

# Age-Optimal Power Allocation in Industrial IoT: A Risk-Sensitive Federated Learning Approach

Yung-Lin Hsu\*, Chen-Feng Liu†, Sumudu Samarakoon†, Hung-Yu Wei\*, and Mehdi Bennis†

\*Graduate Institute of Communication Engineering, National Taiwan University, Taiwan

†Centre for Wireless Communications, University of Oulu, Finland

E-mail: {d04942010, hywei}@ntu.edu.tw, {chen-feng.liu, sumudu.samarakoon, mehdi.bennis}@oulu.fi

**Abstract**—This work studies a real-time environment monitoring scenario in the industrial Internet of things, where wireless sensors proactively collect environmental data and transmit it to the controller. We adopt the notion of *risk-sensitivity* in financial mathematics as the objective to jointly minimize the mean, variance, and other higher-order statistics of the network energy consumption subject to the constraints on the age of information (AoI) threshold violation probability and the AoI exceedances over a pre-defined threshold. We characterize the extreme AoI staleness using results in *extreme value theory* and propose a distributed power allocation approach by weaving in together principles of Lyapunov optimization and *federated learning* (FL). Simulation results demonstrate that the proposed FL-based distributed solution is on par with the centralized baseline while consuming 28.50% less system energy and outperforms the other baselines.

**Index Terms**—5G and beyond, industrial IoT, smart factory, federated learning (FL), age of information (AoI), extreme value theory (EVT).

## I. INTRODUCTION

Environment monitoring and control in smart factory scenarios are mission-critical applications in 5G and beyond, where sensors, meters, and monitors generate and upload data to a central controller with real-time ultra-low latency. In particular, for real-time monitoring and control, the elapsed time since the data was generated is a key factor for the control performance. Such time duration is referred to as the *age of information* (AoI). If the AoI of the controller’s available data grows unexpectedly, the outcome of the real-time environmental monitoring and control will be poorly degraded [1]. The impact of AoI-aware resource allocation has been investigated in various communication systems [1]–[5]. In [1], a mean-field game approach was proposed in a dense IoT monitoring system. The work [2] considered a multi-sensor industrial IoT (IIoT) scenario with finite blocklength transmission in which the controller instructs all devices to sample and upload environment data based on its AoI records. The objective therein was to minimize the sensors’ power consumption. Aiming at minimizing the average AoI and average peak AoI using a Markov decision process. The work [3] investigated the tradeoff between AoI and energy cost and proposed an action policy for devices. In [4], the authors assumed that the devices could not upload data while wirelessly harvesting energy from the base station. As a result, the AoI exponentially increases during the energy harvesting period. Finally, the authors in

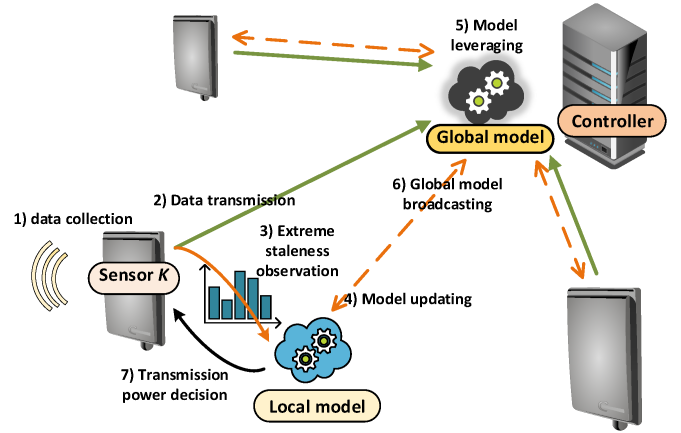


Figure 1: Sensor-Controller system architecture.

[5] considered a remote monitoring problem trading off the expected AoI and the AoI threshold violation probability.

While reducing the AoI in a centralized manner, the proposed resource allocation approaches in [1]–[5] incurred not only supernumerary power for statistical information exchanging but also tremendous signaling overheads, which are not negligible in ultra-low latency real-time monitoring, e.g., industrial automation. Under the ultra-low latency constraints, delegating the transmission decisions to the sensors provides a more realistic avenue, especially when the data-sampling time is uncertain. Therefore, while accounting for the AoI threshold violation and the threshold-exceeding events, our work aims at proposing a decentralized and proactive power allocation scheme by minimizing the sensors’ entropic risk measure energy consumption [6]. We further leverage *extreme value theory* at the sensor to characterize the AoI exceedances over a threshold using the locally observed historical data. However, the accuracy of characterization is limited by the amount of the sensor’s data. To improve the accuracy under limited data availability, we resort to *federated learning* (FL), a collaborative and distributed model training framework [7], [8], and propose an FL-based distributed power allocation algorithm along with Lyapunov optimization. Numerical results show that the proposed approach is on par with the centralized baseline while consuming less system energy for data transmission and model training.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

As shown in Fig. 1, we consider an industrial IoT network with a set  $\mathcal{K}$  of  $K$  intelligent sensors that monitor distinct and independent (but of the same type) environments and transmit the sampled data to a controller. We assume that the sensors' sampling operations are event-triggered by the environmental changes without any available statistical information. After the sensor samples the status data, the new data is immediately transmitted to the controller if the previous data's uploading procedure has been completed. Otherwise, the (new) data is queued in the buffer. Such a queueing time is called the procrastinated time in this paper. Let us index the sequentially sampled data of each sensor by  $i \in \mathbb{Z}^+$  and denote the procrastinated time of the  $i$ th data of sensor  $k \in \mathcal{K}$  as  $\eta_k(i) \geq 0$ . Then to send the  $i$ th data with size  $N$ , sensor  $k$  allocates transmission power  $p_k(i)$  over its dedicated bandwidth  $B$  with the transmission time

$$t_k(i) = \frac{N}{B \log_2 \left( 1 + \frac{h_k(i)p_k(i)}{N_0 B} \right)}, \quad (1)$$

in which  $h_k(i)$  is the channel gain, including path loss and channel fading, between sensor  $k$  and the controller, and  $N_0$  is the power spectral density of the additive white Gaussian noise. Further, assume that the controller issues control commands to the actuator right after receiving the status data. Hence, we focus on the AoI of the time instant at which the status data is fully received at the controller. In this regard, we express the concerned AoI of the  $i$ th data as  $a_k(i) = \eta_k(i) + t_k(i)$  which incorporates the elapsed procrastinated and transmission time. The data sampling and transmitting instances are schematically illustrated in Fig. 2. Therein,  $\tau_k(i)$  denotes the sampling time instant of the  $i$ th data, and the procrastinated time is shown as the overlapping period. From Fig. 2, we can straightforwardly find  $\eta_k(i) = [\tau_k(i-1) + a_k(i-1) - \tau_k(i)]^+$  with  $[x]^+ = \max\{x, 0\}$  and rewrite the AoI in a recursive manner as

$$a_k(i) = [\tau_k(i-1) + a_k(i-1) - \tau_k(i)]^+ + t_k(i). \quad (2)$$

Since the information ages in consecutive transmissions are entangled as per (2), we should proactively account for the impacts of the allocated power on future AoI.

Note that the factory environment varies continuously, in which the stale information can degrade the control system performance. As a remedy, we impose a set of constraints on the AoI in terms of the weighted expectation, AoI outage probability, and AoI exceedance. We firstly introduce a staleness function from [9] as  $f_k(i) = \frac{[a_k(i)]^{(1-\beta)}}{1-\beta}$  for the AoI staleness with a predetermined value  $\beta \leq 0$  and then consider a long-term time-averaged constraint for every sensor  $k$ , i.e.,

$$\lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \mathbb{E}[f_k(i)] \leq f_0, \forall k \in \mathcal{K}, \quad (3)$$

where  $f_0$  is a pre-defined threshold. Additionally, we impose a probabilistic constraint on the AoI for each sensor  $k$  as

$$\lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \Pr\{f_k(i) > f_0\} \leq \epsilon, \forall k \in \mathcal{K}, \quad (4)$$

in which  $\epsilon \ll 1$  is the tolerable threshold violation prob-

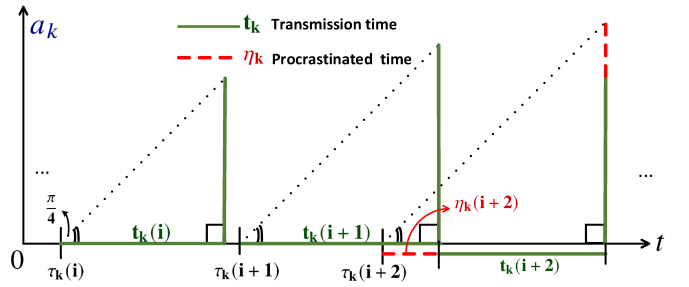


Figure 2: The AoI instances with data sampling and transmission. Since only  $\eta_k(i+2) > 0$ , we have  $a_k(i) = t_k(i)$ ,  $a_k(i+1) = t_k(i+1)$ , and  $a_k(i+2) = \eta_k(i+2) + t_k(i+2)$ .

ability. Although a very low occurrence probability of the AoI exceedance is ensured in (4), the uploaded data with an extremely large age can hinder the control performance. To mitigate this effect, a constraint on the AoI exceedances  $q_k(i) = f_k(i) - f_0 > 0$  is imposed as follows:

$$\lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \mathbb{E}[q_k(i)] \leq e_0, \forall k \in \mathcal{K}. \quad (5)$$

Here, we define the set of AoI exceedance of sensor  $k$  as  $\mathcal{Q}_k \triangleq \{q_k(i) | \mathbb{1}_{\{q_k(i) > 0\}}, \forall i\}$ , and  $e_0$  is a pre-defined threshold. Finally, taking into account the sensors' limited-energy, we aim at minimizing not only the sensors' average energy consumption but also the variance for data uploading to reduce superabundant energy between each sensor. To this end, we invoke the entropic risk measure  $\frac{1}{\rho} \ln(\mathbb{E}_X[\exp(\rho X)])$ , which incorporates the mean, variance, and higher-order statistics of  $X$  [6], in our global objective, and formulate the studied optimization problem as follows:

$$\text{minimize}_{p_k(i)} \frac{1}{\rho} \ln \left( \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \frac{1}{K} \sum_{k \in \mathcal{K}} \exp(\rho E_k(i)) \right) \quad (6a)$$

$$\text{subject to } 0 \leq p_k(i) \leq p_{max}, \forall i \in \mathbb{Z}^+, k \in \mathcal{K}, \quad (6b)$$

$$(3), (4), \text{ and } (5),$$

where  $E_k(i) = p_k(i)t_k(i)$  denotes sensor  $k$ 's energy consumption for transmitting its  $i$ th data,  $\rho > 0$  reflects the weights of the variance and higher-order statistics in risk minimization, and  $p_{max}$  is the sensor's data transmission power budget. Moreover, in the objective (6a), we take the average over time and sensors instead of expectation in the entropic risk measure. In the next section, we will utilize Lyapunov optimization techniques [10] to solve problem (6) since the objective (6a) and constraints (3), (4), and (5) are long-term time-averaged functions.

### III. FL-BASED DISTRIBUTED POWER ALLOCATION

#### A. Lyapunov Optimization Framework

In order to ensure the time-averaged constraints by Lyapunov optimization, we first introduce the virtual queues

$$\Gamma_k(i+1) = [\Gamma_k(i) + f_k(i) - f_0]^+, \quad (7)$$

$$\Upsilon_k(i+1) = [\Upsilon_k(i) + [q_k(i) - e_0] \mathbb{1}_{\{q_k(i) > 0\}}]^+, \quad (8)$$

for constraints (3) and (5), respectively. Additionally, by applying  $\Pr\{f_k(i) > f_0\} = \mathbb{E}[\mathbb{1}_{\{q_k(i) > 0\}}]$  and scaling both sides of

(4) as  $\lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I f_k(i) \mathbb{E}[\mathbb{1}_{\{q_k(i) > 0\}}] \leq \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I f_k(i) \epsilon$ , the virtual queue

$$\Lambda_k(i+1) = [\Lambda_k(i) + (\mathbb{1}_{\{q_k(i) > 0\}} - \epsilon) f_k(i)]^+ \quad (9)$$

is considered for the constraint (4). Due to the fact that minimizing  $\lim_{I \rightarrow \infty} \frac{1}{IK} \sum_{i=1}^I \sum_{k \in \mathcal{K}} \exp(\rho E_k(i))$  is equivalent to minimizing (6a), we consider the conditional Lyapunov drift-plus-penalty

$$\frac{1}{K} \sum_{k \in \mathcal{K}} \mathbb{E} \left[ \frac{1}{2} \Phi_k(i+1) \Phi_k(i+1)^T - \frac{1}{2} \Phi_k(i) \Phi_k(i)^T + V \exp(\rho E_k(i)) \right], \quad (10)$$

where  $\Phi_k(i) = [\Gamma_k(i), \Upsilon_k(i), \Lambda_k(i), \forall k \in \mathcal{K}]$  is the combined queue vector for notational simplicity, and  $(\cdot)^T$  represents the transpose of a vector. Subsequently, by applying  $([Q+y]^+)^2 \leq Q^2 + 2Qy + y^2$ , we can derive the inequality

$$(10) \leq \frac{1}{K} \sum_{k \in \mathcal{K}} \mathbb{E} [\Delta_0 + F_k(i) + V \exp(\rho E_k(i))], \quad (11)$$

in which  $F_k(i) = \theta_k^1(i) [f_k(i)]^2 + \theta_k^2(i) f_k(i)$  with  $\theta_k^1(i) = \frac{1}{2}(1+\epsilon^2) + (1-\epsilon) \mathbb{1}_{\{q_k(i) > 0\}}$  and  $\theta_k^2(i) = \Gamma_k(i) - f_0 - \epsilon \Lambda_k(i) + [\Lambda_k(i) + \Upsilon_k(i) - f_0 - e_0] \mathbb{1}_{\{q_k(i) > 0\}}$ . Additionally,  $\Delta_0 = \frac{1}{2}[f_0^2 - \Gamma_k(i) + [(f_0 + e_0)^2 - \Lambda_k(i)(f_0 + e_0)] \mathbb{1}_{\{q_k(i) > 0\}}$  is a constant. Note that the solution of (6) can be obtained by minimizing the derived upper bound on the conditional Lyapunov drift-plus-penalty function [10], i.e., (11), in each transmission  $i$  by optimizing over the transmit power  $p_k(i)$ . Furthermore,  $V \geq 0$  is the tradeoff parameter between the lengths of the virtual queues and the optimality of the energy consumption in (6).

### B. Distributed Power Allocation with Federated Learning for AoI Exceedances

Obtaining the optimal solution of minimizing the upper bound (11) requires the global information about  $\Phi_k(i), \forall k$ . However, centrally allocating power (i.e., the centralized model-training (CENT) scheme in Section IV) consumes tremendous power overheads on queue state information exchanges. To improve power efficiency, we propose a decentralized FL-Based scheme, wherein each sensor  $k \in \mathcal{K}$  locally allocates its transmit power for each transmission  $i \in \mathbb{Z}^+$  by minimizing the individual upper bound of (11),

$$\underset{p_k(i)}{\text{minimize}} \quad V \exp(\rho E_k(i)) + F_k(i) \quad \text{subject to (6b)}. \quad (12)$$

In (12), we can straightforwardly prove the convexity of  $F_k(i)$ , but  $\exp(\rho E_k(i))$  is non-convex with respect to  $p_k(i)$ . In order to tractably solve the non-convex problem (12), we adopt the notion of the convex-concave procedure (CCP) [11] by which we iteratively convexify the non-convex part  $\exp(\rho E_k(i))$  by the first-order Taylor series expansion with respect to a reference point  $\hat{p}_k$  as

$\exp(\rho E_k(\hat{p}_k)) \left[ t_k(\hat{p}_k) \left( 1 - \frac{\hat{p}_k h_k t_k(\hat{p}_k)}{N(\hat{p}_k h_k + N_0 B) \ln 2} \right) (p_k - \hat{p}_k) + 1 \right]$  and solve the convexified problem. Specifically, given the reference point  $\hat{p}_k^r$  in the  $r$ th iteration, we focus on

$$\underset{\hat{p}_k^{r+1}}{\text{minimize}} \quad V J(\hat{p}_k^r) \hat{p}_k^{r+1} + F_k(\hat{p}_k^{r+1}) \quad (13a)$$

$$\text{subject to} \quad 0 \leq \hat{p}_k^{r+1} \leq p_{max}, \quad (13b)$$

in which  $J(\hat{p}_k^r) = \exp(\rho E_k(\hat{p}_k^r)) t_k(\hat{p}_k^r) \left[ 1 - \frac{\hat{p}_k^r h_k t_k(\hat{p}_k^r)}{N(\hat{p}_k^r h_k + N_0 B) \ln 2} \right]$ . The optimal solution to (13) is subsequently used as the reference point  $\hat{p}_k^{r+1}$  in the next iteration. The initial reference point  $\hat{p}_k^0$  is randomly selected from  $(0, p_{max}]$ . When  $|\hat{p}_k^r - \hat{p}_k^{r-1}| \rightarrow 0$  after a number of iterations, sensor  $k$  uses the transmit power  $p_k^*(i) = \hat{p}_k^r$  to upload the  $i$ th data and then updates (2) and all virtual queues (7), (8), and (9).

However, since the AoI exceedance happens rarely, the deviation of virtual queue (8) between each sensor becomes severe with local statistics only. To address this issue, we invoke results from extreme value theory and the principles of FL. Let us rewrite the virtual queue  $\Upsilon_k(i)$  in (8) as

$$\begin{aligned} \Upsilon_k(i+1) &= \sum_{j=1}^i [q_k(j) - e_0]^+ \mathbb{1}_{\{q_k(j) > 0\}} \\ &\geq \left[ \sum_{j=1}^i [q_k(j) - e_0] \mathbb{1}_{\{q_k(j) > 0\}} \right]^+ \\ &\stackrel{(a)}{=} \left[ \frac{\sum_{j=1}^i q_k(j) \mathbb{1}_{\{q_k(j) > 0\}}}{\sum_{j=1}^i \mathbb{1}_{\{q_k(j) > 0\}}} - e_0 \right]^+ \sum_{j=1}^i \mathbb{1}_{\{q_k(j) > 0\}} \end{aligned}$$

in which the first term in (a) represents the empirical average, which may have large variance due to limited historical data of the excess AoI. Nevertheless, if the mean of the AoI exceedance is available, we can estimate the steady-state average length of the virtual queue (8).

**Theorem 1** (Pickands–Balkema–de Haan theorem [12]). *Given a random variable  $A$  with the cumulative distribution function  $F_A(a)$  and a threshold  $a_0$ , as  $a_0 \rightarrow F_A^{-1}(1)$ , the excess value  $Q = A - a_0 > 0$  can be approximately characterized by a generalized Pareto distribution (GPD) with the scale  $\sigma > 0$  and shape  $\xi \in \mathbb{R}$  parameters. The mean of the GPD is  $\frac{\sigma}{1-\xi}$ .*

Leveraging the results in Theorem 1, we characterize the statistics of  $q_k(j)$  as a GPD whose parameters  $\sigma$  and  $\xi$  (i.e., the mean) can be estimated using maximum likelihood estimation. Given a sufficient amount of historical data, the GPD model (i.e., scale and shape parameters) of the excess AoI can be trained. However, owing to the sparsity of the excess AoI data at the sensor, it is time-consuming for each sensor to train the GPD model independently. To overcome this hurdle, we utilize the FL framework in which all sensors periodically update their locally-trained model to the controller. Then the controller aggregates the updated local models and feeds back the aggregated model to the sensors. Our FL-based model training is detailed as follows.

Assume that the local-model updating time interval is  $M$ , and each interval is indexed by  $m \in \mathbb{Z}^+$ . In every updating time interval, each sensor trains its model locally. In order to have sufficient independent data for local training, we set  $W$  observation time windows within which the sensor selects the largest excess AoI as a training sample. The observation time windows are indexed by  $w \in \mathbb{Z}^+$ , and the window size is  $O$  with  $M/O = W \in \mathbb{Z}^+$ , which should be sufficiently large to minimize the correlation between the selected data while being sufficiently small to prevent filtering out the data

overmuch. Moreover, the selected extreme data at sensor  $k$  in the  $w$ th time window of the  $m$ th time interval is denoted by  $\hat{q}_k^{m,w} = \max_{\tau_k(i) \in \mathcal{T}^{m,w}} \{q_k(i) \mathbb{1}_{\{q_k(i) > 0\}}\}$ , where  $\mathcal{T}^{m,w} \in [M(m-1) + O(w-1), M(m-1) + Ow]$ . The selected data set within the  $m$ th time interval is denoted by  $\mathcal{Q}_k^m = \{\hat{q}_k^{m,w}\}_{w=1}^W$ . After collecting the samples  $\mathcal{Q}_k^m$ , we train the sensor  $k$ 's GPD model  $\theta_k^m = \{\sigma_k^m, \xi_k^m\}$  via *tilted empirical risk minimization* (TERM) [13], i.e.,

$$\begin{aligned} & \underset{\theta_k^m}{\text{minimize}} \quad \tilde{\mathcal{L}}(\theta_k^m | t, \mathcal{Q}_k^m) \\ & \equiv \underset{\theta_k^m}{\text{minimize}} \quad \frac{1}{t} \ln \left( \frac{1}{|\mathcal{Q}_k^m|} \sum_{\mathcal{Q}_k^m} G(\theta_k^m | \mathcal{Q}_k^m)^{-t} \right). \end{aligned} \quad (14)$$

Here,  $G(\sigma, \xi | q) = \frac{1}{\sigma} (1 + \frac{\xi q}{\sigma})^{-(\frac{1}{\sigma} + 1)}$  is the GPD's likelihood function while  $t$  is the tilted factor used to leverage the effects of the outliers in model training. In this regard, by setting  $t < 0$  in (14), we can account for the outliers and other extreme events. Subsequently, based on the global model  $\theta^{m-1} = \{\sigma^{m-1}, \xi^{m-1}\}$  received in the  $(m-1)$ th interval, each sensor  $k$  updates the local model parameters as per

$$\theta_k^m = \theta^{m-1} - \delta_\theta \nabla_{\theta_k^m} \tilde{\mathcal{L}}(\theta_k^m | t, \mathcal{Q}_k^m), \quad (15)$$

with the initial value  $\theta^0$  and sends  $\theta_k^m$  to the controller. In (15),  $\delta_\theta$  is the step size. The gradient of  $\tilde{\mathcal{L}}(\theta_k^m | t, \mathcal{Q}_k^m)$  with respect to  $\theta_k^m$  is given by

$$\begin{aligned} \nabla_{\theta_k^m} \tilde{\mathcal{L}}(\theta_k^m | t, \mathcal{Q}_k^m) &= \\ & \frac{-\sum_{\mathcal{Q}_k^m} \nabla_{\theta_k^m} G(\theta_k^m | \mathcal{Q}_k^m) \times G(\theta_k^m | \mathcal{Q}_k^m)^{-t-1}}{\sum_{\mathcal{Q}_k^m} G(\theta_k^m | \mathcal{Q}_k^m)^{-t}}, \end{aligned} \quad (16)$$

and

$$\begin{aligned} \frac{\partial G(\theta | \mathcal{Q})}{\partial \sigma} &= \frac{Q - \sigma}{\sigma^3} \left( 1 + \frac{\xi Q}{\sigma} \right)^{-2 - \frac{1}{\xi}}, \\ \frac{\partial G(\theta | \mathcal{Q})}{\partial \xi} &= \frac{(1 + \frac{\xi Q}{\sigma})^{-1 - \frac{1}{\xi}}}{\sigma \xi} \left( \frac{-(\xi + 1)Q}{\sigma + \xi Q} + \frac{\ln(1 + \frac{\xi Q}{\sigma})}{\xi} \right). \end{aligned}$$

Here, the notations  $m$  and  $k$  are neglected for simplicity. The controller then calculates the global model of the  $m$ th updating time interval as

$$\theta^m = \frac{\sum_{k=1}^K (|\mathcal{Q}_k^m| \theta_k^m)}{\sum_{k=1}^K |\mathcal{Q}_k^m|}. \quad (17)$$

Finally, after receiving the global feedback model  $\theta^m$ , each sensor  $k$  replaces the virtual queue value with

$$\Upsilon_k(\hat{i} + 1) = \left[ \frac{\sigma^m}{1 - \xi^m} - e_0 \right]^+ \mathbb{E}_{\mathcal{K}} \left[ \sum_{j=1}^{\hat{i}} \mathbb{1}_{\{q_k(j) > 0\}} \right], \quad (18)$$

in which  $\hat{i} = \arg \min_{\forall i} \{\tau_k(i) - Mm \geq 0\}$ , and proceeds with the next local-model training  $\theta_k^{m+1}$ . The proposed FL-based distributed power allocation is outlined in Algorithm 1. Therein, each sensor  $k \in \mathcal{K}$  iteratively optimizes the transmission power  $p_k(i_k)$  for each local-sampled data  $i_k$  based on the time instance  $\tau_k(i_k)$  and the FL updated virtual queue states  $A(i_k), \Gamma_k(i_k), \Lambda_k(i_k), \Upsilon_k(i_k)$ .

#### IV. NUMERICAL RESULTS

We simulate a factory environment with  $K = 50$  sensors with 50 Hz data-sampling frequency in Poisson. The considered path loss model is  $32.45 + 31.9 \log 20 + 20 \log 3.5$

#### Algorithm 1 FL-Based Distributed Power Allocation Mechanism

---

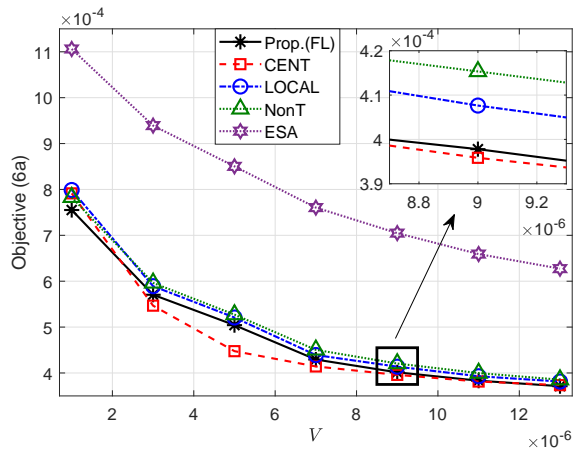
**Input:**  $\tau_k(i_k) \forall k \in \mathcal{K}$ , transmission parameters in (1),  $\beta, \theta^0, \delta_\theta, M, O$ .  
**Output:**  $p_k^*(i_k), A(i_k), \Gamma_k(i_k), \Lambda_k(i_k), \Upsilon_k(i_k) \forall i_k, \forall k \in \mathcal{K}$ .  
**Initialize:** create  $i_k$  for data indicator,  $i_k = 1, \forall k$ ,  
 $a_k(0) = 0, \eta_k(1) = 0, \tau_k(i_k | i_k = 0, 1) = 0, \forall k$ ,  
 $\{A_k(1), \Gamma_k(1), \Lambda_k(1), \Upsilon_k(1)\} = 0$ .

- 1: **for**  $m = 1, 2, \dots$  **do** // for global model updating interval
- 2:   **for**  $k = 1, \dots, |\mathcal{K}|$  **do**
- 3:      $w = 1, \mathcal{Q}_k^m = \{\}$ , // check each observation window
- 4:     **if**  $\tau_k(i_k) \leq Mm$  &&  $w \leq W$  **then**
- 5:       **while**  $\tau_k(i_k) \leq M(m-1) + Ow$  **do**
- 6:          update  $\eta_k(i_k), \theta_k^1(i_k), \theta_k^2(i_k), \theta_k^3(i_k)$ ,
- 7:          form  $u_k(i_k)$  by (11) and solve  $p_k^*(i_k)$  by CCP.
- 8:          count  $t_k^*(i_k)$  by (1),  $E_k^*(i_k), a_k(i_k), f_k(i_k), q_k(i_k)$ ,
- 9:          update state  $A(i_k), \Gamma_k(i_k), \Lambda_k(i_k), \Upsilon_k(i_k)$ ,
- 10:           $i_k = i_k + 1$ ,
- 11:       **end while**
- 12:       Count  $\hat{q}_k^{m,w}$ , update  $\mathcal{Q}_k^m = \{\mathcal{Q}_k^m, \hat{q}_k^{m,w}\}$ ,
- 13:        $w = w + 1$ ,
- 14:     **else**  $\mathcal{Q}_k^m = \{\}$ ,
- 15:     **end if**
- 16:     train  $\theta_k^m$  with (15), (16) by  $\mathcal{Q}_k^m$  and update them to Controller,
- 17:   **end for**
- 18:   Controller updates  $\theta^m$  with  $\theta_k^m$  by (17), and broadcasts to IDs,
- 19:    $\forall k, \Upsilon_k(i_k)$  is replaced with  $\hat{\theta}^m$  by (18).
- 20: **end for**

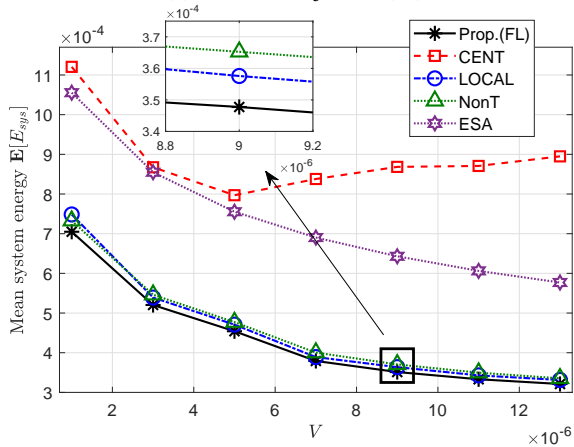
---

(dB) given 3.5 GHz carrier frequency and a 20-meter sensor-controller distance [14]. The wireless channel experiences Rayleigh fading with unit variance. The remaining simulation parameters are  $N = 3000$  bytes,  $B = 180$  kHz,  $N_0 = -174$  dBm,  $p_{max} = 23$  dBm,  $\beta = -2$ ,  $f_0 = 5 \times 10^{-4}$ ,  $e_0 = 10^{-4}$ ,  $\epsilon = 2 \times 10^{-3}$ ,  $M = 30$  ms,  $O = 10$  ms,  $\theta^0 = [0.0002, 0.02]$ ,  $\delta_\theta = [10^{-9}, 10^{-3}]$ , and  $t = -0.1$ . For performance comparison, we consider four baselines: *i*) **Centralized model-training** (CENT) scheme which trains the extreme staleness GPD model at the central controller with all sensors' AoI exceedance data; *ii*) **Local model-training** (LOCAL) scheme in which the sensors only train the GPD model individually without any information exchange; *iii*) **Non-model-training** (NonT) scheme which directly solves problem (12) without renewing the virtual queue  $\Upsilon_k(i)$  in (8) via the training result of GPD model; *iv*) **Excess staleness-agnostic** (ESA) scheme which does not take extreme staleness into consideration, i.e., neglecting constraints (4) and (5) in problem (6). In addition to the performance of the objective (6a), we further investigate the expected system energy consumption  $\mathbb{E}[E_{sys}] = \mathbb{E}[E_k] + \mathbb{E}[E_{comp}^{train}] + \mathbb{E}[E_{tx}^{train}]$ , including the sensor's monitored data-updating energy  $E_k$ , the computation energy in GPD-model training  $E_{comp}^{train} = 10^{-27} f_{cpu}^2 N_{tr} l_{req}$ , and the energy consumption in model-uploading  $E_{tx}^{train} = p_{max} N_{tr} / (B \log_2 (1 + \frac{h_k p_{max}}{N_0 B}))$ . Here,  $f_{cpu} = 2 \times 10^{11}$  cycle/s and  $f_{cpu} = 10^9$  cycle/s are the controller's and sensor's computation capabilities for model training [6].  $N_{tr} = 30$  bytes is the single-data size in GPD-model training, and  $l_{req} = 87.8$  cycle/bit is the required computation frequency.

The impact of the tradeoff parameter  $V$  in the Lyapunov optimization on energy consumption is shown in Fig. 3. We first examine the performance of the objective (6a) in Fig. 3a as a function of  $V$ . Note that the objective is a decreasing function of  $V$  since the importance of energy reduction increases with



(a) Value of objective (6a)

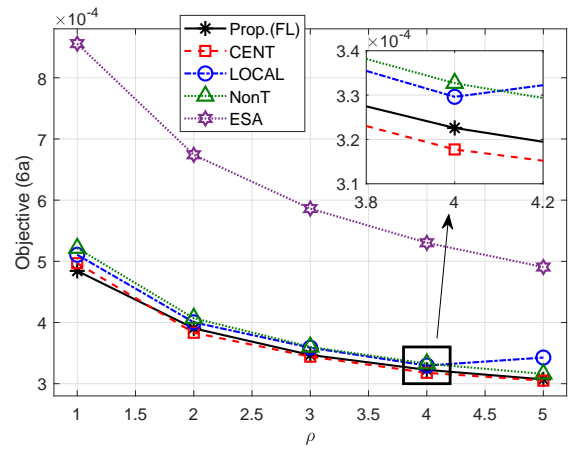


(b) System energy  $\mathbb{E}[E_{sys}]$

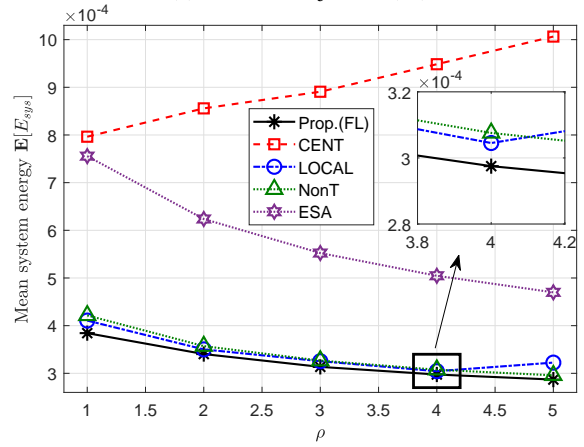
Figure 3: Energy consumption versus  $V$  with  $\rho = 2$ .

increasing  $V$ . Although CENT outperforms the other schemes across various  $V$  values due to the GPD model training, the performance of our proposed scheme is close to CENT with 0.48% – 11.26% increments in the objective. Compared with LOCAL and NonT, the proposed FL scheme decreases the objective cost up to 5.45% and 5.56%, respectively, showing the benefit of FL model training. In addition, LOCAL outperforms NonT, manifesting the advantage of model training even without any information exchange. Being agnostic to extreme staleness, ESA may incur a high instantaneous AoI, resulting in larger energy consumption. In this regard, ESA has a higher objective cost (up to 44.75% increase) than the cost of the proposed scheme.

In Fig. 3b, we further verify the advantages of our proposed approach in terms of the expected system energy consumption. In CENT, all sensors have to deliver every observed extreme staleness data to the controller, consuming high energy for information exchange. Note that as  $V$  increases, all schemes tend to save more power in environmental data transmission based on the Lyapunov optimization framework but increases the occurrence chance of extreme staleness. In this situation, the CENT sensors have to consume more energy to upload more model-training data to the controller. Therefore, the



(a) Value of objective (6a)



(b) System energy  $\mathbb{E}[E_{sys}]$

Figure 4: Energy consumption versus  $\rho$  with  $V = 10^{-5}$ .

expected system energy of CENT grows gradually with  $V$  in Fig. 3b. In the proposed FL-based approach, since the sensors train the GPD models locally and only upload the model parameters to the controller, the energy consumption is significantly reduced. In this regard, our approach can save up to 64.11% in system energy compared to CENT. In contrast with NonT and LOCAL, our approach spends extra energy on FL model training and information exchange. Nevertheless, our energy-saving benefit with respect of  $\mathbb{E}[E_k]$  compensates this expenditure. The proposed scheme can save up to 5.82% and 6.21% in system energy then LOCAL and NonT. The objective performance and expected system energy consumption by varying  $\rho$  are shown in Fig. 4. As per (12), all the schemes put more focus on energy deduction as  $\rho$  grows.

The capability of extreme staleness control is manifested in Fig. 5. Given a specific amount of expected status-updating energy  $\mathbb{E}[E_k]$ , the proposed approach, CENT, and LOCAL benefited from the GPD-model training showcase the lowest excess staleness values. In contrast, ESA, which is agnostic to the extreme AoI, has much higher excess staleness. However, if we further consider the energy consumption in model training, our proposed approach (compared with CENT and LOCAL) consumes less energy while achieving the same



## V. CONCLUSIONS

We studied an industrial IoT network in which the sensors autonomously and proactively allocate transmit power for uploading monitored status data. To avoid stale data delivery hindering the monitoring performance, we have formulated an entropic risk-minimizing problem for energy consumption subject to data staleness constraints. By applying extreme value theory results, we trained an GPD model to obviate extreme staleness regimes and have further proposed a distributed FL-based scheme to improve the power efficiency versus its centralized counterpart. Numerical results have shown that the proposed FL scheme is on par with the centralized model-training scheme while consuming significantly less system energy. Future work will extend the current framework towards a system that can jointly optimize the data transmission power and sampling rate rather than obey a default number.

## ACKNOWLEDGMENTS

This work is funded by the Ministry of Science and Technology (MOST) of Taiwan under the Graduate Students Study Abroad Program grant 109-2917-I-002-007 and supported by the CHIST-ERA LeadingEdge and CONNECT projects.

## REFERENCES

- [1] B. Zhou and W. Saad, "Age of information in ultra-dense iot systems: Performance and mean-field game analysis," *CoRR*, vol. abs/2006.15756, pp. 1–30, Jun. 2020.
- [2] C.-F. Liu and M. Bennis, "Taming the tail of maximal information age in wireless industrial networks," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2442–2446, Dec. 2019.
- [3] E. Gindullina, L. Badia, and D. Gündüz, "Average age-of-information with a backup information source," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2019, pp. 1–6.
- [4] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and K. B. Letaief, "Online transmission policy in wireless powered networks with urgency-aware age of information," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf.*, Jun. 2019, pp. 1096–1101.
- [5] A. Elgabli, H. Khan, M. Krouka, and M. Bennis, "Reinforcement learning based scheduling algorithm for optimizing age of information in ultra reliable low latency networks," in *Proc. IEEE Symp. Comput. Commun.*, Jun. 2019, pp. 1–6.
- [6] S. Batewela, C.-F. Liu, M. Bennis, H. A. Suraweera, and C. S. Hong, "Risk-sensitive task fetching and offloading for vehicular edge computing," *IEEE Commun. Lett.*, vol. 24, no. 3, pp. 617–621, Mar. 2020.
- [7] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, "Wireless network intelligence at the edge," *Proc. IEEE*, vol. 107, no. 11, pp. 2204–2239, Nov. 2019.
- [8] J. Park, S. Samarakoon, A. Elgabli, J. Kim, M. Bennis, S.-L. Kim, and M. Debbah, "Communication-efficient and distributed learning over wireless networks: Principles and applications," *CoRR*, vol. abs/2008.02608, pp. 1–20, Aug. 2020.
- [9] A. B. Atkinson, "On the measurement of inequality," *Journal of Economic Theory*, vol. 2, no. 3, pp. 244–263, Sep. 1970.
- [10] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. San Rafael, CA, USA: Morgan and Claypool, Jun. 2010.
- [11] T. Lipp and S. Boyd, "Variations and extension of the convex–concave procedure," *Optimization and Engineering*, vol. 17, 06 2016.
- [12] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. London, U.K.: Springer, 2001.
- [13] T. Li, A. Beirami, M. Sanjabi, and V. Smith, "Tilted empirical risk minimization," *CoRR*, vol. abs/2007.01162, pp. 1–40, Jul. 2020.
- [14] 3GPP, "Scenarios, frequencies and new field measurement results from two operational factory halls at 3.5 ghz for various antenna configurations," November 2018, 3GPP TSG RAN WG1 Meeting 95, Spokane, USA.

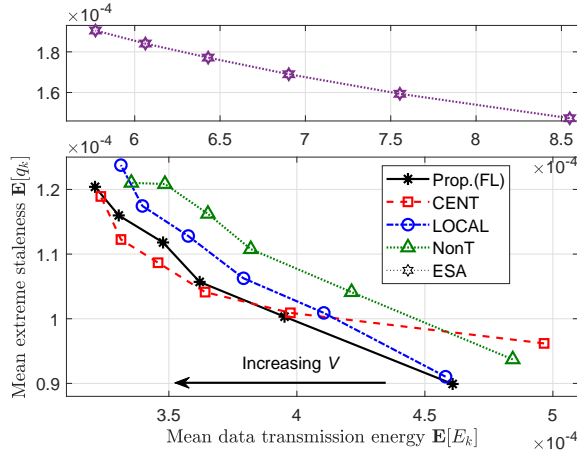


Figure 5: Extreme AoI versus average transmission energy with  $\rho = 2$ .

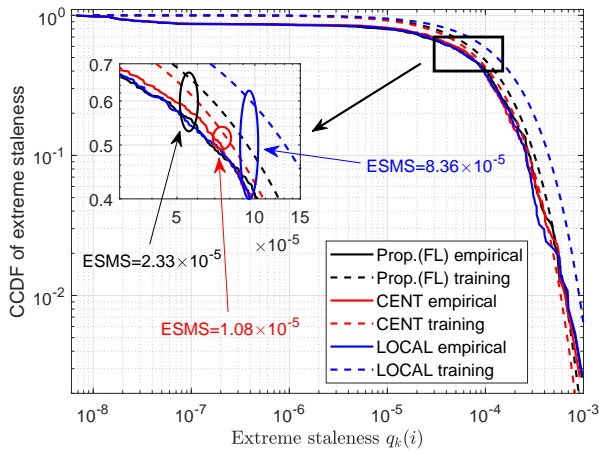


Figure 6: CCDF of extreme AoI with  $\rho = 2$  and  $V = 10^{-5}$ .

extreme staleness performance.

Finally, we discuss the performance of GPD-model training in terms of the complementary cumulative distribution function (CCDF) of extreme staleness in Fig. 6. As shown, the predicted extreme staleness mean  $\sigma/(1 - \xi)$  is always higher than the empirical one. The reason is that the predicted extreme staleness mean values in both schemes are leveraged to suppress the extreme staleness further. The more precise the estimation is, the more accurate the decision. The centralized approach, i.e., CENT, estimates the GPD model closer to the empirical one, saving more energy to control the extreme staleness. In this regard, the proposed FL scheme and CENT, respectively, yield  $2.33 \times 10^{-5}$  and  $1.08 \times 10^{-5}$  in terms of estimation-statistic mean surplus (ESMS) value between the trained model and empirical curve. On the other hand, the localized approach, i.e., LOCAL, has the highest ESMS  $8.36 \times 10^{-5}$  due to the lack of global estimation. Such results reflect on the least objective performance in Fig. 3a and 4a. However, the mean extreme staleness from the proposed scheme is 3.54% higher than CENT, underscoring that increasing transmission energy based on inaccurate predictions cannot effectively suppress extreme values.