Goal-Oriented Communications for Remote Inference with Two-Way Delay

Cagri Ari[†], Md Kamran Chowdhury Shisher[‡], Elif Uysal[†], and Yin Sun[‡]

Department of Electrical and Electronics Engineering, Middle East Technical University, Ankara, Turkiye[†]
Department of Electrical and Computer Engineering, Auburn University, Auburn, AL[‡]
ari.cagri@metu.edu.tr, mzs0153@auburn.edu, uelif@metu.edu.tr, yzs0078@auburn.edu

Abstract—We study the design of goal-oriented communication strategies for remote inference, where an inferrer (e.g., a trained neural network) on the receiver side predicts a time-varying target signal (e.g., the position of the car in front) using the data packet (e.g., video clip) most recently received from a sensor (e.g., camera). The communication between the sensor and the receiver is carried out over a two-way channel. The objective is to minimize the expected inference error per time slot by exploiting the memory of the delay in the channel. It turns out that the optimal policy is an index-based threshold policy. The scheduler submits a packet at suitable time slots for which the index function exceeds a threshold. The index function depends on the current age of information on the receiver side and the prior knowledge about the delay in the subsequent packet transmission.

Index Terms—Goal-Oriented Communications, Markovian Delay, Two-way Delay, Remote Inference, Age of Information, Delay with Memory.

I. INTRODUCTION

Orders of magnitude enhancement is desired in terms of performance, coverage, capacity, and energy efficiency in 6G, with respect to existing cellular networks. Such striking objectives accompany the demand for intelligent applications that run based on remotely collected data, sometimes involving non-terrestrial links. Based on the envisioned coexistence of terrestrial and non-terrestrial connections in 6G, it will not be unusual for network connections to be serviced through multiple alternative paths, with highly variable delay. The relays in the network may force the utilization of a selected path for a particular application, and this path may be changed in time to manage the variety of flows in the network. Therefore, while developing efficient policies to optimize the performance of intelligent applications, considering their adaptability to significantly varying delay conditions, specifically the time dependence in the delay statistics, is important.

On the other hand, for the increasing numbers of these intelligent applications including remote monitoring, control and inference over networks, the scalability of the network architecture requires a paradigm shift in the communication system and protocol design. The paradigm shift is for networks to focus on the effective accomplishment of the sensing task at the destination side, rather than be a vehicle to solve the transmission problem [1] whose goal is to reliably transmit the data produced by a source. Solving the effectiveness problem in remote intelligence with efficient use of network resources requires combining the data generation and transmission problems whereby data samples that are useful to the computation at the destination side are delivered in a timely manner. We will refer to this paradigm as *Goal-oriented* communication.

A goal-oriented communication network discards the assumption that data arrives to the communication system interface as part of an exogenous process. This implies that the communication link can pick and choose which data packets to transmit to the source for effective computation at the destination, based on the state of the network (e.g., the delay state). As we do not expect the link layer to operate jointly with the application layer for all possible applications, a surrogate metric that the link layer can use when making the decisions of which packets to choose and send is useful. In this paper, following the model in [2], we use Age of Information (AoI) as an intermediate metric that can be tracked by the link, irrespective of the particular application. A mapping of application performance with respect to the AoI is then all that is needed for the link layer to operate in a goal oriented manner. This paper builds on the work in the series of papers [2]–[5] that first demonstrated the usefulness of AoI metric as a surrogate for effective communication for remote tracking or inference of a process.

AoI is an indicator of the freshness of the data at the destination of a data flow [6]. The AoI at time t, $\Delta(t)$, is given by the relation $\Delta(t) = t - U_t$, where U_t is the generation time of the most recently delivered data packet. This formulation is perhaps most meaningful when the data packets contain status updates, such that the most recent update makes all the past ones obsolete, yet, age and functions thereof have also served very well in capturing performance related to freshness in a variety of other applications, through optimizing nonlinear functions of age. Examples of the use of generalized functions of AoI include control system scenarios [7], [8], remote estimation [3], [4], and remote inference [2], [5], [9]. While a number of studies considered the analysis of age in queuing models, closest to the spirit of this paper is the control of age via replacement of exogenous data arrivals with the generation of data "at will" [7], [10]–[17]. A generalization of this approach is to incorporate jointly optimal sampling and scheduling policies to control not only age, but a more sophisticated end to end distortion criterion, by using age as an

auxiliary parameter [2]–[4], [9], [18]. While strikingly more demanding of analysis, these formulations take us closer to goal oriented communication system design.

Among the previous studies containing the generation of data at will [3], [4], [7], [11]–[18], most considered monotonic AoI penalty functions (where, essentially, data does not get more valuable with age). However, recent works [2], [5], [9] have demonstrated that this monotonicity assumption does not capture the age dependency of performance in some scenarios, notably for certain remote inference problems. Specifically, when the input and target data sequences in a system can be closely approximated as a Markov chain, the inference error increases as AoI grows. However, if the data sequence significantly deviates from a Markovian structure, the monotonic assumption does not hold [2], [5], [9].

The most closely related work to this paper is [2], which developed scheduling policies for remote inference. For a single source-predictor system, the scheduling policy outlined in [2] is an index-based threshold policy, where a new packet transmission is initiated when the previously transmitted packet is delivered, and an index exceeds a threshold. The paper [2] considered random IID delay for packet transmissions from the transmitter to the receiver and assumed a delay-free feedback channel from the receiver to the transmitter. In practical scenarios, there may be a feedback delay [19], [20], and this may affect the performance of the remote inference algorithm. Moreover, in the light of the application scenarios mentioned above (e.g., satellite and space communication), the IID delay may not be a good model, requiring a model that captures the memory in the delay process as in [11]. Hence the main contributions of this paper are:

- We developed an index-based threshold policy to minimize the steady-state time-average inference error of an intelligent application whose input data packets are transmitted from a distant location. Given the age penalty function corresponding to a particular application, not necessarily monotonic, the policy achieves the optimal in terms of practical performance. In other words, the policy makes the scheduler behave in a *Goal-oriented* manner.
- We extended the formulation in [2] to the Markovian delay case, with finite number of states, and remove the delay-free feedback channel assumption. Therefore, the developed policy provides a solution for systems where channel delay statistics vary significantly with memory. Such systems will likely be abundant due to the envisioned coexistence of terrestrial and non-terrestrial connections in 6G, as mentioned above.
- The simulation results highlight the significance of (i) the policy being adaptive to non-monotonic age penalty functions and (ii) considering the memory of the delay. Numerically, the generate-at-will + zero-wait policy achieves 31% to 47% higher inference error than the developed policy. In addition, the performance gain obtained by considering the memory of the delay increases up to 11% with the increasing memory of the delay.
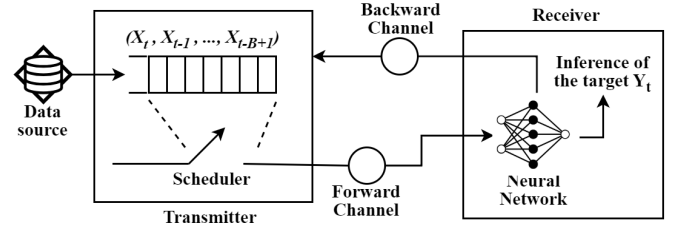

Fig. 1. The remote inference system with two-way communication where the neural network on the receiver side predicts the current value of a target variable using the most recently delivered packet.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a discrete-time remote inference system, as depicted in Fig. 1. The transmitter incorporates a data buffer and a scheduler. At each time slot t, a new data packet $X_t \in \mathcal{X}$ is sampled from the source and added to the buffer; meanwhile, the oldest is discarded. By this, the buffer contains the most recently sampled B packets $(X_t, X_{t-1}, ..., X_{t-B+1})$ at any time slot t. The scheduler can submit any packet stored in the buffer to the channel at suitable time slots. The Age of Information (AoI), $\Delta(t)$, is the time difference between the current time t and the generation time $t - \Delta(t)$ of the most recently delivered packet $X_{t-\Delta(t)}$. The value of AoI depends on (i) which packet in the buffer is selected by the scheduler during the submission of the most recently delivered packet and (ii) communication delays. A trained neural network on the receiver side takes the packet $X_{t-\Delta(t)}$ and the AoI $\Delta(t)$ as inputs and produces an action $a \in \mathcal{A}$ to infer the current value of a target signal $Y_t \in \mathcal{Y}$ at each time slot t. Given the AoI $\Delta(t) = \delta$, the average inference error is [2]

$$h(\delta) = \mathbb{E}_{Y, X \sim P_{Y_t, X_{t-\delta}}}[L(Y, \phi(\delta, X_{t-\delta})], \quad (1)$$

where $\phi : \mathbb{Z}^+ \times \mathcal{X} \mapsto \mathcal{A}$ is the function representing the neural network, $L : \mathcal{Y} \times \mathcal{A} \mapsto \mathbb{R}$ is a loss function, and $P_{Y_t, X_{t-\delta}}$ is the joint distribution of the target Y_t and the packet $X_{t-\delta}$. $L(y, a)$ is the incurred loss if the output a is used for prediction when the target signal $Y_t = y$

The loss function L is determined by the purpose of the application. For example, a quadratic loss function $L_2(y, \hat{y}) = \|y - \hat{y}\|_2^2$ is used in neural network based minimum mean-squared estimation, where the action $a = \hat{y}$ is an estimate of the target $Y_t = y$ and $\|y\|_2$ is the Euclidean norm. In softmax regression (i.e., neural network based maximum likelihood classification), the action $a = \mathcal{Q}_Y$ is a distribution of Y_t, and the loss function $L_{log}(y, \mathcal{Q}_Y) = -log\mathcal{Q}_Y(y)$ is the negative log-likelihood function of the value $Y_t = y$.

The communication is carried out between the transmitter and receiver over a two-way channel. The two-way channel has a finite number of states $\mathcal{C} = (1, 2, ..., C)$. Each state c has a distinct transmission and feedback delay distributions represented by the discrete random variables Q_c and R_c, respectively, where $\mathbb{E}[Q_c] < \infty$ and $\mathbb{E}[R_c] < \infty$ for all c. The delay duration is at least one time slot for each packet or

feedback transmission. In practice, the state c of the channel could be used to characterize time varying physical phenomena that affect the transmission and feedback delay distributions, Q_c and R_c. Examples of such phenomena incluude (i) the channel strength in slow fading, (ii) the on/off availability of a light-of-sight propagation path, and (iii) long or short network routes that packets may go through. Upon the delivery of the i-th packet, the receiver is able to detect the channel state c_i valid during the transmission of packet i. Then, the channel state c_i is fed back to the transmitter along with an ACK. Both channels are reliable; that is, no packet or message is lost during the transmission.

Recent studies [2], [21] have exhibited scenarios where the inference error in (1) is not monotonic with age, where fresher packets do not always lead to a smaller inference error than stale ones. In other words, in certain scenarios, an older packet with AoI > 0 can be more relevant than a fresher packet, in the sense that this packet can lead to a smaller inference error than a packet with AoI $= 0$ [2], [21]. In these cases, the "generate-at-will" approach which samples a fresh packet upon making a transmission decision, will be general enough to achieve optimal performance. In recognition of this fact, the "selection-from-buffer" model was proposed in [2], [9]. In this paper we adopt this latter model.

The system starts to operate at time slot $t = 0$. We assume that the buffer initially contains the packets $(X_0, X_{-1}, ..., X_{-B+1})$. By this, the buffer is kept full at all time slots $t \geq 0$. At time slot S_i, the packet $X_{S_i - b_i}$ is submitted to the channel, which is the $(b_i + 1)$-th freshest packet in the buffer, with $b_i \in \mathcal{B} = \{0, 1, \ldots, B - 1\}$. The packet is delivered to the receiver at time slot D_i, and the transmitter receives the corresponding ACK and channel state c_i at time slot A_i, such that $S_i < D_i < A_i$. We assume that the transmitter always waits for the ACK before submitting a new packet to the channel. In other words, the scheduler always remains silent at the time slots between S_i and A_i for all i. Let i-th epoch be composed of the time slots in the interval $[A_{i-1}, A_i)$. The channel state in the i-th epoch, c_i, evolves following a finite-state ergodic Markov chain with the transition probabilities p_{ij}, where $i, j \in \mathcal{C}$, and p_{ij} is the probability of a transition from state i to state j. The Markov chain makes a single transition at the time slots A_i and none otherwise. Let $T_i = D_i - S_i$ and $F_i = A_i - D_i$ be the incurred transmission and feedback delays in i-th epoch, respectively. The distributions of T_i and F_i are determined by the channel state in the i-th epoch. That is, $p_{T_i|(c_i=c)}(q) = p_{Q_c}(q)$ and $p_{F_i|(c_i=c)}(r) = p_{R_c}(r)$, where $p_K(k)$ is the probability mass function (PMF) of the dicrete random variable K. The AoI at time slot t is given by

$$\Delta(t) = t - S_i + b_i, \text{ if } D_i \leq t < D_{i+1}. \quad (2)$$

The initial conditions of the system are $S_0 = 0$, $\Delta(0) < \infty$, $c_0 \in \mathcal{C}$, and $b_0 \in \mathcal{B}$.

We strive to minimize the expected inference error per time slot in steady-state. Accordingly, the scheduler on the transmitter side has to determine (i) when to submit a packet to the channel and (ii) which packet in the buffer to submit. A scheduling policy is a tuple (f, g), where the buffer position sequence $f = (b_1, b_2, ...)$ controls which packet in the buffer to submit in each epoch, and the packet submission time sequence $g = (S_1, S_2, ...)$ specifies when to submit the packet to the channel and start the transmission in each epoch. Let Π be the set of all policy tuples (f, g). Then, the problem formulation is given by

$$h_{opt} = \inf_{(f,g)\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{(f,g)}\left[\sum_{t=0}^{T-1} h(\Delta(t))\right], \quad (3)$$

where $h(\Delta(t))$ is the inference error at time slot t, and h_{opt} is the optimum value of (3).

This problem formulation extends the single-source scheduling problem in [2] to more practical scenarios by considering the memory of the transmission delay and by removing the delay-free feedback assumption.

III. Optimal Scheduling Policy

We provide an optimal scheduling policy to (3) in two steps. First, we consider a class of policy tuples $(f_\psi, g) \in \Pi$ such that $f_\psi = (b_1, b_2, \ldots)$ is a stationary deterministic buffer position sequence, where

$$b_{i+1} = \psi(c_i), \quad (4)$$

for all $i = \{0, 1, 2, \ldots\}$, and Ψ is the set of all mapping functions $\psi : \mathcal{C} \mapsto \mathcal{B}$. Given a stationary deterministic buffer position sequence $f_\psi = (\psi(c_0), \psi(c_1), \ldots)$, we first optimize the packet submission time sequence $g = (S_1, S_2, \ldots)$:

$$h_{\psi,opt} = \inf_{(f_\psi, g)\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{(f_\psi, g)}\left[\sum_{t=0}^{T-1} h(\Delta(t))\right], \quad (5)$$

where $h_{\psi,opt}$ is the optimum value of (5). In the second step, we present an optimal scheduling policy for the original problem (3) using the solution to (5).

The optimal packet submission time sequence g for (5) will be described by using an index function defined by

$$\gamma(\delta, c) = \inf_{\nu\in\mathbb{Z}^+} \frac{1}{\nu} \sum_{k=0}^{\nu-1} \mathbb{E}\left[h(\delta + T_{i+1} + k)\Big| c_i = c\right], \quad (6)$$

where $\delta \in \mathbb{Z}^+$ and $c \in \mathcal{C}$.

Theorem 1. *If $|h(\delta)| < M$ for all $\delta = 1, 2, \ldots$, then the packet submission time sequence $g = (S_1(\beta_\psi), S_2(\beta_\psi), ...)$ is an optimal solution to* (5)*, where*

$$S_{i+1}(\beta_\psi) = \min\{t \geq A_i(\beta_\psi) : \gamma(\Delta(t), c_i) \geq \beta_\psi\}, \quad (7)$$

and β_ψ is the unique root of

$$\mathbb{E}\left[\sum_{t=A_i(\beta_\psi)}^{A_{i+1}(\beta_\psi)-1} h(\Delta(t))\right] - \beta_\psi\mathbb{E}[A_{i+1}(\beta_\psi)-A_i(\beta_\psi)] = 0, \quad (8)$$

where $A_i(\beta_\psi) = S_i(\beta_\psi) + T_i + F_i$ is the i-th ACK feedback time and $\Delta(t) = t - S_i(\beta_\psi) + b_i$ is the AoI at time slot t. Moreover, β_ψ is exactly the optimum value of (5)*, i.e., $\beta_\psi = h_{\psi,opt}$.*

Proof sketch. The problem (5) is an infinite-horizon average-cost semi-Markov decision process (SMDP) [22, Chapter 5.6]. Define $\tau = S_{i+1} - A_i$ as the waiting time for sending the $(i+1)$-th packet after the ACK feedback of the i-th packet is delivered to the transmitter. Given $\Delta(A_i) = \delta$ and $c_i = c$, the Bellman optimality equation of the average-cost SMDP (5) is

$$V_\psi(\delta, c) =$$

$$\inf_{\tau \in \{0,1,\ldots\}} \mathbb{E}\left[\sum_{k=0}^{\tau+T_{i+1}-1} \left(h(\delta + k) - h_{\psi,opt} \right) \middle| c_i = c \right]$$

$$+ \mathbb{E}\left[\sum_{j=0}^{F_{i+1}-1} \left(h(\psi(c_i) + T_{i+1} + j) - h_{\psi,opt} \right) \middle| c_i = c \right]$$

$$+ \mathbb{E}[V_\psi(\psi(c_i) + T_{i+1} + F_{i+1}, c_{i+1})|c_i = c], \quad (9)$$

for all $\delta \in \mathbb{Z}^+$, $\psi(\cdot) \in \Psi$, and $c \in \mathcal{C}$, where $V_\psi(\cdot)$ is the relative value function of the average-cost SMDP (5). Theorem 1 is proven by directly solving the Bellman optimality equation (9). The details are provided in Appendix A. $\square$

Theorem 1 signifies that the optimal solution to (5) is an index-based threshold policy, where the index function depends on both the AoI and the channel state in the previous epoch. The scheduler submits $(i+1)$-th packet according to (4) at the first time slot for which the following conditions are satisfied: (i) the ACK of the previous transmission is received, i.e., $t \geq A_i(\beta_\psi)$, and (ii) the index function exceeds the threshold. The threshold β_ψ is exactly the optimum value of (5) and can be found by solving equation (8) using the low-complexity algorithms, e.g., bisection search, provided in [4, Algorithms 1-3]. The index function in (6) is obtained by directly solving the Bellman optimality equation (9).

Recent studies [2], [9] have revealed a connection between the AoI optimization problems, such as (5), and the *restart-in-state* formulation of the Gittins index in [23, Ch. 2.6.4]. The optimal scheduling policy in Theorem 1 has been obtained by employing that connection.

Now, we present an optimal solution to problem (3).

Theorem 2. *If $|h(\delta)| < M$ for all $\delta = 1, 2, \ldots$, then the policy tuple (f^*, g^*) is an optimal solution to (3) such that*

i.) $f^* = (b_1^*, b_2^*, \ldots)$, *where*

$$b_{i+1}^* = \psi^*(c_i), \ i = 0, 1, 2, \ldots, \quad (10)$$

$$\psi^* = arg\min_{\psi \in \Psi} h_{\psi,opt}, \quad (11)$$

and $h_{\psi,opt}$ is the optimal objective value to (5).

ii.) $g^* = (S_1^*, S_2^*, \ldots)$, *where*

$$S_{i+1}^* = \min\{t \geq S_i^* + T_i + F_i : \gamma(\Delta(t), c_i) \geq h_{opt}\}, \quad (12)$$

and

$$h_{opt} = \min_{\psi \in \Psi} h_{\psi,opt}. \quad (13)$$

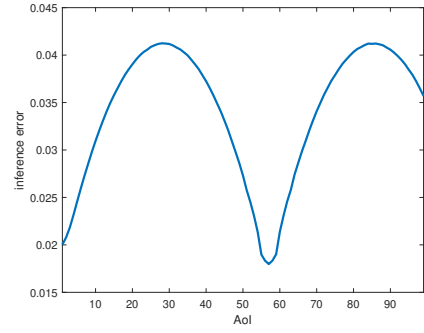*Proof.* See Appendix B. $\square$

Fig. 2. The incurred inference error when the current sample of the target Y_t is predicted using a past sample $X_{t-\delta}$ with AoI δ ranging from 1 to 99.

Theorem 2 points out that the optimal buffer position sequence to problem (3) is $f^* = f_{\psi^*} = (\psi^*(c_0), \psi^*(c_1), \ldots)$, where $\psi^* \in \Psi$ is the optimal mapping function. This implies that the optimal packet in the buffer to submit, at any submission time slot, is independent of the current AoI and only depends on the channel state in the previous epoch. Since $\psi^* \in \Psi$, the optimal packet submission time sequence is obtained by Theorem 1.

In the special case that the inference error $h(\delta)$ is a non-decreasing function of the AoI δ, the index function $\gamma(\delta, c)$ reduces to $\mathbb{E}[h(\delta+T_{i+1})|c_i = c]$. Moreover, the optimal buffer position sequence is $f^* = (0, 0, \ldots)$.

IV. SIMULATION RESULTS

In this section, we demonstrate the performance of our optimal scheduling policy in Theorem 2. To obtain an inference error function $h(\delta)$, we use a discrete-time autoregressive model of the order p (AR(p)):

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \ldots + a_p X_{t-p} + W_t, \quad (14)$$

where the noise $W_t \in \mathbb{R}$ is zero-mean Gaussian with variance σ_W^2 and $X_t \in \mathbb{R}$. Let $Y_t = X_t + N_t$ be the target variable, where the noise N_t is zero-mean Gaussian with variance σ_N^2. Fig. 2 indicates the average inference error values incurred when the target Y_t is predicted using $X_{t-\delta}$, where the value of AoI δ ranges from 1 to 99. We set $\sigma_W^2 = 0.01$, $\sigma_N^2 = 0.001$ and construct an AR(63) with coefficients $a_{51} = a_{52} = a_{62} = a_{63} = 0.015$, $a_{53} = a_{61} = 0.03$, $a_{54} = a_{60} = 0.065$, $a_{55} = a_{59} = 0.145$, $a_{56} = a_{58} = 0.15$, $a_{57} = 0.16$. The remaining coefficients are zero. We consider a quadratic loss function $L_2(y, \hat{y}) = \|y - \hat{y}\|_2^2$. Because $X_{t-\delta}$ and Y_t are jointly Gaussian, and the loss function is quadratic, the optimal inference error performance is achieved by a linear MMSE estimator. Hence, a linear regression algorithm is adopted in our simulations. However, our study can be readily applied to other neural network-based predictors.

We compare the performance of three scheduling policies: (i) the optimal policy in Theorem 2, (ii) the single-source scheduling policy in [2], and (iii) generate-at-will + zero wait policy: (f, g) is such that $f = (0, 0, \ldots)$ and $g = (A_1, A_2, \ldots)$.

The simulation parameters related to the channel statistics have been set as follows: (i) the number of channel states
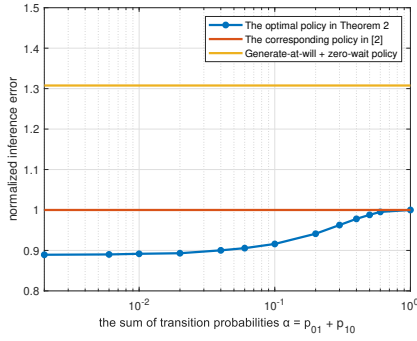
Fig. 3. The time-average inference error achieved by the three different scheduling policies. (normalized by the inference error value achieved by the single-source scheduling policy in [2]).

$N = 2$, (ii) the random variable Q_1 has the PMF $p_{Q_1}(3) = 0.45$, $p_{Q_1}(4) = 0.25$, and $p_{Q_1}(5) = p_{Q_1}(6) = 0.15$, (iii) the random variable Q_2 has the PMF $p_{Q_2}(18) = p_{Q_2}(19) = 0.15$, $p_{Q_2}(20) = 0.4$, and $p_{Q_2}(21) = 0.3$, (iv) the random variable R_1 has the PMF $p_{R_1}(2) = 1$, and (v) the random variable R_2 has the PMF $p_{R_2}(6) = 1$. The memory of the transmission delay is adjusted by varying the parameter $\alpha = p_{01} + p_{10}$. We set $p_{01} = p_{10}$, which ensures that the fraction of epochs with convenient delay condition is always $\frac{1}{2}$. The scheduling policy in [2] assumes both channel states are equally likely at each epoch independent of the history for any value of α.

Fig. 3 presents the normalized time-average inference error values achieved by the three scheduling policies mentioned above. The normalization is based on the value achieved by the single-source scheduling policy in [2]. The results highlight (i) submitting the freshest packet consistently with zero waiting time yields 31% to 47% higher inference error than the other two policies and (ii) the optimal scheduling policy in Theorem 2 outperforms the corresponding policy in [2] with a performance gain rising to 11% with the increasing memory of the delay. The channel distribution is almost IID for α values close to 1. Therefore, the performance gain for that region is quite small. The memory of the delay rises as α decays.

V. CONCLUSION

In this paper, we considered a remote inference problem where an intelligent application predicts the current value of a target signal using the data packets sent from a distant location over a two-way channel with varying delay conditions. We developed an optimal index-based threshold policy that minimizes the expected inference error per time slot in steady-state. Finally, we demonstrated, with the simulation results, the performance gain that can be achieved in the remote inference problems by considering the memory of the delay.

REFERENCES

[1] C. E. Shannon and W. Weaver, "The mathematical theory of communication," 1949.
[2] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?" *ACM MobiHoc*, 2022.
[3] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the Wiener process for remote estimation over a channel with random delay," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1118–1135, 2020.
[4] T. Z. Ornee and Y. Sun, "Sampling and remote estimation for the Ornstein-Uhlenbeck process through queues: Age of information and beyond," *IEEE/ACM Trans. Netw.*, vol. 29, no. 5, pp. 1962–1975, 2021.
[5] M. K. C. Shisher, H. Qin, L. Yang, F. Yan, and Y. Sun, "The age of correlated features in supervised learning based forecasting," in *IEEE INFOCOM Age of Information Workshop*, 2021.
[6] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *IEEE INFOCOM*, 2012, pp. 2731–2735.
[7] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "AoI-penalty minimization for networked control systems with packet loss," in *IEEE INFOCOM Age of Information Workshop*, 2019, pp. 189–196.
[8] T. Soleymani, J. S. Baras, and K. H. Johansson, "Stochastic control with stale information–part I: Fully observable systems," in *IEEE CDC*, 2019, pp. 4178–4182.
[9] M. K. C. Shisher, B. Ji, I.-H. Hou, and Y. Sun, "Learning and communications co-design for remote inference systems: Feature length selection and transmission scheduling," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 524–538, 2023.
[10] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *2015 Information Theory and Applications Workshop (ITA)*. IEEE, 2015, pp. 25–31.
[11] 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, 2017.
[12] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, 2018.
[13] I. Kadota, A. Sinha, and E. Modiano, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Trans. Netw.*, no. 27(4), pp. 1359–1372, 2019.
[14] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Nonlinear age functions," *J. Commun. Netw.*, no. 21(3), pp. 204–219, 2019.
[15] V. Tripathi and E. Modiano, "A Whittle index approach to minimizing functions of age of information," in *IEEE Allerton*, 2019, pp. 1160–1167.
[16] J. Sun, Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Closed-form Whittle's index-enabled random access for timely status update," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1538–1551, 2019.
[17] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.
[18] A. M. Bedewy, Y. Sun, R. Singh, and N. B. Shroff, "Optimizing information freshness using low-power status updates via sleep-wake scheduling," in *ACM MobiHoc*, 2020, pp. 51–60.
[19] C.-H. Tsai and C.-C. Wang, "Unifying aoi minimization and remote estimation—optimal sensor/controller coordination with random two-way delay," *IEEE/ACM Trans. Netw.*, vol. 30, no. 1, pp. 229–242, 2021.
[20] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimal sampling for data freshness: Unreliable transmissions with random two-way delay," *IEEE/ACM Trans. Netw.*, vol. 31, no. 1, pp. 408–420, 2022.
[21] M. K. C. Shisher, Y. Sun, and I.-H. Hou, "Timely communications for remote inference," *in preparation*, 2023.
[22] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.
[23] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.
[24] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

APPENDIX A: PROOF OF THEOREM 1

The scheduling problem (5) is an infinite-horizon average-cost semi-Markov decision process (SMDP) [22, Chapter 5.6]. According to [22, Chapter 5.6], we describe the components of this problem in detail below:

- **Decision Time:** Each i-th ACK feedback time $A_i = S_i + T_i + F_i$ is a decision time of the problem (5), where S_i is the submission time of the i-th packet, which takes $T_i \geq 1$ time slots to be delivered to the receiver, and

the i-th ACK feedback takes $F_i \geq 1$ time slots to be delivered to the transmitter.

- **State:** At time A_i, the state of the system is represented by $(\Delta(A_i), c_i)$, where $\Delta(A_i)$ is the AoI at time slot A_i and c_i is the channel state in i-th epoch.

- **Action:** Let $\tau_i = S_{i+1} - A_i$ represent the waiting time for sending the $(i + 1)$-th packet after the i-th ACK feedback is received. As we consider $S_0 = 0$, $S_i = \sum_{j=0}^{i-1}(T_j + F_j + \tau_j)$ and $A_i = S_i + T_i + F_i$ for each $i \in \{0, 1, \dots\}$. Given $(T_0, T_1, \dots)$ and $(F_0, F_1, \dots)$, the sequence $(S_1, S_2, \dots)$ is uniquely determined by $(\tau_0, \tau_1, \dots)$. Hence, one can also use $(\tau_0, \tau_1, \dots)$ to represent a sequence of actions instead of $(S_1, S_2, \dots)$. The scheduler decides the waiting time τ_i at time slot A_i. Then, the packet in the buffer position $b_{i+1} = \psi(c_i)$ is submitted to the channel at time slot S_{i+1}.

- **State Transitions:** The AoI $\Delta(t)$ evolves as $\Delta(t) = T_i + b_i$ if $t = D_i$ and $\Delta(t) = \Delta(t - 1) + 1$ otherwise for $i \in \{0, 1, \dots\}$. The state of the channel in i-th epoch, c_i, evolves following a finite-state ergodic Markov chain with the transition probabilities p_{ij}, where $i, j \in \mathcal{C}$, and p_{ij} is the probability of a transition from state i to state j. The Markov chain makes a single transition at the time slots A_i and none otherwise.

- **Expected Transition Time:** The expected time difference between the decision times A_i and A_{i+1} given $c_i = c$ is $\mathbb{E}[A_{i+1} - A_i | c_i = c] = \mathbb{E}[\tau_i + T_{i+1} + F_{i+1} | c_i = c]$.

- **Expected Transition Cost:** Given $c_i = c$, the expected cumulative cost incurred during the interval between the decision times A_i and A_{i+1} can be calculated as

$$\mathbb{E}\left[\sum_{t=A_i}^{A_{i+1}-1} h(\Delta(t)) \bigg| c_i = c\right] =$$
$$\mathbb{E}\left[\sum_{k=0}^{\tau_i + T_{i+1} - 1} h(\Delta(A_i) + k) \bigg| c_i = c\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} h(\psi(c_i) + T_{i+1} + k) \bigg| c_i = c\right] \quad (15)$$

The average-cost SMDP (5) can be solved by using dynamic programming [22], [24]. Given $\Delta(A_i) = \delta$ and $c_i = c$, the Bellman optimality equation of the average cost problem is

$$V_\psi(\delta, c) =$$
$$\inf_{\tau_i \in \{0,1,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\tau_i + T_{i+1} - 1} \left(h(\delta + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\psi(c_i) + T_{i+1} + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]$$
$$+ \mathbb{E}[V_\psi(\psi(c_i) + T_{i+1} + F_{i+1}, c_{i+1}) | c_i = c], \quad (16)$$

where $V_\phi(\delta, c)$ is the relative value function associated with the average cost problem (5).

Let $\bar{\tau} = \tau(\delta, c, h_{\psi,opt})$ be an optimal solution to (16). If $\Delta(A_i) = \delta$ and $c_i = c$, then the optimal waiting time τ_i for sending the $(i + 1)$-th packet is $\bar{\tau}$. Because the terms

$$\mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\psi(c_i) + T_{i+1} + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]$$

and

$$\mathbb{E}[V_\psi(\psi(c_i) + T_{i+1} + F_{i+1}, c_{i+1}) | c_i = c]$$

do not depend on the waiting time $\bar{\tau}$, the optimization problem can be reformulated as

$$\inf_{\bar{\tau} \in \{0,1,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\bar{\tau} + T_{i+1} - 1} \left(h(\delta + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]. \quad (17)$$

From (17), we can show that $\bar{\tau} = 0$ if

$$\inf_{\bar{\tau} \in \{1,2,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\bar{\tau} + T_{i+1} - 1} \left(h(\delta + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]$$
$$\geq \mathbb{E}\left[\sum_{k=0}^{T_{i+1}-1} \left(h(\delta + k) - h_{\psi,opt}\right) \bigg| c_i = c\right]. \quad (18)$$

After some rearrangement, the inequality (18) can also be expressed as

$$\inf_{\bar{\tau} \in \{1,2,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\bar{\tau}-1} \left(h(\delta + T_{i+1} + k) - h_{\psi,opt}\right) \bigg| c_i = c\right] \geq 0. \quad (19)$$

The inequality (19) holds if and only if

$$\inf_{\bar{\tau} \in \{1,2,\dots\}} \frac{1}{\bar{\tau}} \mathbb{E}\left[\sum_{k=0}^{\bar{\tau}-1} h(\delta + T_{i+1} + k) \bigg| c_i = c\right] \geq h_{\psi,opt}. \quad (20)$$

The left-hand side of (20) is equal to the index function $\gamma(\delta, c)$ given by (6). Similarly, $\bar{\tau} = 1$, if $\bar{\tau} \neq 0$ and

$$\inf_{\bar{\tau} \in \{2,3,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\bar{\tau}-1} \left(h(\delta + T_{i+1} + k) - h_{\psi,opt}\right) \bigg| c_i = c\right] \geq 0. \quad (21)$$

The inequality (21) can be rewritten as

$$\gamma(\delta + 1, c) \geq h_{\psi,opt}. \quad (22)$$

By repeating the same steps as (21)-(22), we can obtain $\bar{\tau} = k$ is optimal, if $\bar{\tau} \neq 0, 1, \dots, k - 1$ and

$$\gamma(\delta + k, c) \geq h_{\psi,opt}. \quad (23)$$

Hence, the optimal waiting time $\tau_i = \bar{\tau}$ is determined by

$$\bar{\tau} = \tau(\delta, c, h_{\psi,opt}) = \min_{k \in \mathbb{Z}}\{k \geq 0 : \gamma(\delta + k, c) \geq h_{\psi,opt}\}. \quad (24)$$

Next, we need to compute the optimal objective value $h_{\psi,opt}$. Let φ_i denote the random variable $\psi(c_{i-1}) + T_i$.

Because $\Delta(A_i(\beta_\psi)) = \varphi_i + F_i$ and $A_{i+1}(\beta_\psi) = A_i(\beta_\psi) + \tau(\varphi_i + F_i, c_i, \beta_\psi) + T_{i+1} + F_{i+1}$, we can write

$$\mathbb{E}\left[\sum_{t=A_i(\beta_\psi)}^{A_{i+1}(\beta_\psi)-1} h(\Delta(t))\right] - \beta_\psi \mathbb{E}[A_{i+1}(\beta_\psi) - A_i(\beta_\psi)] =$$
$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,\beta_\psi)+T_{i+1}-1} h(\varphi_i + F_i + k)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} h(\varphi_{i+1} + k)\right]$$
$$- \beta_\psi \mathbb{E}[\tau(\varphi_i + F_i, c_i, \beta_\psi) + T_{i+1} + F_{i+1}]. \quad (25)$$

Equation (25) implies that $h_{\psi,opt} = \beta_\psi$ is the root of (8), if the following holds:

$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,h_{\psi,opt})+T_{i+1}-1} h(\varphi_i + F_i + k)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} h(\varphi_{i+1} + k)\right]$$
$$- h_{\psi,opt}\mathbb{E}[\tau(\varphi_i + F_i, c_i, h_{\psi,opt}) + T_{i+1} + F_{i+1}] = 0. \quad (26)$$

From (16), we get

$$V_\psi(\varphi_i + F_i, c_i) =$$
$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,h_{\psi,opt})+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - h_{\psi,opt}\right)\bigg| c_i\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - h_{\psi,opt}\right)\bigg| c_i\right]$$
$$+ \mathbb{E}[V_\psi(\varphi_{i+1} + F_{i+1}, c_{i+1})|c_i]. \quad (27)$$

By the law of iterated expectations, if we take expectation over c_i on both sides of (27), we obtain

$$\mathbb{E}[V_\psi(\varphi_i + F_i, c_i)] =$$
$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,h_{\psi,opt})+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - h_{\psi,opt}\right)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - h_{\psi,opt}\right)\right]$$
$$+ \mathbb{E}[V_\psi(\varphi_{i+1} + F_{i+1}, c_{i+1})]. \quad (28)$$

Since c_i evolves following a finite-state ergodic Markov chain with a unique stationary distribution, $\mathbb{E}[V_\psi(\varphi_i + F_i, c_i)] = \mathbb{E}[V_\psi(\varphi_{i+1} + F_{i+1}, c_{i+1})]$. Hence, we get

$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,h_{\psi,opt})+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - h_{\psi,opt}\right)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - h_{\psi,opt}\right)\right] = 0, \quad (29)$$

which yields equation (26).

Finally, we need to show that the root of equation (8) is unique. The right-hand side of (25) can be expressed as

$$\mathbb{E}\left[\sum_{k=0}^{\tau(\varphi_i+F_i,c_i,\beta_\psi)+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - \beta_\psi\right)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - \beta_\psi\right)\right] =$$
$$\inf_{\tau\in\{0,1,\dots\}} \mathbb{E}\left[\sum_{k=0}^{\tau+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - \beta_\psi\right)\right]$$
$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - \beta_\psi\right)\right]. \quad (30)$$

This expression is concave, continuous, and strictly decreasing in β_ψ because the first term is the pointwise infimum of the linear decreasing functions of β_ψ and the second term is a linear decreasing function of β_ψ. Therefore, since

$$\lim_{\beta_\psi\to-\infty} \mathbb{E}\left[\sum_{k=0}^{\tau+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - \beta_\psi\right)\right] = \infty$$

and

$$\lim_{\beta_\psi\to\infty} \mathbb{E}\left[\sum_{k=0}^{\tau+T_{i+1}-1} \left(h(\varphi_i + F_i + k) - \beta_\psi\right)\right] = -\infty$$

for any $\tau \geq 0$, and

$$\lim_{\beta_\psi\to-\infty} \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - \beta_\psi\right)\right] = \infty$$

and

$$\lim_{\beta_\psi\to\infty} \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1} \left(h(\varphi_{i+1} + k) - \beta_\psi\right)\right] = -\infty$$

equation (8) has a unique root. This completes the proof.

APPENDIX B: PROOF OF THEOREM 2

The scheduling problem (3) is an infinite-horizon average-cost semi-Markov decision process (SMDP) [22, Chapter 5.6]. The components [22, Chapter 5.6] of this problem are the same as the SMDP in the proof of Theorem 1 with one exception. In this case, at any decision time A_i, the buffer position b_{i+1} is not determined by an arbitrary mapping function $\psi(\cdot) \in \Psi$, and the scheduler can freely select the buffer position b_{i+1} from the set $\mathcal{B}$.

The average-cost SMDP (3) can be solved by using dynamic programming [22], [24]. Given $\Delta(A_i) = \delta$ and $c_i = c$, the Bellman optimality equation of the average cost problem is

$$V(\delta, c) =$$

$$\inf_{\substack{\tau_i \in \{0,1,2,\dots\} \\ b_{i+1} \in \{0,1,\dots,B-1\}}} \mathbb{E}\left[\sum_{k=0}^{\tau_i+T_{i+1}-1}\left(h(\delta+k)-h_{opt}\right)\Big|c_i = c\right]$$

$$+ \mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1}\left(h(b_{i+1}+T_{i+1}+k)-h_{opt}\right)\Big|c_i = c\right]$$

$$+ \mathbb{E}[V(b_{i+1}+T_{i+1}+F_{i+1}, c_{i+1})|c_i = c], \tag{31}$$

where $V(\delta, c)$ is the relative value function associated with the average cost problem (3). Because the term

$$\mathbb{E}\left[\sum_{k=0}^{\tau_i+T_{i+1}-1}\left(h(\delta+k)-h_{opt}\right)\Big|c_i = c\right]$$

does not depend on the buffer position b_{i+1}, the optimal buffer position b_{i+1}^* when $\Delta(A_i) = \delta$ and $c_i = c$ is

$$b_{i+1}^* = arg\min_{b_{i+1} \in \{0,1,\dots,B-1\}}$$

$$\mathbb{E}\left[\sum_{k=0}^{F_{i+1}-1}\left(h(b_{i+1}+T_{i+1}+k)-h_{opt}\right)\Big|c_i = c\right]$$

$$+ \mathbb{E}[V(b_{i+1}+T_{i+1}+F_{i+1}, c_{i+1})|c_i = c]. \tag{32}$$

Equation (32) indicates that the optimal buffer position b_{i+1}^* is independent of the AoI δ at the decision time slot A_i and is a function of the channel state in the previous epoch $c_i = c$. Therefore, we can construct a mapping $\psi^* : \mathcal{C} \mapsto \mathcal{B}$ such that

$$\psi^*(c) =$$

$$arg\min_{b \in \mathcal{B}} E\left[\sum_{k=0}^{F_{i+1}-1}\left(h(b+T_{i+1}+k)-h_{opt}\right)\Big|c_i = c\right]$$

$$+ \mathbb{E}[V(b+T_{i+1}+F_{i+1}, c_{i+1})|c_i = c] \tag{33}$$

for all $c \in \mathcal{C}$. Therefore, in the optimal scenario, this problem is a special case of problem (5), where the buffer position sequence f_ψ is set to $f_{\psi^*} = (\psi^*(c_0), \psi^*(c_1), \dots)$. Since $\psi^* \in \Psi$, the optimal packet submission time sequence $g^* = (S_1^*, S_2^*, \dots)$ is obtained by Theorem 1. This completes the proof.