Staleness Minimization of Sensing Under Multiple Real-Time Observations

Melike Arslan
Dept. Computer Engineering
Bülent Ecevit University
email: melike.arslan@mf.karaelmas.edu.tr

Leonardo Badia
Dept. of Information Engineering (DEI)
University of Padova, Italy
email: leonardo.badia@unipd.it

***Abstract*—Cyber-physical systems result from the integration of digital technologies into everyday material entities to enable an interaction between computational and physical components. This study extends an existing framework of information freshness optimization, which focused on monitoring and timely reporting of events, to multiple sensing/transmission opportunities in the observation window. We present an analytical framework for joint staleness minimization for an event with known distribution, as commonly encountered in real-world applications like industrial IoT and healthcare, deriving the closed-form solutions for optimal reporting times t_1 and t_2. This creates a foundation for cyber-physical systems requiring timely monitoring of events under limitation preventing persistent monitoring.**

***Index Terms*—Internet of Things; Age of Information; Optimal scheduling; Event detection.**

I. Introduction

In recent years, mobile devices and real-time applications have significantly amplified the demand for timely data updates across numerous domains [1]. Services such as news feeds, weather alerts, traffic notifications, email messages, stock prices, social media, and mobile advertisements heavily rely on up-to-date content. In parallel, real-time updates are essential for cyber-physical control systems, including sensor-based environmental monitoring [2], autonomous vehicles [3], and biomedical platforms [4]. This dependence on timely information is driven by an ever-increasing need for immediacy across a variety of applications and systems [5], [6].

Meeting this need is non-trivial, as practical constraints such as restricted bandwidth and limited energy resources, as well as unpredictable delays in real-world networks pose significant challenges. This calls for adaptive, intelligent decision-making mechanisms capable of dynamically scheduling observations and data transmissions based on probabilistic knowledge about the occurrence of events, rather than assuming continuous data generation [7]. As such, the strategic timing of information updates becomes a crucial design concern, particularly whenever events are rare but impactful, and any delay or failure in detecting such events can result in severe consequences [8].

To quantify the freshness of status updates in network systems, the seminal reference [9] introduced age of information (AoI) as a performance metric defined as the time between the generation of the most recent update at the source and its eventual reception at the destination point [10], [11]. Since then, AoI has emerged in the literature in multiple contributions due to its analytical approach to assess real-time performance and has become a key performance indicator in environments involving sporadic sensing, real-time decision-making, and machine-to-machine (M2M) communication [12].

Originally proposed for a monitoring node that transmits time-stamped status updates to a far-away destination, AoI is a good way to summarize in one indicator both frequency and delay of updates. Depending on the application, one may want to consider a minimization of average or peak AoI, as well as more refined and application-specific metrics such as age violation probability [13]. Additionally, other metrics can be combined such as energy consumptions or transmission costs, and correlation among multiple sources [14]–[16].

However, most research focusing on AoI assumes a continuous data generation model. In contrast, many real-world systems operate under context-aware and event-driven paradigms, where data are not constantly available but instead generated sporadically and must be transmitted selectively and efficiently [17]. In such scenarios, solely targeting a low average AoI may not yield efficient outcomes, particularly when events occur infrequently.

As a response to these challenges, this study introduces an optimization framework aimed at monitoring an event taking place at a random time x that is known to happen within a bounded time window $[0, 1]$. Building upon a previous model developed in [18] for a single observation, our work extends the analysis to handle two optimal monitoring times t_1 and t_2 that minimize expected penalties derived from the AoI metric. For the sake of analytical tractability, we take a triangular distribution for x, with a tunable parameter a corresponding to the peak of the triangle (hence, the statistical mode of the distribution). We are able to identify how the results of the single observation extend to a multiple observation case, namely by combining a split of the interval for a complete monitoring and saturation effects around the border. Especially, when event more frequently happens at the end of the window, the optimal penalty is lower, coherently with the value of AoI staying contained.

The rest of this paper is organized as follows. In Section II, we review related work. Section III presents the system model and solves it analytically. We present numerical results in Section IV and we conclude in Section V.

II. Related work

The concept of AoI was introduced in pioneering papers like [9] and further elaborated by many follow-ups (see [19] for a detailed review) to describe the timeliness of real-time status reporting in sensing applications, specifically for vehicular networks. Over the last years, it has become increasingly popular in light of the increasing popularity of real-time applications for smart environment and cyber-physical systems. In particular, it has found application in industrial IoT [20], mission-critical systems [21], eHealth [4], [22], and smart agriculture [23].

This marks a converging trend of distributed multi-agent systems towards time-sensitive communications, which is one of the main pillars of next generation communication systems [1]. As such, AoI becomes critical as a way to represent, and possibly minimize, information staleness and ensuring timely delivery. As a side note, this is also a fertile ground for the application of game theory, which fundamentally relies on the coexistence of multiple objectives and multiple agents in the same system; in this spirit, coexisting sources of information may compete when resources are limited [24] or just their absence of coordination may result in inefficiencies [25].

At the same time, the widespread adoption of digital twins as a virtual counterpart of cyber-physical systems, requires a tight synchronization, and AoI is often adopted as way to represent it [26]. This reflects to many authors evaluating AoI in the context of semantic communications [27], possibly leading to further extensions of the metric itself such as age of task information (AoTI) proposed in [28].

However, in this paper we take a slightly different perspective, by focusing not just on the value itself of AoI, but rather on how timely measurements of a critical event are made possible by an AoI-inspired framework. Thus, our approach concerns more the practical collection of data towards AoI minimization, rather than assuming it for granted and optimizing other network aspects, i.e., upper layers and/or coordination among agents. Indeed, there exist studies where the concept of information freshness is approached by also considering practical details on the message exchange. For example, in [16] it is discussed how preserving accurate timestamps is crucial, along that different transport layer protocols and congestion control mechanisms influence AoI performance.

In [13], the probability of peak-age violations and delay violations are analyzed in point-to-point communication systems utilizing short information packets, with an evaluation of how parameters like frame size and undetected error probability influence data freshness. In [17], it is argued that sporadic generation of data, as typical of many sensing systems, can lead to a generally different optimization of AoI as opposed to the more standard scenario of generate-at-will that is considered in the literature [11].

On the one hand, this may lead to statistical inference approaches to estimate the actual AoI from the statistics of the underyling process, which is the idea behind some recent contributions [29], [30]. On the other hand, this implies to carefully select the observation instant for a system whenever persistent monitoring is too expensive especially since the precise temporal location of the relevant events, albeit confined to a certain horizon, is not known with certainty.

For this reason, in this paper we expand upon the work of [18], which proposed an optimal AoI-driven monitoring of an event whose precise instant is not certain, but it is known that it falls within a given finite window and according to a certain probability density function. For the sake of simplicity, that paper considered only one observation, whereas we expand the analysis to the case of multiple observations (we consider two observations in practice, but the design is easily generalizable to more points). This can be seen as an extension along the lines of assuming sporadic event generation and/or anomaly reporting, and we show how an optimized decision mechanism achieves a vanishing additional penalty to AoI when reporting updates about an event whose actual timing is uncertain.

III. System Model

In this study, we examine a time-sensitive event monitoring problem in a cyber-physical system. We assume that a relevant event of interest takes place at a random time x, which is known to be confined within an observation window, which is taken with a normalized length as $[0, 1]$. The system seeks to detect the event occurring at time x by making observations pre-established time instants. In our analysis, we consider two observation instants t_1 and t_2, thereby satisfying $0 < t_1 < t_2 \leq 1$. The exact order between x, t_1, and t_2 is unknown and will be discussed later as depending on its realization, the system suffers a different AoI-based penalty.

As a further assumption, we follow [18] in assuming that the occurrence time of the event x is modeled using a triangular probability distribution with mode equal to a, where $a \in (0, 1)$. This assumption is made both for the sake of analytical tractability but also because it reflects situations where it is known where a random event is more likely to happen. In particular, we consider a triangular distribution around a specific central point, a scenario frequently encountered in industrial and healthcare systems [17]. We also remark that, according to [23], a Gaussian distribution can be equivalently used, but at the price of losing a concise closed-form expression and necessarily requiring numerical solutions.

Thus, the probability density function (pdf) of x is defined as follows:

$$f(x) = \begin{cases} \frac{2x}{a}, & 0 \leq x \leq a \\ \frac{2(1-x)}{1-a}, & a < x \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The system attempts to detect the event at one of the two observation times depending on when it occurs:

- If $x \leq t_1$: the event is detected at time t_1 (early detection).
- If $t_1 < x \leq t_2$: the event is detected at time t_2 (mid detection).

- If $x > t_2$: the event is detected at the end of the window, i.e., at time 1 (late detection).

The cost of delayed detection is quantified using a penalty function based on the square of the detection delay, to mimic an AoI-based penalty [19]. These expressions represent the instantaneous penalty incurred for a given realization of x, depending on which region it falls into.

To account for the randomness of x, the expected penalty over the time interval is calculated by integrating the penalty over the pdf $f(x)$. This results in the following total expected cost function:

$$\mathbb{E}[P(t_1,t_2)] = \int_0^{t_1} \frac{(t_1-x)^2}{2} f(x)dx + \int_{t_1}^{t_2} \frac{(t_2-x)^2}{2} f(x)dx + \int_{t_2}^{1} (1-x)^2 f(x)dx \tag{2}$$

This function represents the average cost of using observation times t_1 and t_2 to detect an event randomly occurring within the interval $[0,1]$. It captures the trade-off between early resource consumption (checking too soon) and potential penalties from delayed detection (checking too late).

By minimizing this expected cost with respect to t_1 and t_2, we aim to find the optimal observation instants that provide the best freshness-efficiency trade-off under a given triangular distribution. This model is aligned with task-aware concepts such as AoTI [28].

IV. Analytical Framework

Building upon the system model introduced in the previous section, we now formalize the analytical framework for evaluating and minimizing the expected penalty associated with two-stage event monitoring under a triangular event distribution.

The goal is to minimize the expected total penalty $\mathbb{E}[P(t_1,t_2)]$ as a function of the sampling times t_1 and t_2, where $0 < t_1 < t_2 \leq 1$. The expectation is taken over the random variable x, the event occurrence time, the event occurrence time, drawn from a triangular distribution peaking at.

Following the AoI-based penalty framework [18], the expected penalty $P(t_1,t_2)$ is computed as a function of t_1 and t_2, with the following closed-form solution.

Case A: $t_1 \leq t_2 \leq a$:

$$P(t_1,t_2) = \frac{1}{12a}\Big[-2t_1^4 + 8t_1^3 - 6t_1^2 + 2t_2^4 - 8t_2^3 + 6t_2^2 + a^3 - 3a^2 + 3a\Big]$$

Case B: $t_1 \leq a \leq t_2$:

$$P(t_1,t_2) = \frac{1}{12a(1-a)}\Big[(1-a)(-2t_1^4+8t_1^3-6t_1^2+a^3-3a^2+3a) + a(-2t_2^4+8t_2^3+6at_2^2-18t_2^2-4a^2t_2+12t_2+a^3-3)\Big]$$

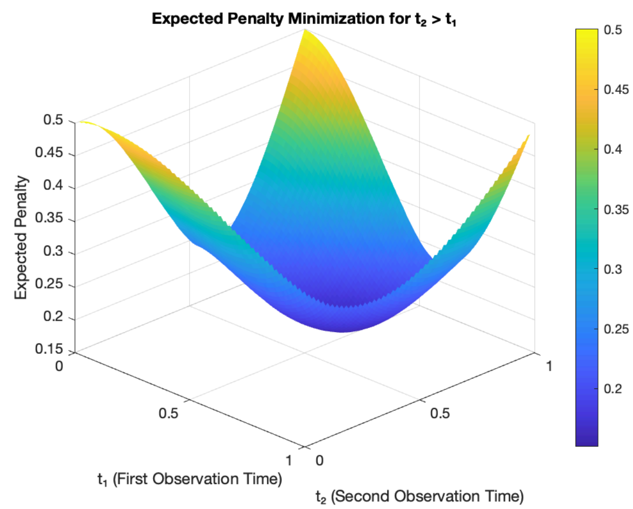

Fig. 1. Illustration of the penalty minimization point through `fmincon` for different values of a.

Case C: $a \leq t_1 \leq t_2$:

$$P(t_1,t_2) = \frac{1}{12(1-a)}\Big[-2t_1^4+8t_1^3+6at_1^2-18t_1^2-4a^2t_1+12t_1 + 2t_2^4-8t_2^3-6at_2^2+18t_2^2+4a^2t_2-12t_2+a^3-3\Big]$$

These closed-form solutions enable the evaluation of expected penalty across different triangular distribution configurations and allow for an efficient optimization. Also, the expression can be generalized to more than 2 observation points but with an even more cumbersome derivation, which is therefore omitted here. In the following section, we will discuss practical numerical evaluations of this analysis.

V. Results

We evaluated the performance of a two-stage event monitoring strategy over a normalized observation window $[0,1]$, where the event occurrence follows a triangular probability distribution with varying mode a. Using both brute-force grid search and MATLAB `fmincon` optimization algorithm, we identified the optimal observation times t_1 and t_2 that minimize the expected penalty. Fig. 1 reports a three dimensional plot of the resulting penalty as a function of t_1 and t_2 showing the unique minimization point in the domain $0 < t_1 < t_2 < 1$.

More in general, as it is immediate to expect, the optimal observation times shift to later values as the value of the distribution peak a increases. This is reported in Fig. 2 that shows the optimal positioning of the observation points t_1 and t_2, as well as the resulting expected penalty. A saturation effect is also observed, in that when the most likely values of x are those around the end of the window, with a flat trend near the end of the interval as the distribution becomes more skewed. When a surpasses a threshold at around 0.75, it becomes convenient to set the observation points earlier than that, to avoid the case of a high penalty even in the unlikely case of an early event. This is consistent with the fact that, even though

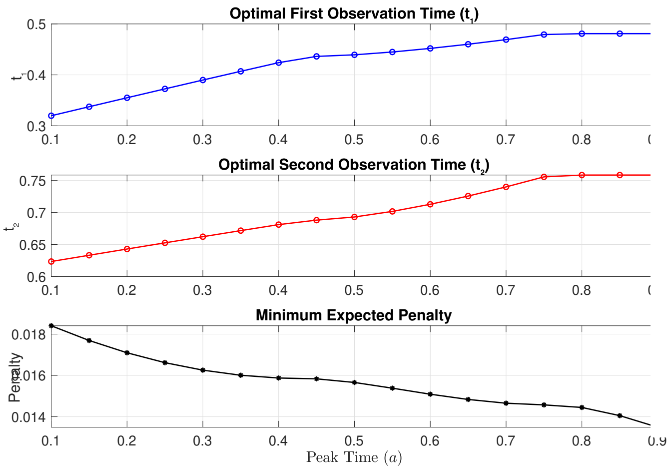

Fig. 2. Penalty minimization under two transmission points

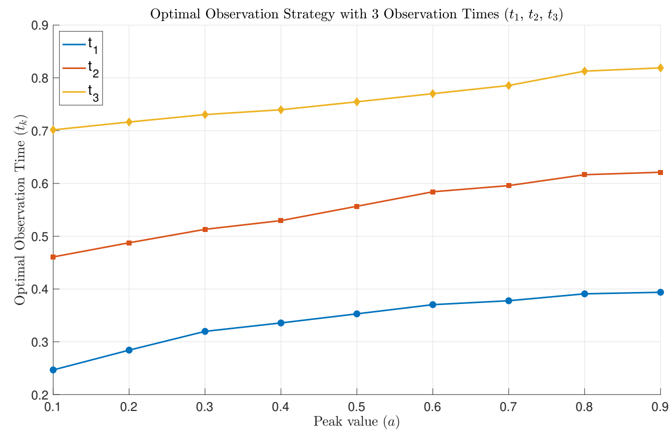

Fig. 3. Optimal observation point for three measurements

the event is ultimately not captured by either observation, the notion that nothing happened in the initial observations still gives valuable information through logical inference [30].

Such results are in line with what obtained in [18]. In addition, they also imply that a two-point monitoring reduces the late detection penalty compared to the single-observation case, which is visible by comparing the expected penalties. Finally, these results can also be generalized to more than 2 observation, even though the equations become even more complex and they are not reported here. We report in Fig. 3 the trend for 3 observation points, which shows qualitative similarity with the results presented in Fig. 2 and highlights how the saturation point is pushed even further, but the trend is overall flatter due to the presence of multiple observations.

VI. Conclusions

Real-time decision-making influenced by the Age of Information (AoI) in cyber-physical systems involves complex dynamics, providing both theoretical perspectives and real-world relevance [16]. The coordination of prompt event recognition and efficient transmission planning is vital for maintaining control performance and meeting the demands of time-critical communication systems [21].

We considered a multiple-observation detection scenario for an event taking place within a finite observation window, where the event time is only statistically characterized via a probability distribution. Specifically, the event is taken as following a triangular probability distribution with a variable peak. We studied the case of two observation points, aiming to minimize the expected AoI-related penalty by determining their optimal placement.

Building on this foundation, we formulated a mathematical optimization model that quantifies the penalty associated with delayed or missed event detection in our framework. We derived closed-form expressions and we discussed the resulting trends. This basic analysis can be extended to different analytical frameworks, with other probability distributions [23] or more observation points and/or multiple events, with the goal to align decision-making with the probabilistic nature of the system.

Indeed, besides the development of real-time applications, it is also important to gain insight on how intelligent systems can react to timing uncertainty and event distribution asymmetry in dynamic environments. For this reason, our analytical framework lays a foundation for further research into alternative probability models, event correlations, or even monitoring of multiple concurrent events.

Additionally, future extensions may explore scenarios involving multiple intelligent agents making sequential or decentralized decisions, either within a dynamic system or from a game-theoretic perspective [25]. Such work would contribute to understanding the strategic behavior and resulting efficiency of collaborative event detection in complex cyber-physical infrastructures.

References

[1] R. B. Abreu, G. Pocovi, T. H. Jacobsen, M. Centenaro, K. I. Pedersen, and T. E. Kolding, "Scheduling enhancements and performance evaluation of downlink 5G time-sensitive communications," *IEEE Access*, vol. 8, pp.128106-128115, 2020.

[2] A. R. Al-Ali, I. Zualkernan, and F. Aloul, "A mobile GPRS-sensors array for air pollution monitoring," *IEEE Sensors Journal*, vol. 10, no. 10, pp. 1666–1671, Oct. 2010.

[3] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 413–430, Sep. 2007.

[4] G. Cisotto, A. V. Guglielmi, L. Badia, and A. Zanella, "Classification of grasping tasks based on EEG-EMG coherence," *Proc. IEEE Int. Conf. e-Health Netw. Applic. Services (Healthcom)*, pp. 1–6, 2018.

[5] E. Uysal, O. Kaya, S. Baghaee, and H. B. Beytur, "Age of information in practice," in: *Age of Information: Foundations and Applications*, ch. 12, 2023.

[6] M. Rossi, L. Badia, and M. Zorzi, "On the delay statistics of SR ARQ over Markov channels with finite round-trip delay," *IEEE Trans. Wirel. Commun.*, vol. 4, no. 4, pp. 1858-1868, 2005.

[7] L. Badia, A. Baiocchi, A. Todini, S. Merlin, S. Pupolin, A. Zanella, and M. Zorzi, "On the impact of physical layer awareness on scheduling and resource allocation in broadband multicellular IEEE 802.16 systems," *IEEE Wirel. Commun.*, vol. 14, no. 1, pp. 36–43, 2007.

[8] A. Buratto, A. Munari, and L. Badia, "Strategic backoff of slotted ALOHA for minimal age of information," *IEEE Commun. Lett.*, vol. 29, no. 1, pp. 155–159, 2025.

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

[10] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *Proc. IEEE ISIT*, 2013, pp. 66–70.

[11] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[12] A. Zancanaro, G. Cisotto, A. Munari, and L. Badia, "Impact of transmission delays over age of information under finite horizon scheduling," in *Proc. IEEE Int. Wkshp Comp. Aided Modeling Design Commun. Links Netw. (CAMAD)*, 2023.

[13] R. Devassy, G. Durisi, G. C. Ferrante, O. Simeone, E. Uysal, "Reliable transmission of short packets through queues and noisy channels under latency and peak-age violation guarantees," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 721–734, 2019.

[14] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 193–204, Mar. 2018.

[15] L. Crosara, A. Zancanaro, G. Cisotto, N. Laurenti, and L. Badia, "Analytical evaluation of age of information in networks of correlated sources," in *Proc. IEEE Int. Wkshp Metrology Agric. Forestry (MetroAgriFor)*, 2022, pp. 323–328.

[16] E. Uysal, O. Kaya, S. Baghaee, H. B. Beytur, *Age of Information in Practice*, Cambridge Univ. Press, 2022.

[17] L. Badia and A. Munari, "Exogenous update scheduling in the industrial Internet of things for minimal age of information," *IEEE Trans. Ind. Informat.*, vol. 21, no. 2, pp. 1210–1219, 2025.

[18] Y. Yiğitbaşı, F. Stroppa, and L. Badia, "Optimizing real-time decision-making in sensor networks," in *Proc. Int. Conf. Data Eng. Smart Environments (DeSE)*, 2023.

[19] 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.

[20] C. Lin, B. Hu, C. Shao, Y. Ye, H. M. Tai, K. Xie, D. Jiao, and W. Li, "Event-triggered load frequency control based on age-of-information," *IEEE Trans. Plasma Sci.*, vol. 38, no. 3, pp. 2348–2361, 2022.

[21] H. Farag, M. Gidlund, and Č. Stefanović, "A deep reinforcement learning approach for improving age of information in mission-critical IoT," in *Proc. IEEE GCAIoT*, 2021, pp. 14–18.

[22] Z. Ling, F. Hu, H. Zhang, and Z. Han, "Age-of-information minimization in healthcare IoT using distributionally robust optimization," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 16154–16167, 2022.

[23] M. Lukashkou and L. Badia, "Timely monitoring of events over a finite time horizon for smart agriculture," *Proc. IEEE Int. Wkshp Metrology Agric. Forestry (MetroAgriFor)*, 2024, pp. 533–538.

[24] L. Badia, A. Zanella, and M. Zorzi, "A game of ages for slotted ALOHA with capture," *IEEE Trans. Mob. Comput.*, vol. 23, no. 5, pp. 4878–4889, 2023.

[25] G. Quer, F. Librino, L. Canzian, L. Badia, and M. Zorzi, "Inter-network cooperation exploiting game theory and Bayesian networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4310–4321, 2013.

[26] J. Li, S. Guo, W. Liang, J. Wang, Q. Chen, Z. Xu, and W. Xu, "AoI-aware user service satisfaction enhancement in digital twin-empowered edge computing," *IEEE/ACM Trans. Netw.*, vol. 32, no. 2, pp. 1677–1690, 2023.

[27] F. Chiariotti, J. Holm, A. E. Kalør, B. Soret, S. K. Jensen, T. B. Pedersen, and P. Popovski, , "Query age of information: Freshness in pull-based communication," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1606–1622, 2022.

[28] C. Ying, Z. Zhao, C. Yi, Y. Shi, and J. Cai, "An AoTI-driven joint sampling frequency and access selection optimization for industrial wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12311–12325, Sep. 2023.

[29] Z. Zhao and I. Kadota, "Optimizing age of information without knowing the age of information," *Proc. IEEE Conf. Comp. Commun. (INFOCOM)*, pp. 1–10, 2025.

[30] F. Chiariotti, A. Munari, L. Badia and P. Popovski, "Distributed optimization of age of incorrect information with dynamic epistemic logic," *Proc. IEEE Conf. Comp. Commun. (INFOCOM)*, pp. 1–10, 2025.