dynamic iterations, without any explicit desire for cooperation but just in accordance to the individual objectives of the selfish players [25]. It is even possible that explicit rewarding mechanisms are foreseen and implemented [17]. In our opinion, these can be practical ways to achieve network efficiency in large scale systems, and as such should be certainly pursued.

References

[1] H. Viswanathan and P. E. Mogensen, "Communications in the 6G era," *IEEE Access*, vol. 8, pp. 57 063–57 074, Mar. 2020.
[2] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.
[3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE Infocom*, 2012, pp. 2731–2735.
[4] A. Munari, "Modern random access: an age of information perspective on irregular repetition slotted ALOHA," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3572–3585, Jun. 2021.
[5] J. Delsing, "Local cloud internet of things automation: Technology and business model features of distributed Internet of things automation solutions," *IEEE Ind. Electron. Mag.*, vol. 11, no. 4, pp. 8–21, Dec. 2017.
[6] H. H. Yang, C. Xu, X. Wang, D. Feng, and T. Q. Quek, "Understanding age of information in large-scale wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3196–3210, Jan. 2021.
[7] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE ISIT*, 2012, pp. 2666–2670.
[8] L. Badia and A. Munari, "A game theoretic approach to age of information in modern random access systems," in *Proc. IEEE Globecom Wkshps*, 2021.
[9] K. Saurav and R. Vaze, "Game of ages in a distributed network," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1240–1249, May 2021.
[10] J. Hribar, M. Costa, N. Kaminski, and L. A. DaSilva, "Using correlated information to extend device lifetime," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2439–2448, Sep. 2018.
[11] L. Badia, "Age of information from two strategic sources analyzed via game theory," in *Proc. IEEE CAMAD*, 2021.
[12] R. Aras, A. Dutech, and F. Charpillet, "Cooperation through communication in decentralized Markov games," in *Proc. AISTA*, 2004.
[13] S. Leng and A. Yener, "Age of information minimization for wireless ad hoc networks: A deep reinforcement learning approach," in *Proc. IEEE Globecom*, 2019.
[14] L. Prospero, R. Costa, and L. Badia, "Resource sharing in the Internet of things and selfish behaviors of the agents," *IEEE Trans. Circuits Syst. II*, vol. 68, no. 12, pp. 3488–3492, Dec. 2021.
[15] F. Li, H. Yao, J. Du, C. Jiang, and Y. Qian, "Stackelberg game-based computation offloading in social and cognitive industrial Internet of things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5444–5455, Dec. 2019.
[16] L. Canzian, L. Badia, and M. Zorzi, "Promoting cooperation in wireless relay networks through Stackelberg dynamic scheduling," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 700–711, Dec. 2012.
[17] J. Hu, K. Yang, K. Wang, and K. Zhang, "A blockchain-based reward mechanism for mobile crowdsensing," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 1, pp. 178–191, Jan. 2020.
[18] E. Gindullina, L. Badia, and D. Gündüz, "Age-of-information with information source diversity in an energy harvesting system," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 3, pp. 1529–1540, Sep. 2021.
[19] J. Tourrilhes, "Robust broadcast: improving the reliability of broadcast transmissions on CSMA/CA," in *Proc. IEEE PIMRC*, vol. 3, 1998, pp. 1111–1115.
[20] Y.-C. Pang, S.-L. Chao, G.-Y. Lin, and H.-Y. Wei, "Network access for M2M/H2H hybrid systems: A game theoretic approach," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 845–848, Jun. 2014.
[21] R. D. Yates and S. K. Kaul, "Status updates over unreliable multiaccess channels," in *IEEE ISIT*, 2017, pp. 331–335.
[22] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
[23] H. Chen, Y. Gu, and S.-C. Liew, "Age-of-information dependent random access for massive IoT networks," in *Proc. IEEE INFOCOM Wkshps*, 2020.
[24] L. Badia, A. Zanella, and M. Zorzi, "Game theoretic analysis of age of information for slotted ALOHA access with capture," in *Proc. IEEE INFOCOM Wkshps*, 2022, pp. 1–6.
[25] S. Sorin, "Cooperation through repetition: Complete information," in *Cooperation: Game-Theoretic Approaches*. Springer, 1997, pp. 169–198.