

Optimal Scheduling for Discounted Age Penalty Minimization in Multi-Loop Networked Control

Onur Ayan, Mikhail Vilgelm, Wolfgang Kellerer

Chair of Communication Networks

Technical University of Munich, Germany

{onur.ayan, mikhail.vilgelm, wolfgang.kellerer}@tum.de

Abstract—Age-of-information (AoI) is a metric quantifying information freshness at the receiver. Since AoI combines packet generation frequency, packet loss, and delay into a single metric, it has received a lot of research attention as an interface between communication network and application. In this work, we apply AoI to the problem of wireless scheduling for multi-loop networked control systems (NCS), i.e., feedback control loops closed over a shared wireless network. We model the scheduling problem as a Markov decision process (MDP) with AoI as its observable states and derive a relation of control system error and AoI. We further derive a stationary scheduling policy to minimize control error over an infinite horizon. We show that our scheduler outperforms the state-of-the-art scheduling policies for NCS. To the best of our knowledge, this is the first work proposing an AoI-based wireless scheduling policy that minimizes the control error over an infinite horizon for multi-loop NCS.

I. INTRODUCTION

Cyber-Physical Systems (CPS) are engineered systems that integrate computation and control with physical processes. Power grids, smart transportation, and industrial robotics are some prominent applications of CPS [1]. From a system theoretical point of view, many CPS fall under the category of Networked Control Systems (NCS), i.e., feedback control loops closed over a communication network. Each loop consists of a plant, a sensor, and a controller with state estimator. In a typical NCS scenario, multiple heterogeneous NCS share a wireless communication medium and compete for network resources to transmit their latest sensor measurements to the controllers. The received measurement is then used to estimate the plant state remotely and determine the best control input for the plant. In such a setting, imperfections of the wireless channel or time-critical requirements of the underlying control loops impose challenges for both communication and control. In particular, the end-to-end delay or information loss induced by the network reduces the precision of control actions, hence degrading the control quality.

Age-of-information (AoI) is a metric that measures the information freshness about a remote process at the receiver side. It is defined as the time elapsed since the generation of the latest received information [2], [3]. As the name suggests, AoI increases linearly in time for all types of applications until a new update is received. AoI combines packet generation frequency, end-to-end delay, and packet loss in a single metric. For example, the absence of information increases AoI on the controller side irrespective of the cause: high delay, packet

loss, or a low information update frequency. Due to its ability to connect control and communication layers, AoI has been widely adopted as a cross-layer metric by the researchers from both communities.

A. Related Work

The initial research interest on AoI has concentrated solely on the communication side. The effect of various queuing and packet management techniques has been studied for single-hop [2]–[6] and multi-hop [7], [8] transmissions. [8]–[10] study the AoI minimization as a resource scheduling problem and provide average optimal policies, i.e., policies that are optimal to achieve minimum average AoI in the network. The notion of age penalty functions, that is, staircase or non-linear evolution of AoI over time, has been proposed to reflect the heterogeneous demands of different applications sharing a communication network as in typical industrial internet of things or CPS scenarios. [6] conducts a study on update policies for age penalty minimization in the presence of random transmission delays. They consider a single-user link and show that generating a new information right after the delivery of the previous packet is not optimal under certain conditions. In [11], authors expand the concept of AoI by introducing Cost of Update Delay and Value of Information of Update Delay metrics as a function of information staleness. They consider single-user link with first-come-first-served (FCFS) queue discipline for a M/M/1 model.

With the introduction of application-dependent age penalty functions, the AoI metric has been adopted in NCS research as well. In [12] authors consider multiple heterogeneous feedback control loops communicating over a base station where the network resources are limited. They define estimation error as a function of AoI and compare age- and value of information (VoI) scheduling of wireless resources. In the context of NCS, they show that optimal AoI scheduling leads to worse estimation accuracy than greedy VoI scheduling. Similarly, in [13] authors study the estimation problem for stochastic processes in a single-loop and single-hop scenario. They propose average optimal sampling strategies for age penalty minimization for selected age and transmission cost functions. [14] proposes a heuristic scheduling policy for wireless NCS where network resources are allocated greedily to the sub-systems with the highest expected error.

B. Contributions

The key contribution of this work is the derivation of an approximately optimal scheduling strategy for a *multi-loop NCS*, where the centralized scheduler takes a decision based on observed AoI. We re-use the concept of age penalty function, but, unlike the related work, we derive *control-aware age penalty* which take into account control system dynamics. By modeling the system as a Markov decision process (MDP) with AoI as its states, we further derive a γ -optimal stationary scheduling policy to minimize the network-induced estimation error over an infinite horizon. We show that our scheduling policy outperforms the related work age-optimal scheduling proposed in [10] and greedy error scheduling proposed in [14] at minimizing the average network-induced error.

The remainder of the paper is organized as follows. Section II presents the system model and formulates the problem. Section III introduces the optimal scheduler design. Section IV presents simulative evaluation and benchmarking of the proposed scheduler, and Section V concludes the paper.

C. Notations

Throughout the paper \mathbf{v}^T and \mathbf{M}^T stand for the transpose of a vector \mathbf{v} and a matrix \mathbf{M} , respectively. $\text{tr}(\cdot)$ is the trace operator. The expected value of a random variable X is denoted by $\mathbb{E}[X]$. The euclidean norm of a vector \mathbf{v} is denoted by $\|\mathbf{v}\|$ with $\|\mathbf{v}\| = \sqrt{\mathbf{v}^T \mathbf{v}}$. The normal distribution with mean μ and standard deviation σ is denoted by $\mathcal{N}(\mu, \sigma^2)$. $P(A | B)$ denotes the occurrence probability of an event A given B . \mathbb{N}_+ is the set of positive natural numbers.

II. SCENARIO AND PROBLEM STATEMENT

Suppose N independent, linear time-invariant (LTI) control systems sharing a wireless communication network. Each sub-system i consists of a plant \mathcal{P}_i , a sensor \mathcal{S}_i and a controller \mathcal{C}_i with an estimator \mathcal{E}_i . Each sensor \mathcal{S}_i samples the plant periodically and transmits the latest state information to \mathcal{C}_i . On the controller side, \mathcal{E}_i estimates periodically the current plant state based on the latest received information. Estimated state is then used by \mathcal{C}_i to calculate the next control input. We assume ideal links between each controller and plant pair while sensor is operating remotely over the shared wireless channel as illustrated in Fig. 1.

A. Network Model

All transmissions are managed by a centralized entity called *scheduler*. In each time slot t , the scheduler decides which $R \leq N$ subsystems are allocated transmission resources. Transmission resources are assumed non-overlapping, thus, contention-free medium access without collisions is guaranteed. Let us define a scheduling decision variable $q_i[t] \in \{0, 1\}$ being $q_i[t] = 1$ if the sub-system i is allowed to transmit in slot t and $q_i[t] = 0$ otherwise. Therefore, the following holds for all time slots:

$$\sum_{i=1}^N q_i[t] \leq R \quad , \forall t. \quad (1)$$

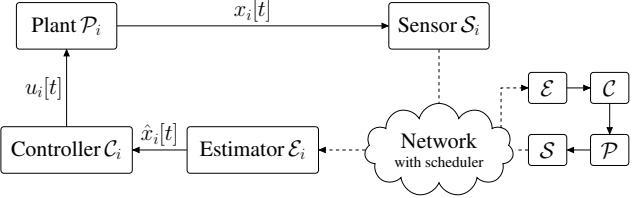


Figure 1: Considered scenario with N heterogeneous LTI networked control systems. Solid lines indicate ideal controller-to-plant and plant-to-sensor links. Sensor-to-controller link is closed over a shared wireless channel. Medium access is granted by a scheduler.

The smallest time unit we consider is a time slot which is also equal to the sampling period of sub-systems. In other words, when a sensor is granted channel access, it always transmits its latest state information. We assume a constant delay of one time slot, i.e., any transmission scheduled in time slot t is assumed to be received in $t + 1$. We leverage the results in [4] and suppose that sensors replace their packet in the transmit buffer with a more recent measurement if they do not receive any resource grant.

We consider *packet erasure channel* where each sub-system i has a constant probability p_i for a successful packet reception and decoding. That is, if we define a success indicator variable $\delta_i[t] \in \{0, 1\}$ representing a failed $\delta_i[t] = 0$ or a successful $\delta_i[t] = 1$ reception of a packet, then the success probability is given by $P(\delta_i[t] = 1 | q_i[t] = 1) = p_i \in (0, 1]$. Similarly, packet loss probability is $P(\delta_i[t] = 0 | q_i[t] = 1) = 1 - p_i$. For the sake of completeness, note that $P(\delta_i[t] = 1 | q_i[t] = 0) = 0$ and $P(\delta_i[t] = 0 | q_i[t] = 0) = 1$.

B. Age-of-information Model

Age-of-information of a sub-system i is defined as the time elapsed since the generation of the freshest information available at \mathcal{C}_i . If we denote AoI of the i -th sub-system by $\Delta_i[t]$, it can be obtained from $\Delta_i[t] \triangleq t - r_i$ with $r_i = \sup\{\tau \in \mathbb{N} : \tau < t, q_i[\tau] \cdot \delta_i[\tau] = 1\}$ with time evolution:

$$\Delta_i[t+1] = \begin{cases} 1, & \text{if } q_i[t] \cdot \delta_i[t] = 1 \\ \Delta_i[t] + 1, & \text{otherwise.} \end{cases} \quad (2)$$

Due to the constant one slot delay, a successfully received packet resets AoI to one, hence $\Delta_i[t] \in \mathbb{N}_+, \forall i, t$. Otherwise, AoI increases linearly as time passes.

C. Control System Model

The system state of the i -th plant \mathcal{P}_i evolves by the following LTI model over discrete time t :

$$\mathbf{x}_i[t+1] = \mathbf{A}_i \mathbf{x}_i[t] + \mathbf{B}_i \mathbf{u}_i[t] + \mathbf{w}_i[t], \quad (3)$$

with state vector $\mathbf{x}_i \in \mathbb{R}^{n_i}$, control input $\mathbf{u}_i \in \mathbb{R}^{m_i}$, system and input matrices $\mathbf{A}_i \in \mathbb{R}^{n_i \times n_i}$ and $\mathbf{B}_i \in \mathbb{R}^{n_i \times m_i}$, respectively. We assume the identically and independently distributed (i.i.d.) disturbance vector $\mathbf{w}_i[t] \in \mathbb{R}^{n_i}$ to have zero-mean Gaussian distribution with diagonal covariance matrix Σ_i . Each sub-system is initialized with $\mathbf{x}_i[0] \sim \mathcal{N}(\mathbf{0}, \Sigma_i)$.

In the estimator \mathcal{E}_i , the current system state is estimated given the latest observed state, i.e., $\hat{\mathbf{x}}_i[t] \triangleq \mathbb{E}[\mathbf{x}_i[t] | \mathbf{x}_i[t - \Delta_i[t]]]$ as:

$$\hat{\mathbf{x}}_i[t] = \mathbf{A}_i^{\Delta_i[t]} \mathbf{x}_i[t - \Delta_i[t]] + \sum_{q=1}^{\Delta_i[t]} \mathbf{A}_i^{q-1} \mathbf{B}_i \mathbf{u}_i[t-q] \quad (4)$$

with $\hat{\mathbf{x}}_i \in \mathbb{R}^{n_i 1}$. The control input is obtained from the control law $\mathbf{u}_i[t] = -\mathbf{L}_i \hat{\mathbf{x}}_i[t]$ with feedback gain matrix $\mathbf{L}_i \in \mathbb{R}^{m_i \times n_i}$ ². We leverage our results in [12] and define state estimation error of \mathcal{E}_i at t , i.e., $\mathbf{e}_i[t] \triangleq \mathbf{x}_i[t] - \hat{\mathbf{x}}_i[t]$, as a function of AoI:

$$\mathbf{e}_i[t] = \sum_{q=1}^{\Delta_i[t]} \mathbf{A}_i^{q-1} \mathbf{w}_i[t-q]. \quad (5)$$

Analogously, the mean-square norm of $\mathbf{e}_i[t]$ is derived as:

$$\mathbb{E} [\|\mathbf{e}_i[t]\|^2] = \sum_{r=0}^{\Delta_i[t]-1} \text{tr} ((\mathbf{A}_i^T)^r \mathbf{A}_i^r \Sigma_i) \triangleq g_i(\Delta_i[t]). \quad (6)$$

It is important to note that $\Delta_i[t]$ is the only time dependent variable in (6) since we deal with time-invariant control systems. In addition, $\mathbb{E} [\|\mathbf{e}_i[t]\|^2]$ is an increasing function in age. Therefore, by providing the knowledge of the current state to \mathcal{E}_i (or any fresher information than the one it already has), the reduction in $\mathbb{E} [\|\mathbf{e}_i[t]\|^2]$ is guaranteed.

D. Scheduling Problem Formulation

We model the centralized scheduling problem for N networked control systems as a *markov decision process (MDP)* [15]. We define the vector of sub-systems' AoIs as a state of the MDP, i.e., $\mathbf{s}[t] \triangleq (\Delta_1[t], \Delta_2[t], \dots, \Delta_N[t])$. Note that the state space \mathcal{S} is equal to \mathbb{N}_+^N thus by definition is a countable N -dimensional infinite set. The action space is defined as the set of all possible scheduling decisions of length N , i.e., $\mathcal{A}' = \{0, 1\}^N$. Here, the i -th index of $\mathbf{a}[t]$ being 1 indicates that the i -th sub-system is scheduled for transmission, i.e., $a_i[t] = q_i[t], \forall i \in \mathbb{N}^+$. Note that, since the scheduler has to select R out of N sub-systems to transmit, we define the set of admissible actions as $\mathcal{A} = \{\mathbf{a} \in \mathcal{A}' : \sum_{i=1}^N a_i \leq R\}$.

Given the states and admissible actions as above, we are interested in deterministic scheduling policies π within the set of admissible policies Π that maps states into admissible actions, i.e., $\Pi : \mathcal{S} \mapsto \mathcal{A}$. Our goal is to find an optimal policy $\pi^* \in \Pi$ that minimizes the *total mean-square estimation error over an infinite horizon*, given any initial state $\mathbf{s}_0 = (\Delta_1[0], \dots, \Delta_N[0])$:

$$J_{\pi^*}(\mathbf{s}_0) = \min_{\pi \in \Pi} J_{\pi}(\mathbf{s}_0) \quad (7)$$

¹The estimator requires the history of control inputs applied to the plant. Since we assume stationary control policy throughout the paper, this assumption does not impose any additional communication between the controller and estimator.

²We assume that the control system is stabilizable and the stabilizing gain \mathbf{L}_i exists.

with

$$J_{\pi}(\mathbf{s}_0) = \lim_{T \rightarrow \infty} \mathbb{E}_{\pi} \left[\sum_{t=0}^{T-1} \sum_{i=1}^N \gamma^t g_i(\Delta_i[t]) \right], \quad (8)$$

where \mathbb{E}_{π} indicates the expected mean-square error under policy π . Here, $\gamma \in (0, 1)$ is a positive scalar value called *discount factor*.

Definition 1. A scheduling policy is called γ -optimal if it minimizes the discounted mean-square estimation error over an infinite horizon, i.e., the right hand side of (8) [15].

The role of γ is to adjust the importance of short-term or long-term penalties which are the total mean-square error norms as in (6). A penalty received in t steps in the future is worth only γ^t times what it would cost if it were received immediately. Hence, as γ approaches 1, the scheduler weights future penalties more and thereby becomes more farsighted.

III. SCHEDULER DESIGN

A. Approximating Sequence of the MDP

The state space of the given MDP is infinite. Since we can not iterate through infinite states in practice, we work with an *augmentation type approximating sequence (ATAS)* of the MDP [16]. First, let us define a nonempty finite state space $\mathcal{S}_M \subset \mathcal{S}$ that divides \mathcal{S} into two disjoint sets, where states with $\Delta_i \leq M, \forall i$ belong to the finite set \mathcal{S}_M and states with $\Delta_i > M, \exists i$ belong to $\mathcal{S} \setminus \mathcal{S}_M$. In other words, as long as AoIs of all sub-systems are smaller than $M+1$, the state is in \mathcal{S}_M . If the AoI of at least one sub-system is larger than M , the state is in the infinite set $\mathcal{S} \setminus \mathcal{S}_M$.

Suppose that in state $\mathbf{s} \in \mathcal{S}_M$ action \mathbf{a} is chosen. For any $\mathbf{s}' \in \mathcal{S}_M$ the transition probability from \mathbf{s} to $\mathbf{s}' P_{ss'}(\mathbf{a})$ is unchanged. Now let us define the *excess probability* as the transition probability from a state $\mathbf{s} \in \mathcal{S}_M$ to a state $\hat{\mathbf{s}} \notin \mathcal{S}_M$ outside the finite space with $P_{s\hat{s}}(\mathbf{a}) \geq 0$. In this case we redistribute all excess probabilities to the states \mathbf{s}' in \mathcal{S}_M according to a probability distribution $f_{s'}(\mathbf{s}, \mathbf{a}, \hat{\mathbf{s}}, M)$. If we denote the resulting MDP by MDP_M the definition of ATAS follows from [16].

Definition 2. A sequence MDP_M is an augmentation type approximating sequence of an MDP iff for each $\hat{\mathbf{s}} \notin \mathcal{S}_M$, given $\mathbf{s}, \mathbf{s}' \in \mathcal{S}_M$ and $\mathbf{a} \in \mathcal{A}$, there exists a probability distribution $f_{s'}(\mathbf{s}, \mathbf{a}, \hat{\mathbf{s}}, M)$ such that

$$P_{ss'}(\mathbf{a}; M) = P_{ss'}(\mathbf{a}) + \sum_{\hat{\mathbf{s}} \in \mathcal{S} \setminus \mathcal{S}_M} P_{s\hat{s}}(\mathbf{a}) f_{s'}(\mathbf{s}, \mathbf{a}, \hat{\mathbf{s}}, M). \quad (9)$$

By this definition we introduce the following ATAS for our scenario: Let $\tilde{\Delta}_i^{(M)}[t]$ be the *augmented age-of-information* for the i -th sub-system with $\tilde{\Delta}_i^{(M)}[t] = \min(\Delta_i[t], M)$. That is, we truncated AoI to M if its value exceeds M . Thus the new state space becomes finite with $\mathcal{S}_M = \{1, \dots, M\}^N$. Fig. 2 illustrates the resulting MDP_M for a single sub-system i given the scheduler decides to schedule the same user at age M , i.e., $a_i[t] = 1$. In the remainder of this paper we refer to

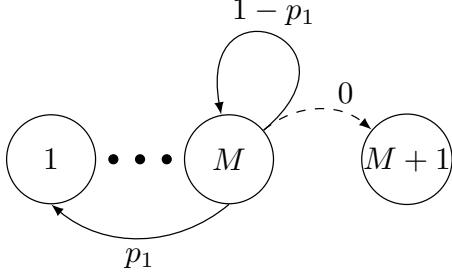


Figure 2: Augmentation type approximating sequence for the exemplary MDP with one sub-system $i = 1$ ($N = 1$) with maximum augmented age M . Given the current state $s[t] = M$ and action $a[t]$ with $a_1[t] = q_1[t] = 1$, we redistribute the excess probability of a transition from M to $M + 1$ back to state M , i.e., $P(s[t+1] = M \mid s[t] = M, a_1[t] = 1) = 1 - p_1$.

the augmented states without changing the notation, $s \in \mathcal{S}_M$ to avoid visual clutter.

As we have the states, actions, and transition probabilities of the finite MDP_M , let us define the immediate cost associated with each augmented state and action pair as:

$$C(s[t], a[t]) = \sum_{i=1}^N g_i(\tilde{\Delta}_i^{(M)}[t]), \quad (10)$$

with states $s \in \mathcal{S}_M$. Note that cost does not depend on the action taken. One can easily introduce an additive communication cost term to the equation above in case different subsystems demand different amount of network resources.

B. γ -Optimal Discounted Error Scheduler (DES)

Now for the given augmentation level M , we propose a γ -optimal stationary deterministic scheduling policy that minimizes the discounted infinite horizon problem from (7) with cost approximations from (10)³. To that end, we apply the standard *value iteration* approach with *dynamic programming* [15].

Each state is associated with an initial value function $J_0(s)$. We iterate through all state-action pairs (s, a) and their possible next states s' as:

$$\begin{aligned} J_{k+1}(s) &= \min_{a \in \mathcal{A}} \{ \mathbb{E}[C(s, a) + \gamma J_k(s')] \} \\ &= \min_{a \in \mathcal{A}} \{ C(s, a) + \gamma \mathbb{E}[J_k(s')] \} \\ &= \min_{a \in \mathcal{A}} \left\{ C(s, a) + \gamma \sum_{s' \in \mathcal{S}_M} P_{ss'}(a) J_k(s') \right\}, \end{aligned} \quad (11)$$

while minimizing the right-hand-side of the Bellman equation (11). Here, $P_{ss'}(a)$ denotes the success probability from a current state s to a successor state s' given action a , i.e., $P(s[t+1] = s' \mid s[t] = s, a[t] = a)$. In [15] it is shown that, as k goes to infinity, value functions will converge to an optimal cost, i.e., $\lim_{k \rightarrow \infty} J_k = J^*$, for any initial

³Note that, the proposed algorithm solves the approximated problem optimally for a given γ and an augmentation level M .

Algorithm 1: Value iteration

```

1  $J(s) \leftarrow 0$  for all augmented states  $s \in \mathcal{S}_M$ 
2 do
3    $u \leftarrow 0$ 
4   foreach  $s \in \mathcal{S}_M$  do
5     Update  $J^+(s)$  as in (11);
6      $u \leftarrow \max(u, |J^+(s) - J(s)|)$  ;
7      $J(s) \leftarrow J^+(s)$ ;
8   end
9 while  $u > \theta$ ;
10  $\pi^*(s) \leftarrow \arg \min_{a \in \mathcal{A}} J(s)$ 
Output:  $\pi^*$ 

```

value function J_0 . In this case actions solving (11) constitute the stationary optimal policy π^* . Algorithm 1 illustrates the practical implementation of the value iteration algorithm with $J_0(s) = 0, \forall s$. A convergence threshold θ is used to break the loop. The precision of the final value functions increases if lower values of θ are selected.

IV. EVALUATION

In this section, we present a simulative analysis and compare our proposed scheduler to selected benchmark schedulers from the existing literature. Section IV-B describes the simulation setup and parameters. Sections IV-A and IV-C introduce the benchmark schedulers and discuss simulation results.

A. Benchmarks

In order to compare the performance of our proposed scheduler, we choose two representative schedulers from the state-of-the-art: the γ -optimal AoI scheduler proposed in [10] and the control-aware greedy scheduler proposed in [14].

1) *γ -Optimal AoI Scheduler (AoIS)*: As the name suggests, the γ -optimal AoI scheduler employs the total age-of-information in the network as the immediate cost of a state:

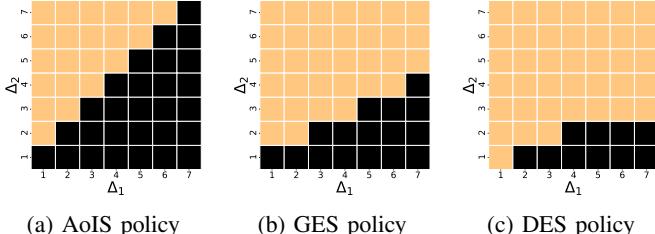
$$C_{AoI}(s[t], a[t]) = \sum_{i=1}^N \tilde{\Delta}_i^{(M)}[t]. \quad (12)$$

The γ -optimal AoI policy can be obtained by simply replacing the cost function in Alg. 1 with the new one from (12). As we are going to show in Section IV-C, the resulting policy provides the highest information freshness among the selected schedulers.

2) *Greedy Error Scheduler (GES)*: The *greedy error scheduler* prioritizes sub-systems with the highest mean squared error:

$$a[t] = \arg \max_{a \in \mathcal{A}} \left\{ \sum_i^N p_i g_i(\Delta_i[t]) \right\}. \quad (13)$$

Here, we introduce the success probabilities as a weighting factor for a *channel-aware* version of the proposed solution in [14]. One can simply ignore it for a *control-aware* implementation only. Fig. 3 illustrates optimal policies of AoIS, GES, and DES for a simplified scenario with $N = 2$



(a) AoIS policy (b) GES policy (c) DES policy

Figure 3: Scheduling policies of (a) AoIS, (b) GES, and (c) DES for two sub-systems with state matrices $A_1 = 1.1$ and $A_2 = 1.3$, loss probabilities $p_1 = p_2 = 0.5$, and with a single resource per time slot $R = 1$. Parameters: $\gamma = 0.9$, $M = 7$, $\theta = 0.1$. Light squares correspond to states $s = (\Delta_1, \Delta_2)$ where user 1 is scheduled ($a_1 = 1, a_2 = 0$), and dark squares correspond to states where user 2 is scheduled ($a_1 = 0, a_2 = 1$).

heterogeneous control loops and one resource $R = 1$. Fig. 3a is symmetrical since AoIS is independent of control system parameters. Fig. 3b and Fig. 3c highlight the asymmetrical distribution of resources among both control loops coming from their discrepancy in system instability.

B. Simulation Set-Up

Suppose the system comprises $N = 5$ heterogeneous control sub-systems sharing a wireless channel with Bernoulli-distributed random packet loss. We assume that all control sub-systems have different plant system dynamics, with increasingly unstable system matrices. To provide intuitive and illustrative results, we consider scalar sub-systems in the evaluation. To that end, we select $A_i = \{1.1, 1.3, 1.5, 1.7, 1.9\}$ for $i = \{1, 2, 3, 4, 5\}$, respectively. We assume equal packet success probability for each sub-system i , i.e., $p_i = 0.9, \forall i$. The feedback gain matrix is given by $L_i = A_i$ which corresponds to the deadbeat control strategy. Input and noise covariance matrices are selected as $B_i = 1$ and $\Sigma_i = 1$, respectively. We further consider single transmission resource per time slot, i.e., $R = 1$. The augmentation level for the approximating sequence of the original MDP is $M = 25$, and the stop condition for Alg. 1 is $\theta = 0.1$. Simulations are run for $T = 20000$ time slots. The discount factor is varied, $\gamma = \{0.1, 0.2, \dots, 0.9\}$ to study its effect on the resulting average quadratic error per time slot, \bar{e} , and average age per time slot, $\bar{\Delta}$, given by:

$$\bar{e} = \frac{1}{T} \frac{1}{N} \sum_{t=0}^{T-1} \sum_{i=1}^N \|e_i[t]\|^2 \quad (14)$$

$$\bar{\Delta} = \frac{1}{T} \frac{1}{N} \sum_{t=0}^{T-1} \sum_{i=1}^N \Delta_i[t] \quad (15)$$

C. Results

We evaluate the performance of the proposed scheduler using average quadratic error, i.e., \bar{e} , as the main metric quantifying estimation accuracy and average age-of-information, i.e., $\bar{\Delta}$, measuring the information freshness⁴. In Fig. 4, we plot the

⁴Results were generated via Monte Carlo experiments with 100 repetitions for any given γ . Each experiment consists of $T = 20000$ time slots.

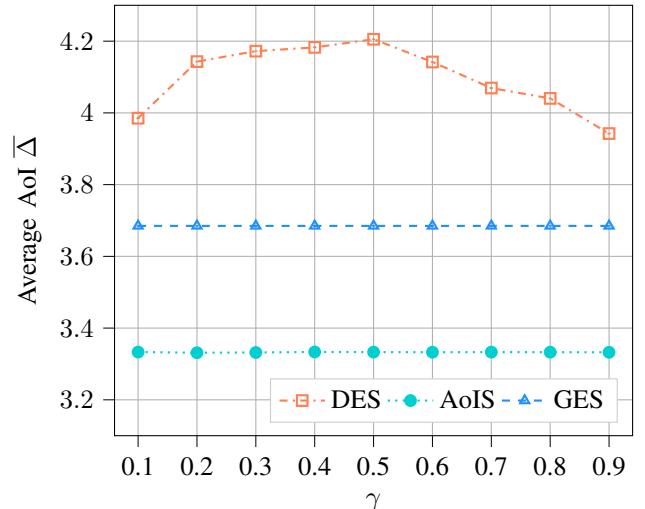


Figure 4: Achieved average age-of-information, i.e., $\bar{\Delta}$, vs. discount factor γ for discounted error (DES), γ -optimal AoI (AoIS) and greedy error (GES) schedulers. Out of five control loops, only one sub-system is allowed to transmit simultaneously due to single channel assumption, i.e., $N = 5$ and $R = 1$. Success probability of each transmission is $p_i = 0.9 \forall i$. 95% confidence intervals are too small to be plotted.

average age-of-information $\bar{\Delta}$ against increasing values of the discount factor γ , corresponding to the decreasing discount of future state. As expected, we observe that the AoI scheduler (AoIS) outperforms the greedy error scheduler (GES) and the γ -optimal discounted error scheduler (DES) with respect to information freshness at the receiver nodes. An interesting finding is that DES results in a concave shape with respect to $\bar{\Delta}$ as γ increases. This can be elaborated with the help of Fig. 5 which illustrates average network shares of individual sub-systems with different task-criticalities. A network share $\alpha_i\%$ indicates that the i -th sub-system was granted channel access $\alpha_i\%$ of the time in average. In the figure, we plot network shares of individual classes $i \in \{1, \dots, 5\}$ with system matrices $A_i = \{1.1, 1.3, 1.5, 1.7, 1.9\}$ against varying γ . One can observe that the discount factor has different impact on different classes. Increasing γ from 0.1 to 0.5 has no impact on classes two, four, and five, but it prioritizes class 3 over class 1 (hence, its value is increasing). Therefore, the gap between the sub-systems with $A_1 = 1.1$ and $A_3 = 1.5$ grows. This leads to reduced fairness with respect to resource distribution and thus the growth in AoI follows for the DES. Further increase of γ , however, reverses this behavior and, at the same time, increases the priority and network share of class 2. Thus the resource distribution among all classes get closer and the decrease in AoI follows. On the other hand, we see that the GES allocates resources in a relatively fairer fashion compared to the DES. This results in a lower average AoI for GES in Fig. 4. In addition, the control system unawareness of AoIS leads to equal distribution of resources among different classes. In other words, network shares of all five loops coincide at 20%. However, we ignore the corresponding figure

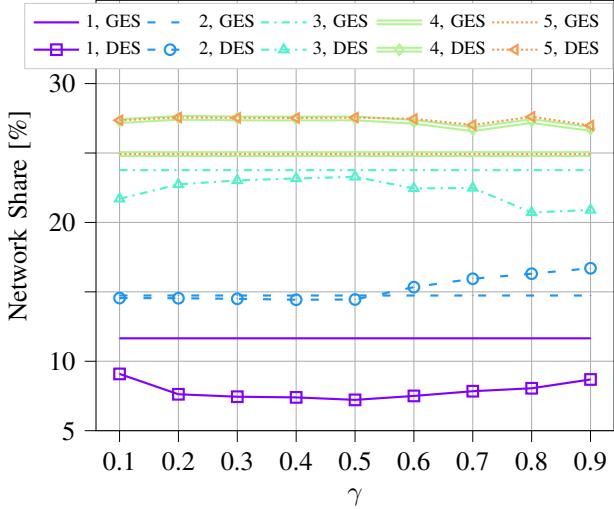


Figure 5: Illustration of network resources shares for different sub-system classes over 100 simulation runs. Numbering in the legend indicate the class i with the respective system matrices $A_{1,2,3,4,5} = \{1.1, 1.3, 1.5, 1.7, 1.9\}$. Lines with and without markers belonging to the same class i illustrate the average network resource share granted to i by the discounted (DES) and greedy (GES) error schedulers, respectively. 95% confidence intervals are too small to be displayed.

due to space considerations.

Next we plot our main performance metric \bar{e} against increasing γ in Fig. 6. One can observe that the AoIS leads to much higher estimation error in average due to its unawareness of control system parameters. Therefore, one can say that AoI scheduling is not suitable for NCS scenarios. On the other hand, DES achieves steadily lower error than the GES and AoIS. In addition, the error is lower for increasing discount factor. Thus, we deduce that being far-sighted and valuing future costs, as done by the DES with higher γ , achieves better estimation accuracy than rushing immediate rewards as done by the GES or low γ values.

D. Discussion on selecting M

One of the key design parameters for the proposed scheduling algorithm is the selection of M . For any given augmentation level M and number of users N , the cardinality of the state space \mathcal{S}_M equals to M^N . Hence, the scalability of the approach suffers heavily from an increase in M . On the other hand, if a larger M is selected, the approximating sequence approaches to the original MDP and may potentially lead to an improvement in performance. As a result, selection of M is a trade-off between complexity and performance for the proposed approach.

In order to investigate the impact of the augmentation level on the performance, we vary the design parameter M and obtain the stationary optimal scheduling policy as described in Algorithm 1. Fig. 7 plots our key performance metric, the average quadratic error \bar{e} , over γ for five different M values, i.e., $M \in \{15, 16, 17, 20, 25\}$. If we look at the $M = 15$ curve, we observe that it is outperformed by the other policies in terms of \bar{e} . This is a result of the relatively inaccurate

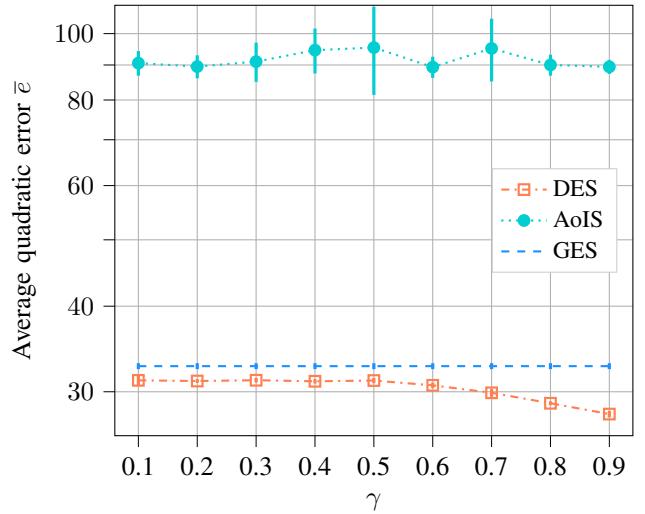


Figure 6: Achieved average quadratic error vs. discount factor γ for discounted error scheduler (DES), greedy error scheduler (GES) and age-of-information scheduler (AoIS). Out of five control loops, only one sub-system is allowed to transmit simultaneously due to single channel assumption, i.e., $N = 5$ and $R = 1$. Success probability of each transmission is $p_i = 0.9 \forall i$. Vertical error bars represent 95% confidence intervals for Monte Carlo simulations with 100 repetitions.

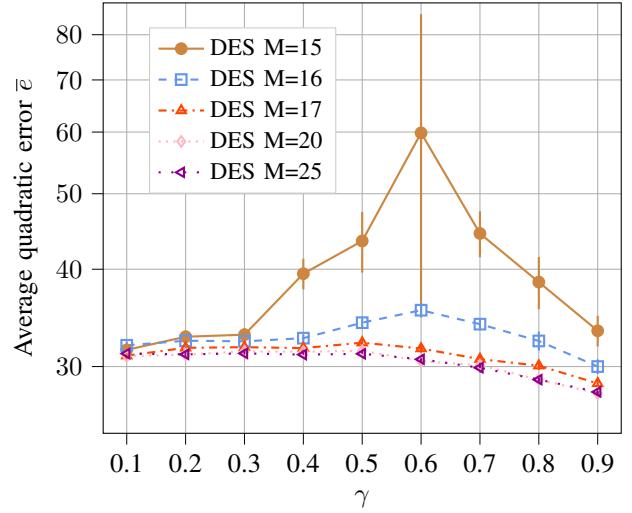


Figure 7: Achieved average quadratic error vs. discount factor γ for discounted error scheduler (DES). Each curve belongs to one of the five selected augmentation levels $M \in \{15, 16, 17, 20, 25\}$. Vertical error bars represent 95% confidence intervals for Monte Carlo simulations with 100 repetitions.

cost functions due to the approximation. That is, since we limit the maximum cost to $g_i(\tilde{\Delta}^{(M)})$ for each class i , the approximating finite sequence does not represent the original MDP as accurate as when higher M is selected. As a result, an increase in M leads to a reduction in \bar{e} . However, the performance gain diminishes after a certain augmentation level which is evident especially from $M = 20$ and $M = 25$ curves. Their overlapping behavior justifies that there is almost no performance gain by introducing any additional complexity to

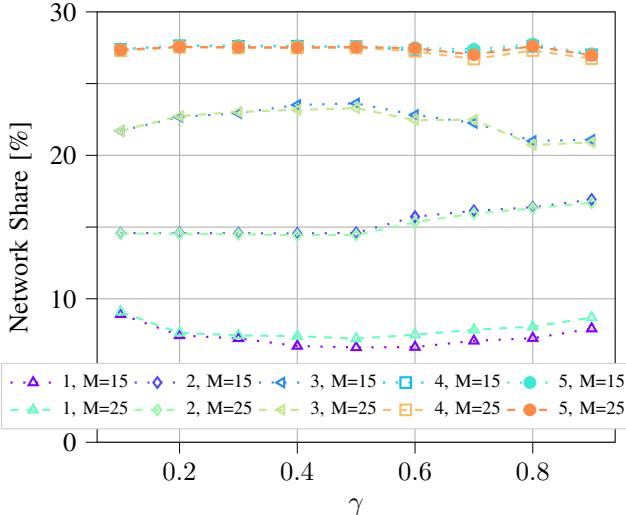


Figure 8: Illustration of network resources shares for different subsystem classes over 100 simulation runs. Numbering in the legend indicate the class i with the respective system matrices $A_{1,2,3,4,5} = \{1.1, 1.3, 1.5, 1.7, 1.9\}$. Identical markers belong to the same class i . Dotted and dashed lines illustrate the average network resource share granted to i by the discounted error scheduler (DES) with augmentation levels $M = 15$ and $M = 25$, respectively. 95% confidence intervals are too small to be displayed.

the proposed scheduler design.

Note that \bar{e} shows a concave behavior for $M = 15$. This can be explained with the help of Fig. 8 which plots average network shares of individual classes i over γ for $M = 15$ and $M = 25$. Both policies follow very similar trajectories for the same class i except for the most stable one with the system matrix $A_1 = 1.1$. This follows from overlooking potential high future costs of class $i = 1$ around $\gamma = 0.6$ due to low bounds we enforce on $g_i(\tilde{\Delta}^{(M)})$ and prioritizing $i = 3$ instead. This effect is reversed by increasing γ further. For $M = 25$ this occurs less dramatically than for $M = 15$.

V. CONCLUSIONS AND FUTURE WORKS

Age-of-Information has been recently used in both control and communication research for data scheduling in multi-user scenarios. In the context of NCS, it serves as an interface between information update frequency and control performance. In this work, we studied the centralized scheduling problem for multiple networked control loops sharing a wireless channel with random packet loss. We were able to formulate the scheduling problem as a cost minimization problem over an infinite horizon. We have used AoI as the states of the MDP and obtained a stationary scheduling policy as a function of states. We have shown that by employing control dependent age penalty functions and solving the infinite horizon cost minimization problem, we can reduce error further than the schedulers existing in the literature.

REFERENCES

- [1] E. A. Lee, "Cyber physical systems: Design challenges," in *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, 2008, pp. 363–369.
- [2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proceedings IEEE INFOCOM*, 2012, pp. 2731–2735.
- [3] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *46th Annual Conference on Information Sciences and Systems (CISS)*, 2012, pp. 1–6.
- [4] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *2014 IEEE International Symposium on Information Theory*, 2014, pp. 1583–1587.
- [5] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 1681–1685.
- [6] 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 Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.
- [7] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal information updates in multihop networks," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 576–580.
- [8] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2017, pp. 486–493.
- [9] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, pp. 844–851.
- [10] Y. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 561–565.
- [11] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 326–330.
- [12] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," in *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, 2019, pp. 109–117.
- [13] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "Aoi-penalty minimization for networked control systems with packet loss," 2019.
- [14] M. Vilgelm, O. Ayan, S. Zoppi, and W. Kellerer, "Control-aware uplink resource allocation for cyber-physical systems in wireless networks," in *23rd European Wireless Conference*, 2017, pp. 1–7.
- [15] D. P. Bertsekas, *Dynamic Programming and Optimal Control Vol. I and II*. Athena Scientific, 2017.
- [16] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons, Inc., 1999.