Age- and Correlation-Aware Information Gathering

Ahmed A. Al-Habob, *Student Member, IEEE*, Octavia A. Dobre, *Fellow, IEEE*, and H. Vincent Poor, *Life Fellow, IEEE*

***Abstract*—Age-of-information (AoI) is a metric that quantifies the freshness of gathered information. In this letter, we expand the concept of AoI by introducing a metric called correlation-aware AoI (CAAoI) to capture both the freshness and the degree of correlation in gathered information. The CAAoI of an information gathering system is evaluated when an unmanned aerial vehicle gathers information about a set of physical processes from a set of ground devices, such that each physical process is sensed by one or more devices. An optimization problem is formulated to minimize the normalized weighted sum of the time-average CAAoI in the considered information gathering system. An ant colony optimization algorithm is developed to solve the formulated problem. Simulation results illustrate that the proposed CAAoI captures both the freshness and diversity of the gathered information.**

***Index Terms*—Age-of-information (AoI), ant colony optimization (ACO), spatial and temporal correlation.**

I. Introduction

AGE-OF-INFORMATION (AoI) has been introduced as a metric to capture the freshness of information at its destination, being defined as the time elapsed since the most recently received information at the destination was generated at the source device [1]. Consequently, the AoI captures the freshness of the information and considers both the generation time and transmission latency of the updates. Freshness of information at the destination is essential for real-time control, monitoring, and decision making systems. As the number of updates at the destination increases, the AoI decreases. However, repeated updating from a device consumes its energy and raises the imbalanced load issue in the network [2]. Moreover, in real-world applications, a physical process is monitored by one or more devices and the updates about a physical process are very likely correlated. Such correlation is either (1) spatial correlation, which is a result of the spatial proximity between the devices observing the physical process, and increases as the inter-device distance decreases [3]; or (2) temporal correlation, which is a result of consecutive observations of the physical process from the same device, and increases as the time difference between consecutive observations decreases [4]. This imposes another important requirement for efficient information gathering, namely diversity in the received information at the destination.

Some existing studies of AoI have focused on developing frameworks to minimize the AoI and maintain information freshness at the destination [1], [2], [5], [6]. Other studies have tried to capture the level of dissatisfaction with data staleness [7], the value of the received information [8], and the cost of update delay [9]. However, to the best of our knowledge, no existing work has investigated the correlation in the received data in this context, despite the fact that it is highly inefficient or even impractical to minimize the AoI using replicas of updates from the same device and disregarding the fact that a physical process can be monitored by one or more devices.

Ant colony optimization (ACO) algorithms are intelligent evolutionary heuristic approaches, whose convergence toward optimum solutions of combinatorial optimization problems was proven in [10, Ch. 4.3]. ACO algorithms have efficiently solved combinatorial optimization problems, such as the travelling salesman problem [11], job-shop scheduling [12], role assignment [13], and vehicle routing [14].

In this letter, we address the following question: *Can we usefully characterize both the freshness and diversity of the received information in information gathering systems?* This question sheds light on a crucial gap in the conventional definition of the AoI which assumes that a physical process is measured using a single device; while it is known that a physical process is monitored by one or more devices and the correlation between the updates of the devices should be considered. The main contributions of this letter are as follows.

- A novel metric referred to as correlation-aware AoI (CAAoI) is introduced that not only captures the information freshness at the receiver, but also reflects the diversity in the gathered information.
- The CAAoI of an information gathering system is studied, in which an unmanned aerial vehicle (UAV) gathers information about a set of physical processes; each process can be measured by one or more ground devices.
- An ACO algorithm is developed to minimizes the normalized weighted sum of the time-average CAAoI of the observed processes.

The remainder of this letter is organized as follows. Section II introduces the concepts of AoI and CAAoI. Section III illustrates the considered information gathering system model and Section IV shows the formulated optimization problem. Section V presents the proposed ACO approach. Section VI shows simulation results for performance evaluation and Section VII concludes this letter.

Manuscript received June 16, 2021; revised October 19, 2021; accepted November 2, 2021. Date of publication November 8, 2021; date of current version February 17, 2022. The work of Ahmed A. Al-Habob and Octavia A. Dobre was supported in part by Memorial University's Research Chair and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) though its Discovery Program. The work of H. Vincent Poor was supported by the U.S. National Science Foundation under Grant CCF-1908308. The associate editor coordinating the review of this article and approving it for publication was H. Chen. *(Corresponding author: Octavia A. Dobre.)*

Ahmed A. Al-Habob and Octavia A. Dobre are with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1B 3X5, Canada (e-mail: aaaalhabob@mun.ca; odobre@mun.ca).

H. Vincent Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Digital Object Identifier 10.1109/LWC.2021.3125625

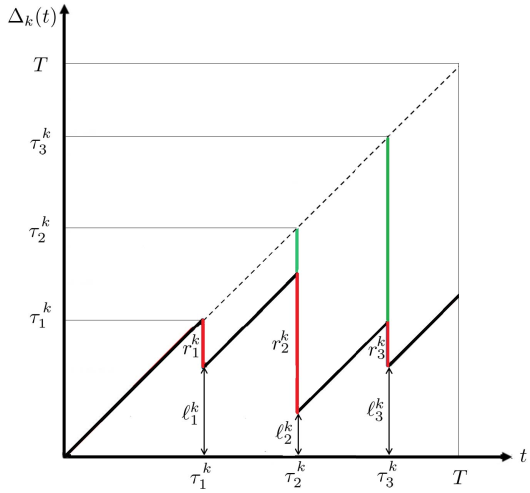

Fig. 1. AoI of the physical process p_k with $U^k(T) = 3$ updates.

II. Correlation-Aware Age of Information

To assess the freshness of the received updates at the destination, Kaul *et al.* [1] define the instantaneous AoI of the physical process p_k at time instant *t* as

$$\Delta_k(t) = t - u^k(t), \tag{1}$$

where $\Delta_k(0) = 0$ and $u^k(t)$ is the time instant at which the last update about p_k was generated (also referred to as the "timestamp" of the last update). An update with a timestamp u_j^k reaches the destination at time instant τ_j^k, such that $\tau_j^k = u_j^k + \ell_j^k$, with ℓ_j^k as the latency of transmitting the j-th update from the source to the destination. Fig. 1 illustrates the AoI of p_k with a total number of updates $U^k(T) = 3$ over a time duration *T*.

Note that the first update reduces the AoI by $r_1^k \triangleq \tau_1^k - \ell_1^k = u_1^k$, the second update reduces the AoI by $r_2^k \triangleq \tau_2^k - \ell_2^k - r_1 = u_2^k - u_1^k$, while the third update reduces the AoI by $r_3^k \triangleq \tau_3^k - \ell_3^k - r_1 - r_2 = u_3^k - u_2^k$. Consequently, the instantaneous AoI of the physical process p_k in (1) can be expressed as

$$\Delta_k(t) = t - \sum_{j=1}^{U^k(t)} (u_j^k - u_{j-1}^k), \tag{2}$$

where $u_0^k = 0$, u_j^k is the timestamp of the *j*-th update, and $U^k(t)$ is the number of updates received before time instant *t*. Measurements from spatially proximal devices or successive measurements from the same device are correlated. To capture this, we define the instantaneous CAAoI of a physical process p_k at time instant *t* as follows:

$$\Gamma_k(t) = t - \sum_{j=1}^{U^k(t)} \alpha_j^k (u_j^k - u_{j-1}^k), \tag{3}$$

where $0 \leq \alpha_j^k \leq 1$ is the novelty factor, which reflects the novelty of the *j*-th update with respect of the previous (*j* −1) updates of the physical process p_k, such that $\alpha_1^k = 1$ and α_j^k is defined as

$$\alpha_j^k = 1 - \frac{\xi_s + \xi_t}{\xi_s \dot{d}_j / \rho_s^k + \xi_t \dot{t}_j / \rho_t^k + \xi_s \xi_t + 1}, \tag{4}$$

with ρ_s^k and ρ_t^k as constant parameters that represent the spatial and temporal correlation extent in the physical process p_k, respectively, $\dot{d}_j$ as the minimum distance between the device that sends the *j*-th update and all the devices that have sent the (*j* −1) updates of the physical process p_k, and $\dot{t}_j$ as the time difference between the current update and the last update about p_k from the same device.[1] Further, ξ_s and ξ_t are introduced to give the decision maker the ability to consider either spatial correlation ($\xi_s = 1, \xi_t = 0$), temporal correlation ($\xi_s = 0, \xi_t = 1$), or both spatial and temporal correlation ($\xi_s = \xi_t = 1$).

III. System Model

In this work, we study a data gathering scenario in which a UAV is given the mission of monitoring a set $\mathcal{P} = \{p_k\}_{k=1}^P$ of *P* physical processes by gathering information from a set $\mathcal{N} = \{n_i\}_{i=1}^N$ of *N* devices. Each device is able to sense one physical process; to represent the devices ability to observe the processes, we define $\boldsymbol{\chi} = [\chi_{ik}]_{N \times P}$ such that

$$\chi_{ik} = \begin{cases} 1, & \text{if } n_i \text{ observes } p_k, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The UAV has a finite battery capacity and can only operate for a finite time interval. Consequently, the total observation time is *T* seconds. Initially, the UAV is placed at a docking station ψ_0 and the updates from devices are scheduled to keep the UAV updated about the status of the physical processes $\mathcal{P}$ during the observation time *T*. To gather information from device n_i, the UAV hovers at $\psi_i = \{x_i, y_i, h\}$, where (x_i, y_i) represents the coordinates of device n_i and *h* is the UAV's altitude.

Assuming additive white Gaussian noise (AWGN), the data rate between device n_i and UAV is

$$R_i = B \log_2\left(1 + \frac{P_i}{\bar{\boldsymbol{\varphi}}_i P_n}\right), \tag{6}$$

where P_i is the transmit power of device n_i, P_n is the power of the AWGN, *B* is the bandwidth of the channel, and $\bar{\boldsymbol{\varphi}}_i$ is the average channel path-loss between device n_i and the UAV. A probabilistic air-to-ground communication model is considered [15], in which $\bar{\boldsymbol{\varphi}}_i$ is expressed as

$$\bar{\boldsymbol{\varphi}}_i = \Pr_i(\text{LoS})\varphi_i(\text{LoS}) + [1 - \Pr_i(\text{LoS})]\varphi_i(\text{NLoS}), \tag{7}$$

where $\Pr_i(\text{LoS})$ is the probability of a line-of-sight (LoS) connection between device n_i and the UAV, which is expressed as [15]

$$\Pr_i(\text{LoS}) = \frac{1}{1 + \beta_1 \exp(-\beta_2[\theta_i - \beta_1])}, \tag{8}$$

where β_1 and β_2 are constants depending on the type of environment, either dense urban, urban, or rural, while θ_i is the elevation angle of the UAV with respect to device n_i during communication. Since the UAV communicates with device n_i while hovering above it, $\theta_i = \frac{\pi}{2}$. Finally, $\varphi_i(\text{LoS})$ and $\varphi_i(\text{NLoS})$ denote the path losses of the LoS and non-LoS (NLoS) connections, respectively, and can be expressed as

$$\varphi_i(\text{LoS}) = \text{PL}_i + \zeta_{\text{LoS}}; \varphi_i(\text{NLoS}) = \text{PL}_i + \zeta_{\text{NLoS}}, \tag{9}$$

[1]The initial value of $\dot{t}_j$ is set to $(1 - \xi_s)\xi_s^{-1}$.

where ζ_{LoS} and ζ_{NLoS} represent the excessive path loss in LoS and NLoS, respectively. Furthermore, $\mathrm{PL}_i = 20\log_{10}(\frac{4\pi h f_c}{c})$ is the free-space path loss, with f_c as the carrier frequency and c as the speed of light.

A scheduling policy is represented by $(\boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{u})$, such that $\boldsymbol{\eta} = [\eta_i]_{1\times N}$ with η_i as the number of updates from device n_i, $\boldsymbol{\mu} = [\mu_\iota]_{1\times F}$ with μ_ι as the index of the device that transmits the ι-th update and $F = \sum_{i=1}^{N} \eta_i$ as the total number of updates received at the UAV about $\mathcal{P}$ from $\mathcal{N}$, and $\boldsymbol{u} = [u_\iota]_{1\times F}$ with u_ι as the timestamp of the ι-th update. The time required for moving the UAV from the device that sends the $(\iota-1)$-th update to the one that sends the ι-th update is $\bar{\mathcal{T}}_\iota = \|\psi_{\mu_{\iota-1}} - \psi_{\mu_\iota}\|/v$, where v is the speed of the UAV and $\|\cdot\|$ is the second norm. The required time for the UAV to receive an update from device n_i is $\mathcal{T}_i = \frac{\sum_{k=1}^{P}\chi_{ik} I_k}{R_i}$, where I_k (in bits) is the payload size of an update of p_k. Device n_i has a finite capacity battery of $E_i^{\max}$ and each time it is scheduled to transmit an update to the UAV, its battery level shrinks by $P_i\mathcal{T}_i$.

IV. Problem Formulation

The objective is to find a scheduling policy that minimizes the time-average CAAoI of the processes of interest. The time-average CAAoI of the physical process p_k over time interval T can be expressed as

$$\langle \Gamma_k(t)\rangle_T \triangleq \frac{1}{T}\int_0^T \Gamma_k(t)\ dt = \frac{T}{2} - \frac{1}{T}\sum_{j=1}^{U^k(T)} \alpha_j^k (u_j^k - u_{j-1}^k)(T - \tau_j^k), \quad (10)$$

where $U^k(T)$ is the total number of updates about the physical process p_k during the observation interval T. It is worth mentioning that the maximum value of $\langle\Gamma_k(t)\rangle_T$ is $T/2$ and $\sum_{k=1}^{P} U^k(T) = F$. The objective function is a normalized weighted sum of time-average CAAoI of the processes of interest and the optimization problem is formulated as

$$\mathbf{P1}\ \min_{\boldsymbol{\eta},\boldsymbol{\mu},\boldsymbol{u}}\ O(\boldsymbol{\mu},\boldsymbol{u}) = \frac{1}{0.5T}\sum_{k=1}^{P}\lambda_k\langle\Gamma_k(t)\rangle_T, \quad (11a)$$

$$\text{s.t.}\ \sum_{\iota=1}^{F}(\mathcal{T}_\iota + \bar{\mathcal{T}}_\iota) \le T, \quad (11b)$$

$$P_i\mathcal{T}_i\eta_i \le E_i^{\max}, \forall n_i \in \mathcal{N}, \quad (11c)$$

$$\eta_i \in \mathbb{Z}^+, \forall n_i \in \mathcal{N}, \quad (11d)$$

$$\mu_\iota \in \{1, 2, \ldots, N\}, \quad (11e)$$

where λ_k is an importance weight for the physical process p_k, such that $\sum_{k=1}^{P}\lambda_k = 1$ and $\mathbb{Z}^+$ represents the set of non-negative integers. Constraint (11b) guarantees that the UAV has enough time to travel and receive all the scheduled updates. Constraint (11c) guarantees that each device is able to transmit all its scheduled updates.

V. Proposed Ant Colony Optimization Algorithm

ACO is a swarm intelligence approach, in which a set of agents "artificial ants" cooperate to solve an optimization problem. The artificial ants mimic the foraging behavior of their biological counterparts in finding a path to the food by indirect communication manner, through depositing a substance called pheromone. Each ant constructs a solution by traveling a tour through the search space of the optimization problem and deposits pheromone to reflect the quality of the corresponding solution. The tour path selection is a stochastic procedure which depends on two parameters, namely the attractiveness and trail pheromone. The points of the search space with higher pheromone concentration will more likely be chosen and thus reinforced.

Algorithm 1 ACO Algorithm for CAAoI Information Gathering

Input: $\mathcal{N}$, T, $E_i^{\max}$, λ_k, ρ_s, ρ_t, A, and I;
Initialize δ_{il}; Calculate ℓ_{il} and $\mathcal{T}_i$;
$O \leftarrow \infty$;
For Iteration $= 1$ to I **do**
  $O_1 \leftarrow \infty$; $O_2 \leftarrow \infty$;
  For $a = 1$ to A **do**
  $\boldsymbol{\eta}^{(a)} = \mathbf{0}_{1\times N}$; $\boldsymbol{\mu}^{(a)} \leftarrow \varnothing$; $\boldsymbol{u}^{(a)} \leftarrow \varnothing$; $t^{(a)} = 0$; $\boldsymbol{s}^{(a)} = \mathbf{0}_{N\times N}$;
   Set $i = 0$; Evaluate $\varepsilon_{il}^{(a)}$ and $\epsilon_l^{(a)}$ $\forall 1 \le l \le N$
   While $\prod_{l=1}^{N}(1-\varepsilon_{il}^{(a)}) + \prod_{l=1}^{N}(1-\epsilon_l^{(a)}) = 0$ **do**
    Select a device n_{l^*} using (12); $t^{(a)} = t^{(a)} + \mathcal{T}_{l^*} + \ell_{il^*}/v$;
    $s_{il^*}^{(a)} = s_{il^*}^{(a)} + 1$; ; $i = l^*$;
    $\eta_i^{(a)} = \eta_i^{(a)} + 1$; $\boldsymbol{\mu}^{(a)} = [\boldsymbol{\mu}^{(a)}\ i]$; $F^{(a)} = F^{(a)} + 1$;
    Re-evaluate $\epsilon_i^{(a)}$ and $\varepsilon_{il}^{(a)}$ $1 \le l \le N$;
   End While
   Evaluate $O(\boldsymbol{\mu}^{(a)}, \boldsymbol{u}^{(a)})$ using (10) and (11a);
   If $O > O(\boldsymbol{\eta}^{(a)}, \boldsymbol{\mu}^{(a)}, \boldsymbol{\mu}^{(a)})$
    $\boldsymbol{\eta}^* \leftarrow \boldsymbol{\eta}^{(a)}$; $\boldsymbol{\mu}^* \leftarrow \boldsymbol{\mu}^{(a)}$; $\boldsymbol{u}^* \leftarrow \boldsymbol{u}^{(a)}$;
   End if
   If $O_1 > O(\boldsymbol{\mu}^{(a)}, \boldsymbol{u}^{(a)})$
    $\boldsymbol{\eta}^{(a_1)} \leftarrow \boldsymbol{\eta}^{(a)}$; $\boldsymbol{\mu}^{(a_1)} \leftarrow \boldsymbol{\mu}^{(a)}$; $\boldsymbol{u}^{(a_1)} \leftarrow \boldsymbol{u}^{(a)}$; $\boldsymbol{s}^{(a_1)} \leftarrow \boldsymbol{s}^{(a)}$;
   Else if $O_2 > O(\boldsymbol{\mu}^{(a)}, \boldsymbol{u}^{(a)})$
    $\boldsymbol{\eta}^{(a_2)} \leftarrow \boldsymbol{\eta}^{(a)}$; $\boldsymbol{\mu}^{(a_2)} \leftarrow \boldsymbol{\mu}^{(a)}$; $\boldsymbol{u}^{(a_2)} \leftarrow \boldsymbol{u}^{(a)}$; $\boldsymbol{s}^{(a_2)} \leftarrow \boldsymbol{s}^{(a)}$;
   End if
  End for
   Deposit pheromone of a_1 and a_2 using (17);
End for
Return $\boldsymbol{\eta}^*$, $\boldsymbol{\mu}^*$, and $\boldsymbol{u}^*$.

The proposed ACO algorithm is presented as Algorithm 1, in which a colony of A ants collaborate to solve P1. The tour of each ant $a \in A$ starts by setting the time indication $t^{(a)} = 0$ and $\boldsymbol{s}^{(a)} = \mathbf{0}_{N\times N}$, and the updates from the devices are embedded in the scheduling policy $(\boldsymbol{\eta}^{(a)}, \boldsymbol{\mu}^{(a)}, \boldsymbol{u}^{(a)})$ until there is no time to receive more updates or none of the devices has sufficient transmission energy. The probability of scheduling the update from device n_l after the current device n_i is

$$\pi_{il}^{(a)} = \frac{\varepsilon_{il}^{(a)}\epsilon_l^{(a)}(\varrho_{il}^{(a)})^{\gamma_1}(\delta_{il})^{\gamma_2}}{\sum_{\substack{n=1\\ n\neq l}}^{N}\varepsilon_{nl}^{(a)}\epsilon_n^{(a)}(\varrho_{il}^{(a)})^{\gamma_1}(\delta_{il})^{\gamma_2}}. \quad (12)$$

The parameters in (12) are as follows.

- $\varrho_{il}^{(a)}$ is the attractiveness of scheduling the update from device n_l after the current device n_i; it is set to be[2]

$$\varrho_{il}^{(a)} = \frac{D\sum_{k=1}^{P}\chi_{lk}\lambda_k}{(1+\eta_l^{(a)})\ell_{il}}, \tag{13}$$

where $\ell_{il} = \|\psi_i - \psi_l\|$ is the distance between n_i and n_l, and D is a constant [11]. ϱ_{il} in (13) suggests more attractiveness to devices that monitor physical processes with higher λ_k, are in close proximity to n_i to minimize the UAV's traveling distances, and have less already scheduled updates.
- δ_{in} is the trail pheromone.
- $\varepsilon_{il}^{(a)}$ indicates that there is enough time to receive the update from n_l; it is set to be

$$\varepsilon_{il}^{(a)} = \begin{cases} 1, & \text{if } t^{(a)} + \mathcal{T}_l + \ell_{il}/v \leq T, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

- $\epsilon_l^{(a)}$ indicates that n_l has enough transmission energy; it is set to be

$$\epsilon_l^{(a)} = \begin{cases} 1, \text{if } \ P_l\mathcal{T}_l\eta_l^{(a)} \leq E_l^{\max}, \\ 0, \text{otherwise.} \end{cases} \tag{15}$$

- γ_1 and γ_2 control the influence of the attractiveness and pheromone, respectively.

Once a device n_{l^*} is selected according to (12), the device index l^* is embedded to $\boldsymbol{\mu}^{(a)}$, $t^{(a)} = t^{(a)} + \mathcal{T}_{l^*} + \ell_{il^*}/v$, the timestamp of the update is embedded to $\boldsymbol{u}^{(a)}$, and $F^{(a)}$ and the corresponding elements $\eta_{l^*}^{(a)}$ and $s_{il^*}^{(a)} \in \boldsymbol{s}^{(a)}$ increase by one. The quality of a solution constructed by ant a is reflected in the deposited pheromone $\varpi^{(a)}$, which is set to be

$$\varpi^{(a)} \triangleq \frac{1}{O(\boldsymbol{\mu}^{(a)}, \boldsymbol{u}^{(a)})} = \frac{0.5T}{\sum_{k=1}^{P}\lambda_k\langle\Gamma_k^{(a)}(t)\rangle_T}. \tag{16}$$

A global updating rule is considered in the proposed ACO, in which the algorithm repeats I colonies and in each colony only two ants with highest and second-highest deposit pheromone according to (16) are allowed to deposit their pheromone [16]. The trail pheromone is updated as follows:

$$\delta_{il} \leftarrow (1-\sigma)\delta_{il} + s_{il}^{(a)}\varpi^{(a)}, \tag{17}$$

where σ is the pheromone evaporation coefficient.

A. Computational Complexity Analysis

The total number of updates that can be gathered by the UAV over a time duration T is upper bounded by

$$\tilde{F} = \left\lceil \frac{T}{\min_{1\leq i,l\leq N, i\neq l}\{\ell_{il}/v\} + \min_{1\leq i\leq N}\{\mathcal{T}_i\}} \right\rceil. \tag{18}$$

The denominator in (18) is the minimum required time to travel between two devices plus the minimum required time to receive an update from a device. The search space of the optimization problem **P1** is $\mathcal{O}(N(N-1)^{\tilde{F}-1})$. An ant a performs $\mathcal{O}(N^2\tilde{F})$ operations to construct a solution and $\mathcal{O}(N\tilde{F}^2\log(\tilde{F}))$ operations to evaluate the objective function. Consequently, the computational complexity of the ACO algorithm is $\mathcal{O}(IA[N^2\tilde{F} + N\tilde{F}^2\log(\tilde{F})] + IN^2) = \mathcal{O}(IA[N^2\tilde{F} + N\tilde{F}^2\log(\tilde{F})])$; this is remarkably lower than the complexity of the exhaustive search approach, which is $\mathcal{O}(N^2(N-1)^{\tilde{F}-1}\tilde{F}^2\log(\tilde{F}))$.

[2]For the conventional AoI, $\varrho_{il}^{(a)} = \frac{D\sum_{k=1}^{P}\chi_{lk}\lambda_k}{\ell_{il}}$.

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value	Parameter	Value
P_i	20 dBm	h	200 m	B	200 kHz
I_k	2560 Bytes	P_n	-110 dBm	$v_{\max}$	12 m/s
f_c	2 GHz	ζ_{LoS}	0 dB	ζ_{NLoS}	20 dB
T	30 min	β_1	10	β_2	0.03
γ_1/γ_2	1	A	100	I	10^3

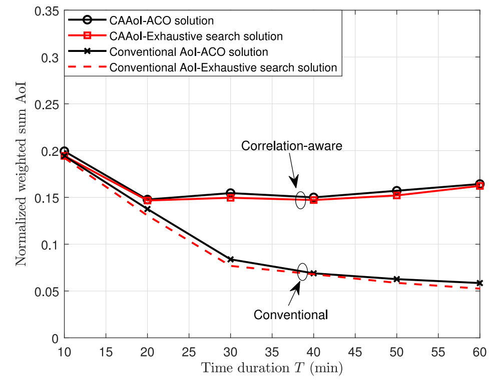

Fig. 2. CAAoI and conventional AoI of the considered system model versus the time duration T with $P = 3$ process, $\xi_s = 1$, $\xi_t = 1$, $\rho_s^k = 100$, and $\rho_t^k = 100$.

VI. Numerical Results and Discussions

In this section, we evaluate the CAAoI of the considered system model, in which the devices are distributed in a 1×1 km^2 area and the main parameters are listed in Table I.

Figure 2 illustrates both the conventional AoI and CAAoI of the considered system model with $P = 3$ physical processes versus a range of the observation time T. It presents the ACO algorithm solution and the optimal solution, obtained through exhaustive search. It is seen that the proposed ACO algorithm achieves near-optimum performance, and as T increases, the gap difference between the CAAoI and conventional AoI increases. This is attributed to the fact that increasing the time to gather data from a fixed number of devices increases the correlation in the gathered data, which increases the CAAoI.

Figure 3 illustrates both the conventional AoI and CAAoI of the considered system model with $P = 3$ physical processes that are monitored by spatially-correlated devices. It is seen that as N increases, both the CAAoI and conventional AoI reduce and the gap difference between them decreases as well. This is attributed to the fact that increasing N increases the diversity. Device diversity augments the diversity in the gathered information, which reduces the CAAoI. While increasing N, the inter-device distance also reduces, which enables the UAV to gather updates about the three physical processes more-frequently. To gain a deeper insight into such behavior, the right-side y-axis of Fig. 3 illustrates the corresponding

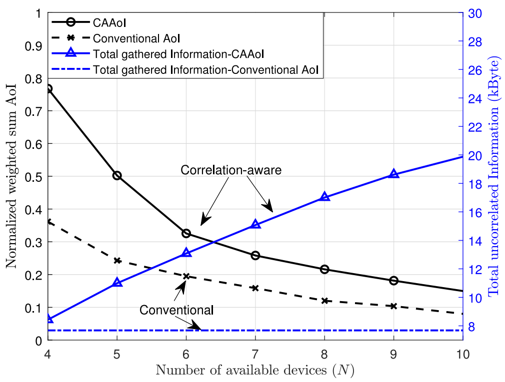

Fig. 3. CAAoI and conventional AoI of the considered system model versus N with $P = 3$ process, $\xi_s = 1$, $\xi_t = 0$, and $\rho_s^k = 100$.

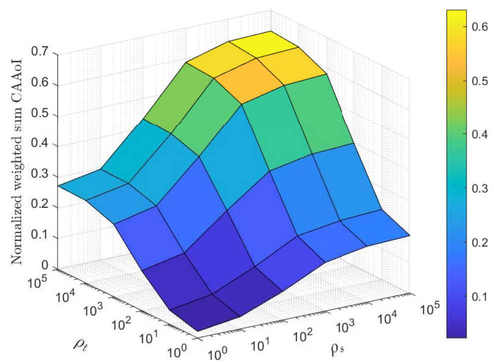

Fig. 4. CAAoI of the considered system model versus $\rho_s = \rho_s^k$ and $\rho_t = \rho_t^k$ with $N = 20$ devices, $P = 3$ process, $\xi_s = 1$, and $\xi_t = 1$.

total uncorrelated information gathered at the UAV, which equals $\sum_{k=1}^{3} \mathcal{H}_k(|\mathcal{N}_k|)$, where $\mathcal{H}_k(|\mathcal{N}_k|)$ is the uncorrelated information of process p_k,[3] $\mathcal{N}_k$ is the subset of devices scheduled to send updates about p_k, and $|\mathcal{N}_k|$ is the number of devices in $\mathcal{N}_k$. $\mathcal{H}_k(|\mathcal{N}_k|)$ can be estimated as $\mathcal{H}_k(|\mathcal{N}_k|) = I_k + I_k \sum_{j=2}^{|\mathcal{N}_k|} [1 - \frac{1}{d_j/\rho_s^k + 1}]$ [3], [13], [17]. It can be seen that the uncorrelated information corresponding to the conventional AoI does not change versus *N*. This can be explained, as in order to reduce the conventional AoI, the UAV gathers data from the closest subset of devices and keeps receiving replicas from the same devices. On the other hand, such replicas do not reduce the CAAoI. Thus, the UAV tries to gather data from all the available devices to minimize the CAAoI, which increases the gathered uncorrelated information.

Figure 4 portrays the CAAoI versus ρ_t^k and ρ_t^k. It is clear that as the correlation among the data decreases (for small values of ρ_t^k and ρ_t^k), the novelty of each update increases, which reduces the CAAoI. The opposite is also valid, i.e., as the correlation in the data increases, the novelty of the updates decreases and the CAAoI increases.

[3]The uncorrelated information of the physical process p_k equals the joint-entropy of the subset of devices scheduled to send updates about p_k to the UAV.

VII. CONCLUSION

In this letter, we have extended the AoI concept by introducing a new CAAoI metric to capture both the freshness and diversity of gathered information. The CAAoI of an information gathering scenario has been studied, in which a UAV monitors a set of physical processes by gathering information from a set of devices. An ACO algorithm has been developed to minimize the CAAoI. Results have illustrated that the proposed CAAoI enables the UAV to maintain the freshness and diversity of the gathered information.

REFERENCES

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

[2] H. H. Yang, C. Xu, X. Wang, D. Feng, and T. Q. S. Quek, "Understanding age of information in large-scale wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3196–3210, May 2021.

[3] S. Pattem, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," *ACM Trans. Sens. Netw.*, vol. 4, no. 4, p. 24, Aug. 2008.

[4] M. C. Vuran, Ö. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Comput. Netw.*, vol. 45, no. 3, pp. 245–259, Jun. 2004.

[5] X. Xie, H. Wang, L. Yu, and M. Weng, "Online algorithms for optimizing age of information in the IoT systems with multi-slot status delivery," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 971–975, May 2021.

[6] H. Tang, J. Wang, L. Song, and J. Song, "Minimizing age of information with power constraints: Multi-user opportunistic scheduling in multi-state time-varying channels," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 5, pp. 854–868, May 2020.

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

[8] P. Gjanci, C. Petrioli, S. Basagni, C. A. Phillips, L. Bölöni, and D. Turgut, "Path finding for maximum value of information in multi-modal underwater wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 404–418, Feb. 2018.

[9] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "The cost of delay in status updates and their value: Non-linear ageing," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4905–4918, Aug. 2020.

[10] M. Dorigo and T. Stützle, *Ant Colony Optimization*, 1st ed. Cambridge, MA, USA: Bradford Book, 2004.

[11] M. Dorigo and L. M. Gambardella, "Ant colony system: A cooperative learning approach to the traveling salesman problem," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 53–66, Apr. 1997.

[12] J. Zhang, X. Hu, X. Tan, J. H. Zhong, and Q. Huang, "Implementation of an ant colony optimization technique for job shop scheduling problem," *Trans. Inst. Meas. Control*, vol. 28, no. 1, pp. 93–108, Mar. 2006.

[13] A. A. Al-Habob, O. A. Dobre, and H. V. Poor, "Role assignment for spatially-correlated data aggregation using multi-sink Internet of Underwater Things," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 3, pp. 1570–1579, Sep. 2021.

[14] J. E. Bell and P. R. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Adv. Eng. Informat.*, vol. 18, no. 1, pp. 41–48, Jan. 2004.

[15] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[16] K. Doerner, W. J. Gutjahr, R. F. Hartl, C. Strauss, and C. Stummer, "Pareto ant colony optimization: A metaheuristic approach to multiobjective portfolio selection," *Ann. Oper. Res.*, vol. 131, no. 1–4, pp. 79–99, Oct. 2004.

[17] A. A. Al-habob, O. A. Dobre, and H. V. Poor, "Energy-efficient spatially-correlated data aggregation using unmanned aerial vehicles," in *Proc. 31st Annu. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, London, U.K., 2020, pp. 1–6.