Game Theoretic Analysis of AoI Efficiency for Participatory and Federated Data Ecosystems

Alessandro Buratto*, Alessio Mora†, Armir Bujari†, and Leonardo Badia*
* Dept. of Information Engineering (DEI), University of Padova, Italy
email: {alessandro.buratto.1@phd. , leonardo.badia@}unipd.it
† Dept. of Computer Science and Engineering, University of Bologna, Italy
email: {alessio.mora, armir.bujari} @unibo.it

***Abstract*—We investigate the Age of Information (AoI) of status updates, resulting from the convergence of multiple and federated data sources subject to both independent and voluntary participation. In this setting, the effectiveness of the status update requires the simultaneous intervention and aggregation of multiple data, according to a parametric function. Given the distributed nature of the problem, such a setup lends itself to a game theoretic approach, whereby strategic intelligent users act in a distributed fashion towards the maximization of their individual utility. The latter is chosen as a combination of the global AoI resulting from the data aggregation and parametrized by a tunable function, and an individual cost term. We compute the Nash equilibrium of the resulting allocation, and show interesting consequences of the strategic decisions made by the players.**

***Index Terms*—Age of Information; Participatory sensing; Federated learning; Mobile crowdsensing; Game Theory.**

I. Introduction

Many modern implementations of communication networks can be linked to the Internet of things (IoT), a system of interconnected physical devices that are able to sense the environment, process data and communicate with other devices over the Internet [1]. These devices tend to have limited and heterogenous capabilities, and generally need simple distributed communication protocols to work efficiently [2].

At the same time, participatory and federated paradigms represent a promising approach for harnessing the power of data, feeding data analytics algorithms for intelligent decision making [3]. Participatory sensing is a method for data acquisition, in which individual nodes are free to collect, compute and share data, e.g., about their environment and surroundings. It can be regarded a type of crowd-sourcing that leverages the widespread availability of mobile devices and their built-in sensors, to gain intelligence from data in real-time [4]. However, receiving data from multiple uncoordinated sources poses a challenge concerning how they are combined. A solution in this sense can be represented by another hot research topic, that of federated learning, whereby the learning process is distributed over a multitude of heterogeneous data sources [5], [6].

This work was supported by the European Union's NextGenerationEU instrument, under the Italian National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3, enlarged partnership "Telecommunications of the Future" (PE0000001), program "RESTART."

Participatory sensing has several advantages over traditional data collection methods [7], [8]. In particular, it can provide real-time data that are more accurate and up-to-date, and it can reach a large number of people and areas that would otherwise be difficult to access. To describe the accuracy and the timeliness of remote sensing applications, a steadily growing research direction proposes to use age of information (AoI) as the main performance metric [9]–[11].

We argue that AoI can be the best way to quantify the effectiveness of distributed data (model) collection through independent, geographically distributed nodes. While a participatory approach to data collection certainly guarantees more diversity and is therefore more resilient to outage events, it may be subject to inefficiency due to the lack of coordination. This inefficiency may be best addressed by evaluating the resulting performance through AoI instead of traditional metrics such as throughput or delay. Indeed, a distributed setup may fail to guarantee the regular status update patterns that are best to minimize AoI [12]–[14].

Finally, in recent years, several contributions have been presented to address the challenges related to the AoI of data acquisition through the use of game theory (GT) [1], [15], [16]. This approach is particularly interesting when dealing with independently-owned devices that cannot be subject to any form of centralized control. Game theoretic decision making usually involves societal environments, but more recently has found considerable applications to distributed engineered systems. Participatory sensing can combine both game theoretic applications, due to the technological nature of IoT applications, but also their societal impact.

In this paper, we present a game theoretic investigation of a scenario where N independent sources monitor the same physical phenomenon and can update a receiver through participatory and federated sensing. This means that the sources choose, over multiple rounds, whether to send data or not. This decision is made according to their own objective that combines the global AoI at the receiver's side, which is the same value for all of them, and an individual cost term that captures the effort for sending data (e.g., the energy expenditure) [17]. Moreover, we model the success or lack thereof of an update, depending on how many nodes transmit in that slot. This is meant to characterize in a parametric form the trait of federated learning [6]. Indeed, we adopt a fairly general representation of the latter part, which can be put in connection with different practical techniques.

The proposed setup is particularly suitable to a game theoretic investigation, allowing to regard nodes as strategic agents having as an objective a low AoI at the receiver's side but are also aware that there is no point in participating in the data collection if enough other nodes are already doing so. As a result, we are able to derive some extremely interesting conclusions. For example, an "anycast" communication requiring a minimum number of participating nodes equal to 1 is normally able to keep the information fresh at the receiver's side, but if more than 1 participants are required, the AoI is significantly increased, due to the lack of coordination among the sources [18].

It is also shown that the resulting AoI can even be decreasing in the number of required participants. This counterintuitive consequence happens due to the strategic character of the sources. In fact, when the number of required participant is set to a high value, knowledge of this is also available to the players, and they are more willing to participate. However, if the number of participant required by the federated learning is not high, the selfish behavior of the users may encourage them to avoid transmission (and hence the associated cost) since they believe that somebody else will take care of it.

The reminder of this paper is organized as follows: Sec. II compares our contribution to some previously conducted studies. Sec. III presents the game theoretical scenario and conducts the analysis. Sec. IV gives the numerical results of our analysis. Finally Sec. V draws the conclusions.

II. Related Work

The analysis of strategic behavior of individual users in network scenarios with distributed control can benefit from game theoretic instruments, in particular when combined with the AoI metric, which offers the advantage of a precise mathematical formalization that allows for closed-form investigations [11].

However, despite being addressed as a development already in the seminal paper [19], relatively few papers consider a game theoretic approach where the individual players are moved by an AoI-related objective. Most of the investigations actually deal with adversarial setups, thereby implying that some players have the objective of increasing the AoI of others [14], [20]. Conversely, just a handful of references consider a resource-constrained game [1], [15], [21], [22], whose setup is still competitive, i.e., individual players have their own AoI term to minimize.

Only [12] addressed for the first time the case of multiple sources with a *common* AoI value to minimize, and explore it via a game theoretic approach. However, the setup is extremely simplified as it consider just two sources and the requirement for the AoI to be reset is just that either node participates with sending data. In this sense, [23] can also be considered similar. In that paper, an out-of-band relay can assist the communication and has an objective that partially overlaps to that of the source's, i.e., AoI minimization, but in that case the roles of the involved nodes are different. At any rate, up to our knowledge no analysis considers multiple nodes, thus making our contribution novel.

A key ingredient that is usually required to spice up any game theoretic investigation is the interaction among multiple players, and how to model it. In our case, it would be convenient to use a tunable function to represent a successful combining of the data received by the nodes through federated learning techniques [24].

The mutual influence of the action of each participant on the decisions of others is also a characteristic trait of mobile crowdsensing, for which game theoretic taxonomies already exist [8]. In particular, the scenario of our considered game is to be regarded as partially cooperative, in that all the players are moved by a common AoI term as an objective (to minimize) but they are also aware of their individual cost and seek to keep it low as well.

Indeed, modeling an individual cost is very important in participatory sensing, and can be related for example to energy consumption, but also to the effort for security and privacy [25]. Both of them are important issues for any mobile crowdsensing platform, but in general there are many equally valid models that can be used to this end. For example, energy consumption depends on many technological aspects, for which reason we adopt a general parametric representation, but ultimately requires to keep under control the frequency of sending updates [26].

As a possible extension for future work, we notice that the residual battery level of devices can also be a key parameter for the recruitment of a user in a measurement campaign, but a GT instrument for that exists already, as this can be framed in the context of Bayesian games [27].

Actually, our transmission cost representation does not necessarily correlate with physical aspects such as power consumption, but also more intangible elements such as privacy, security, confidentiality, and trustworthiness are very important [28]. User may be concerned with sending their own data due to possible collection and profiling, or they may not want to disclose certain specific information. All of this can be represented in our approach thanks to the parametric representation of the utility functions, which essentially creates a linear combination analogous to [16] between the semantic objective of a low AoI with cost considerations.

III. System Model

We consider a discrete time axis divided into slots. Our system involves a set of N sensors, that are potential participants to a given sensing task, and a data collection point R, also referred to as the *receiver*. The latter is interested in keeping fresh information of the process monitored by the sensors, which is tracked by AoI, denoted as δ and formally defined as the current time minus the instant of the last successful update. The sensors can choose, independently of each other, their transmission probability p_i for a specific slot. We further assume that the sensors have always data to transmit (generate-at-will model) and we neglect the transmission delay, as commonly done in the literature [2].

In each slot, the nodes independently may decide to participate in the task. Note that the literature offers some similar analysis in the context of *medium access control*, where the sensors belong to the same collision domain and the goal is to coordinate the transmission from the sensors in a distributed fashion [16], [19].

Here, we look instead at the sensing process from the perspective of upper layers, and the physical location of the nodes is irrelevant, as long as they are all able to collect information about the process being monitored. In this context, we are interested in describing the strategic interaction of the players when making the crowd-sensing decision. Also, we want to capture that, in our system, the phenomenon measured by the sensors might not be fully understandable by the receiver if it gets a number of participants in the same round lower than a certain threshold, denoted as m. We further define a concave function that serves the purpose of reducing the success probability if there are some nodes that decide not to transmit while at least m of them are collaborating. This is to account for the fact that there might be some noisy data collected by the single nodes or some characteristics of the underlying analyzed phenomenon can be better characterized only if most of the node collaborate on the update.

Each source has the objective of reducing the AoI at the receiver and just lacks coordination with others. We introduce a cost term c which models the burden of an update on the single source. To correctly reduce the probability of an update from a source we need to consider $c \geq 1$ otherwise this would be insufficient in making the nodes refrain from transmitting at every slot.

According to [12], the expected AoI for a receiver that is getting independent updates that may or may not be successful can be computed as

$$\mathbb{E}[\delta] = \frac{1}{P_{\text{succ}}} - 1 \tag{1}$$

where P_{succ} is the probability of a successful update to occur. In turn, this event depends on whether enough sensors transmit, and since they are independent sources, P_{succ} follows a Poisson-Binomial distribution. Knowing that for a successful transmission at least m nodes have to transmit simultaneously the probability of success can be written as the survival rate $Q(m)$ of the distribution, defined as

$$P_{\text{succ}} = Q(m) = \sum_{t=m}^{N} P[x = t] \ , \tag{2}$$

where x is the number of participants. This holds true if we consider that every transmission attempt with at least m simultaneous transmissions always results in an update.

In our analysis, we focus on a generalized case, where we want to model the success probability with a soft gradient along the lines of [29], and we want to account not just for the requirement of a minimum number of participants, but also

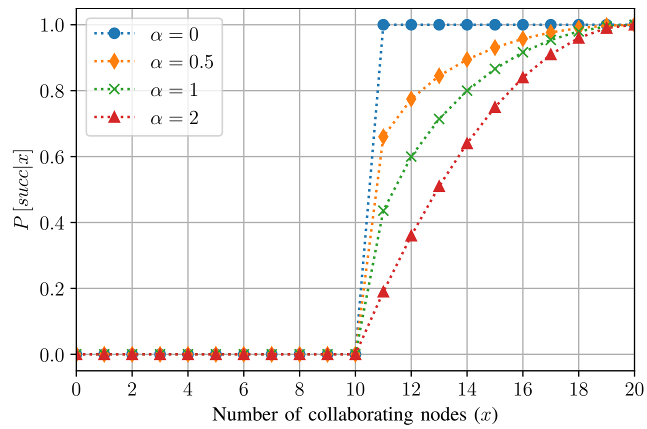

Fig. 1. Conditional success probability function for $m = 11$ and various values for α.

that this does not guarantee success. In particular, the conditional probability of a successful update as a function of the number of nodes transmitting is taken as a strictly increasing concave function, regulated by some tunable parameters, such that it increases if more than m nodes transmit and is equal to 1 only when $m = N$. This results in choosing

$$P[\text{succ}|x] = \begin{cases} \left(\sqrt{1 - \frac{(x-N)^2}{(N-m+1)^2}}\right)^{\alpha} & \text{if } m \leq x \leq N \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where the exponent $\alpha \geq 0$ is to tune the steepness of the growth of the success rate. See Fig. 1 for a graphical display of the resulting function, under various choices of α.

We can write the success probability in a more general form by applying Bayes' rules about conditional probabilities, i.e.,

$$\begin{aligned} P_{\text{succ}} &= \sum_{t=m}^{N} P[\text{succ} \cap x = t] \\ &= \sum_{t=m}^{N} P[x = t] \cdot P[\text{succ} \mid x = t] \end{aligned} \tag{4}$$

Note that $\alpha = 0$ in (3) leads to the case previously shown in (2) as we were previously assuming a probability of update equal to 1 for every $x \geq m$. Expanding further this expression, we can leverage the results of [30] for the probability mass function of the Poisson-Binomial distribution to achieve a closed-form expression for P_{succ}, and consequently for the expected AoI by applying it to (1), as

$$\begin{aligned} P_{\text{succ}} = \sum_{n=0}^{N} \Bigg\{ &\left[\sum_{t=m}^{N} P[\text{succ} \mid x{=}t] \cdot \exp\left(-\frac{2\pi j n t}{N{+}1}\right)\right] \\ &\cdot \prod_{\ell=1}^{N} \left(p_\ell \left(\exp\left(\frac{2\pi j n}{N{+}1}\right) - 1 \right) + 1 \right) \Bigg\} \Big/ (N{+}1) \end{aligned} \tag{5}$$

where j is the imaginary unit.

Consequently, each node has its own utility that is dependent not only on the transmission probability of the node itself,

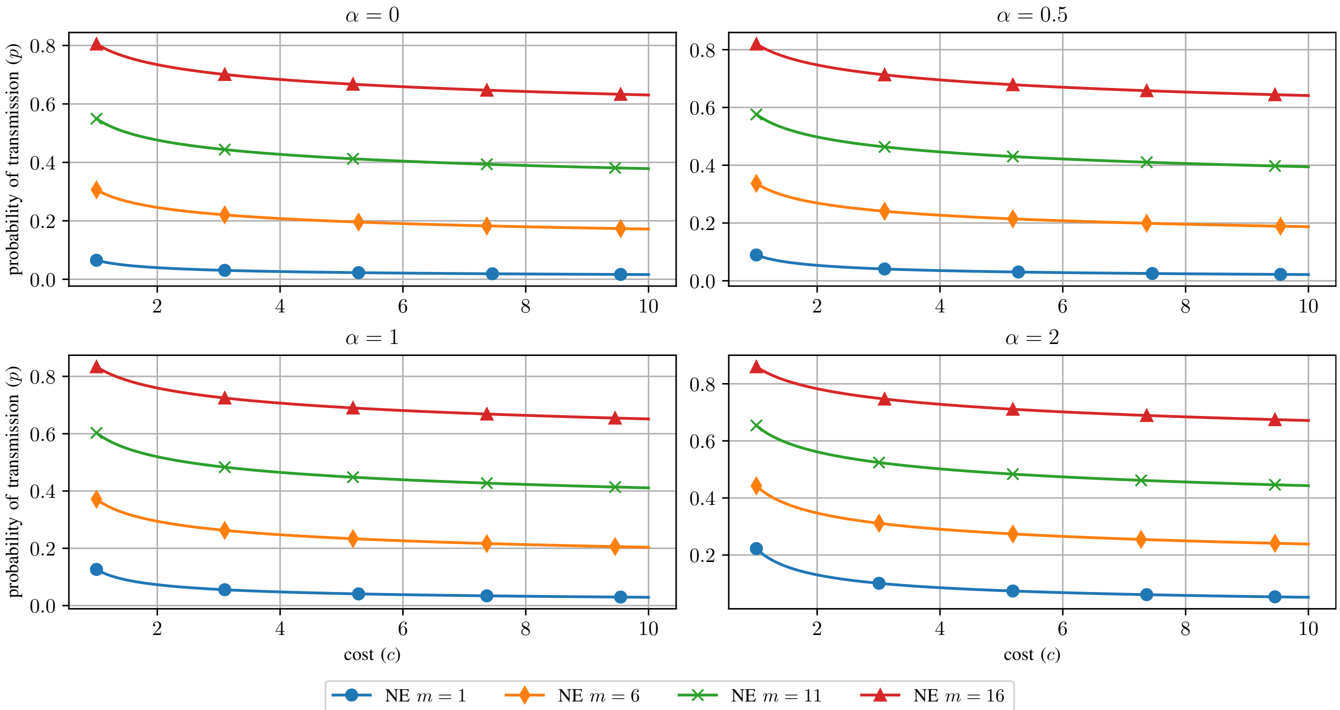

Fig. 2. Transmission probability with the NE for different values of the parameter α, $N = 20$ nodes.

but also all the other nodes' transmission probabilities. More specifically, the utility of a generic node i is defined as [16]

$$u_i = -\mathbb{E}[\delta] - cp_i = -\frac{1}{P_{\text{succ}}} + 1 - cp_i \tag{6}$$

where we consider the linear combination of two terms, the AoI of the source and the individual cost paid, weighted with coefficient c. Note the negative sign for both terms, since the players in the game seek for minimizing both AoI and cost.

From a GT perspective, we can model the interaction between the nodes as a static game of complete information $\mathcal{G} = (\mathcal{S}, \mathcal{A}, \mathcal{U})$ where $\mathcal{S} = \{S_1, S_2, \dots, S_N\}$ is the set of all the players, where $\|\mathcal{S}\| = N$ as the receiver is just a passive entity and therefore not regarded as a player. $\mathcal{A}$ is the set of the possible actions, namely the transmission probability $p \in [0, 1]$ for each player i, and $\mathcal{U}$ is the set of the utilities of the players as written in (6).

The NE of $\mathcal{G}$ is obtained through a one-sided optimization of the utility, i.e. each player looks for a *best response* to the unchanged actions of the other players. Without loss of generality we will focus on player 1 as the solution for the NE of the others is symmetrical. A NE must satisfy the condition

$$\frac{\partial u_1}{\partial p_1} = \frac{\partial P_{\text{succ}}}{\partial p_1} \cdot \frac{1}{(P_{\text{succ}})^2} - c = 0 \tag{7}$$

that can be computed applying the chain rule on (6). The derivative of P_{succ} with respect to p_1 can be obtained in closed-form from (5) with simple derivation rules. The constraint $p_i \in [0, 1]$ leads to a single feasible NE where nodes choose transmission probabilities $p_1 = p_2 = \ldots = p_N \doteq p$ due to the symmetries previously shown.

IV. Results

In this section, we discuss the results obtained by solving the previously derived equations for different values of the number of collaborating nodes m and coefficient α while keeping the total number of nodes $N = 20$ constant.

Fig. 2 shows the transmission probabilities at the NE. As it can be expected, larger values for m indicate the need of bigger values for p as it is inconvenient for the single nodes to choose to reduce their transmission probability too heavily, otherwise the AoI will grow, which is an undesirable outcome. Interestingly, we notice that increasing values of α lead to a larger p. This is due to the fact that as $\alpha \gg 1$ we are implicitly requiring that $x > m$ nodes need to be transmitting simultaneously to achieve a successful transmission and eventually $\alpha \to \infty$ will mean that $x \to N, p = 1$ for all values of m independently of the cost c.

Fig. 3 represents the utility values achieved by the nodes at the NE solution. While for $\alpha = 0$ the utilities decrease linearly with growing cost values, as α increases all the curves start to lean towards the same direction and are indistinguishable from each other for bigger and bigger costs as can be seen when $\alpha = 2$. This phenomenon is again a consequence of the effect of parameter α in the success probability that we already showed in the previous paragraph.

Fig. 4 shows the expected AoI obtained at the NE. Differently from the utility and the transmission probability, we can appreciate sensible differences for the effect of the required number of nodes transmitting m while α increases. In fact, while for $\alpha = 0$ there is a clear advantage in lowering the AoI when m is really small, this situation drastically changes

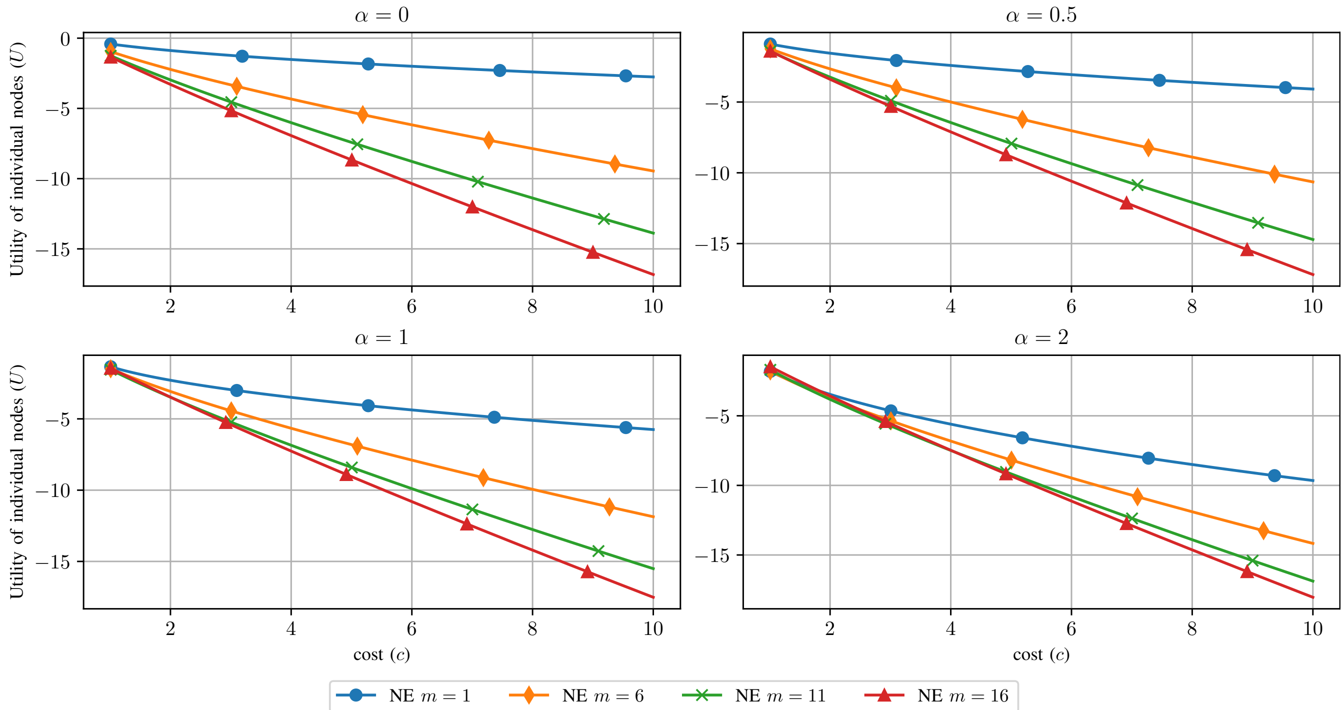

Fig. 3. Utility with the NE for different values of the parameter α, $N = 20$ nodes.

for bigger values of α. It is clear that the more exigent we are with the number of simultaneously transmitting nodes, the higher the efficiency in lowering the expected AoI for large values of the cost coefficient c. This is somehow to be expected as bigger values for m and α lead to bigger transmission probabilities as presented in Fig. 2 and this benefit will eventually wear out when $\alpha \to \infty$ as all the transmission probabilities are equal to 1 independently of m.

An interesting trend related to this last figure is also that the expected AoI $\mathbb{E}[\delta]$ is not necessarily decreasing in m, which is a consequence of the strategic behavior of the sensors. Indeed, when m is high (and this is common knowledge among the players), while the total utility is low as the sensing task is less likely to succeed, nevertheless the sensors are more compelled to participate and therefore the AoI alone is lower. However, if the cost is increased, their voluntary participation to the task is dampened, and the AoI increases. This also seems to imply an amplified importance of the cost coefficient in determining the participation, beyond the individual selfish evaluation. That is, a difficult sensing task sees an increased participation if the cost is low, but a cost increase discourages the sensors to participate and makes the success even less likely.

V. Conclusions

We analyzed, from a game theoretic standpoint, the interaction between nodes collaborating to achieve fresh information at the receiver side. We considered a scenario where a successful update is possible only when a number of nodes must choose to transmit at the same time. We further generalized this condition by applying a monotonically increasing function to reduce the probability of success if fewer than all the nodes send an update at the same time. We derived some closed-form expressions for the expected average AoI and demonstrated the existence of a single NE, where all the nodes choose to transmit with the same transmission probability. We argued that there exist some situations where a condition for the update more difficult to achieve is beneficial in lowering the AoI at the receiver when compared to less strict requirements on the number of simultaneous transmissions.

References

[1] B. Zhou and W. Saad, "Performance analysis of age of information in ultra-dense Internet of things (IoT) systems with noisy channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3493–3507, May 2022.

[2] A. Munari and L. Badia, "The role of feedback in AoI optimization under limited transmission opportunities," in *Proc. IEEE Globecom*, 2022.

[3] A. Bujari, A. Calvio, L. Foschini, A. Sabbioni, and A. Corradi, "A digital twin decision support system for the urban facility management process," *Sensors*, vol. 21, no. 24, p. 8460, 2021.

[4] G. Cisotto, A. V. Guglielmi, L. Badia, and A. Zanella, "Joint compression of EEG and EMG signals for wireless biometrics," in *Proc. IEEE Globecom*, 2018.

[5] Z. Song, J. Cheng, A. Chauhan, and E. Tilevich, "Pushing participatory sensing further to the edge," in *Proc. IEEE EDGE*, 2019, pp. 24–26.

[6] P. Bellavista, L. Foschini, and A. Mora, "Decentralised learning in federated deployment environments: A system-level survey," *ACM Comput. Surv.*, vol. 54, no. 1, pp. 1–38, 2021.

[7] S. S. Kanhere, "Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces," in *Proc. ICDCIT*, 2013, pp. 19–26.

[8] V. S. Dasari, B. Kantarci, M. Pouryazdan, L. Foschini, and M. Girolami, "Game theory in mobile crowdsensing: A comprehensive survey," *Sensors*, vol. 20, no. 7, p. 2055, 2020.

[9] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

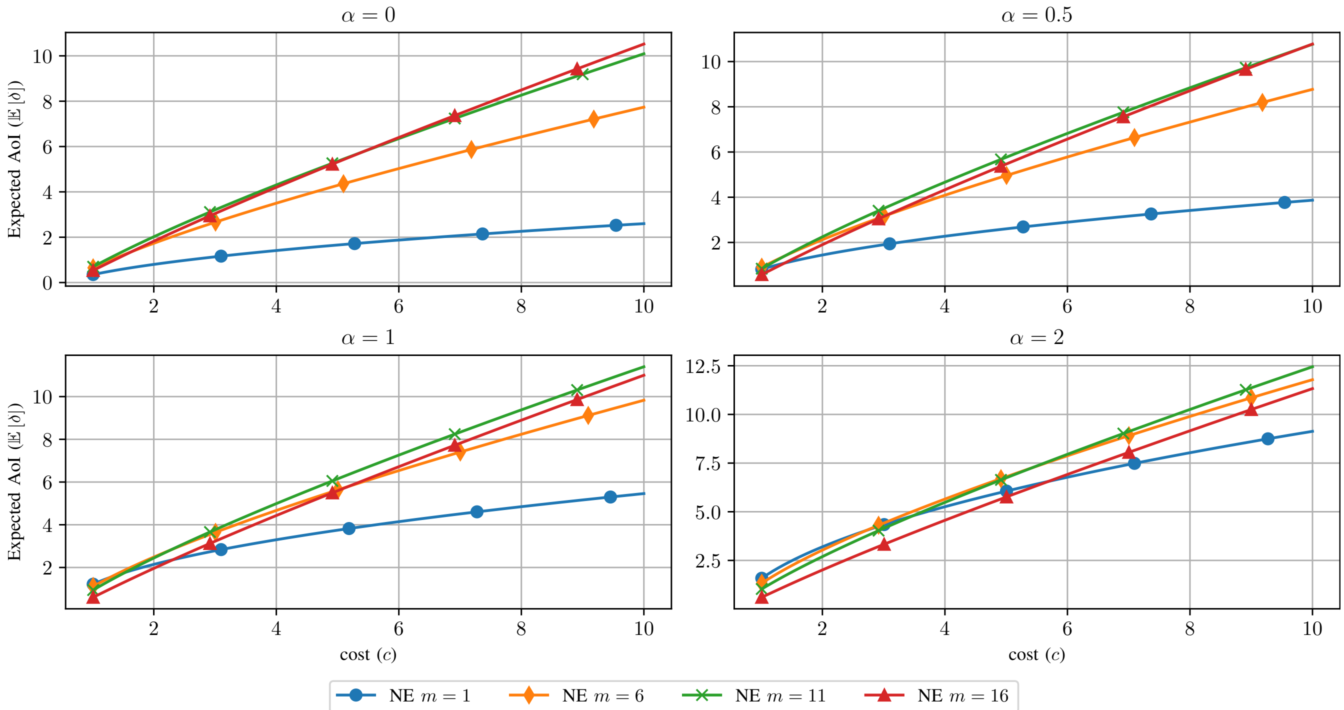

Fig. 4. Expected AoI with the NE for different values of the parameter α, $N = 20$ nodes.

[10] A. Zancanaro, G. Cisotto, and L. Badia, "Modeling value of information in remote sensing from correlated sources," in *Proc. IEEE MedComNet*, 2022.

[11] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6419–6428, Sep. 2018.

[12] L. Badia, "Age of information from two strategic sources analyzed via game theory," in *Proc. IEEE CAMAD*, 2021.

[13] O. T. Yavascan and E. Uysal, "Analysis of slotted ALOHA with an age threshold," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1456–1470, May 2021.

[14] S. Banerjee and S. Ulukus, "Age of information in the presence of an adversary," in *Proc. IEEE Infocom Wkshps*, 2022.

[15] K. Saurav and R. Vaze, "Game of ages," in *Proc. IEEE Infocom Wkshps*, 2020, pp. 942–947.

[16] 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.

[17] A. B. MacKenzie and L. A. DaSilva, *Game theory for wireless engineers*, ser. Synth. Lect. Commun. Morgan & Claypool, 2006, vol. 1, no. 1.

[18] 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.

[19] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE ISIT*, 2012, pp. 2666–2670.

[20] A. Garnaev, W. Zhang, J. Zhong, and R. D. Yates, "Maintaining information freshness under jamming," in *Proc. IEEE Infocom Wkshps*, 2019.

[21] G. D. Nguyen, S. Kompella, C. Kam, and J. E. Wieselthier, "Information freshness over a Markov channel: The effect of channel state information," *Ad Hoc Netw.*, vol. 86, pp. 63–71, Apr. 2019.

[22] 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.

[23] L. Badia and F. Chiariotti, "Age of information resilience with a strategic out-of-band relay," in *Proc. IEEE DRCN*, 2023.

[24] Z. Lian and C. Su, "Decentralized federated learning for internet of things anomaly detection," in *Proc. ACM Asia Conf. Comp. Commun. Sec.*, 2022, pp. 1249–1251.

[25] L. Giaretta, I. Savvidis, T. Marchioro, Š. Girdzijauskas, G. Pallis, M. D. Dikaiakos, and E. Markatos, "PDS 2: A user-centered decentralized marketplace for privacy preserving data processing," in *Proc. IEEE ICDEW*, 2021, pp. 92–99.

[26] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy consumption in mobile phones: a measurement study and implications for network applications," in *Proc. ACM SIGCOMM*, 2009, pp. 280–293.

[27] E. Gindullina and L. Badia, "Asymmetry in energy-harvesting wireless sensor network operation modeled via Bayesian games," in *Proc. IEEE WoWMoM*, 2017.

[28] K. Yang, K. Zhang, J. Ren, and X. Shen, "Security and privacy in mobile crowdsourcing networks: challenges and opportunities," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 75–81, 2015.

[29] L. Badia and M. Zorzi, "On utility-based radio resource management with and without service guarantees," in *Proc. ACM MSWiM*, 2004, pp. 244–251.

[30] M. Fernandez and S. Williams, "Closed-form expression for the poisson-binomial probability density function," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 2, pp. 803–817, 2010.