Age of Information in Control Loop for Industrial IoT

†Thant Zin Oo, and ‡Choong Seon Hong
Department of Computer Science and Engineering, Kyung Hee University, Korea
{†tzoo, ‡cshong }@khu.ac.kr

Abstract

Recent developments in Internet of Things (IoT) enable industrial applications which require time-critical feedback of underlying physical processes in the control loop. We consider a real-time industrial IoT monitoring system for an industrial plant where the IoT sensor sample a physical process based on the control signal received from the IoT controller. This process incurs a sampling cost and an updating cost to send the status packet from the IoT sensor to the IoT controller. This joint status sampling and updating process is formulated as a minimization of Age of Information (AoI) at the IoT controller subject to energy budget constraint of each IoT sensor.

1. Introduction

Recently, the rapid growth of Internet of Thing (IoT) devices is making the delivery of up-to-date status information of the underlying physical processes increasingly more important [1]. Target examples monitoring applications include air quality and water supply in a city, network conditions of the electricity grid, and industrial IoT network for remote monitoring.

For such remote monitoring systems, the freshness of the information status, i.e., age of information (AoI), is the key performance metric [2]. In the industrial IoT, the control signals chosen must be decided based on the real-world status information of the underlying cyber-physical systems. The conventional delay metrics measure the time between generation and delivery of individual packet. AoI considers both the traditional delay and the generation time of each packet, i.e., the sensing interval. Hence, we can optimize AoI to design different control-feedback mechanism for the industrial IoT applications [3].

2. System Model

We consider Internet of Things (IoT) devices deployed in an industrial setting, specially, remote monitoring and automation scenario. As shown in Fig. 1, the industrial building contains multiple sources (areas of interest) which can be either a passive source such as offices and storage area for raw materials and finished products or an active source such as machines and overhead cranes.

Let $\mathcal{S} = \mathcal{S}_{\text{active}} \cup \mathcal{S}_{\text{passive}}$ denote the set of sources in the building. Several IoT devices are deployed in each area or on machines. Let $\mathcal{D}$ denotes the set all IoT devices which can further be subdivided into the sets of controllers and sensors, i.e., $\mathcal{D} = \mathcal{D}_s \cup \mathcal{D}_c$.

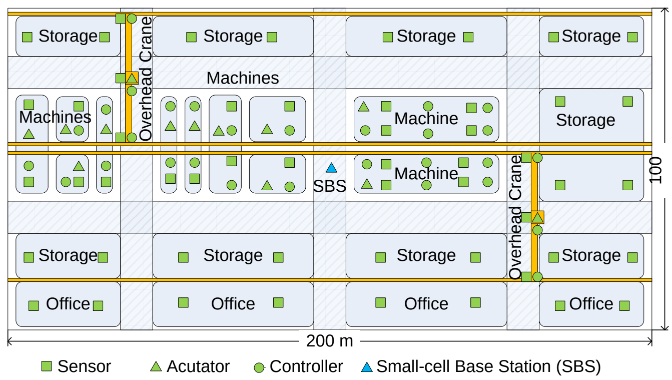

Fig. 1: Industrial IoT devices deployed in a 200 m × 100 m industrial building

A small-cell base station (SBS) is deployed to orchestrate the communications between IoT devices and act as a sink for monitoring and control. Let $\mathcal{K}$ denote the available spectrum resource blocks (RBs) for the IoT network.

The communications between the IoT devices is a transmit job $j \in \mathcal{J}_i$ which is can be defined as a tuple $\langle \hat{v}_j, \hat{R}_j, \hat{\varepsilon}_j, \hat{\delta}_j \rangle$, where $\hat{v}_j$ is the number of bits to be transmitted, $\hat{R}_j$ is the guaranteed bit rate (GBR) required, $\hat{\varepsilon}_j$ is the reliability in terms of block error rate (BLER), and $(t + \hat{\delta}_j)$ is the deadline for the transmit job j. The transmit jobs $\mathcal{J} = \cup_i \mathcal{J}_i$ can be classified into *periodic* and *aperiodic* or *on-demand*. *Periodic* communication is used for control and diagnosis purposes whereas *aperiodic* communication is used to update the changes in the sensing target, i.e. the source.

Fig. 2 shows the remote sensing and monitoring procedure which contains i) *detection*, ii) *reception*, and iii) *control-feedback*. As soon as the sensor detects change in the sensing target, it will transmit packet containing the sensed information to the controller. The controller, on receiving the sensed information, will perform the decision process and choose an action from the predefined candidate set.

The *age of information* (AoI) is defined as follow:

$$\Delta t_{AoI} = t_{\text{received}} - t_{\text{change}} = \delta_{\text{sensing}} + \delta_{\text{sen-con}} \quad (1)$$

where δ_{sensing} is the sensing latency, $\delta_{\text{sen-con}}$ is the transmission latency from sensor to controller. Note that, random variable δ_{sensing} is uniformly distributed and the distribution of random variable $\delta_{\text{sen-con}}$ is governed by the radio access technology of the wireless communication. We assume that the wireless communication employs 5G network. We define transmission latency or delay as the time duration from the arrival of traffic request (i.e., transmit job) until it is transmitted successfully on the RAN, i.e.,

$$\delta_{\text{sen-con}} = \delta_j = t_{j,\text{Tx}} - t_{j,\text{Arv}}, \quad (2)$$

where $t_{j,\text{Tx}}$ and $t_{j,\text{Arv}}$ denote the transmission completion time and job arrival time, respectively.

We index the controller with d_0. Hence, SINR of link $(d, d_0) \in \mathcal{D}_s \times \mathcal{D}_c$ on RB $k \in \mathcal{K}$ is :

$$\Gamma_{d,d_0,k} = \frac{z_{d,d_0,k}\ P_{d,d_0,k}\ h_{d,d_0,k}}{\sum_{d' \in \mathcal{D}_{\text{sen}} \setminus \{d\}} z_{d',d_0,k}\ P_{d',d_0,k}\ h_{d',d_0,k} + \Delta f_\mu\ N_0}, \quad (3)$$

where $z_{d,d_0,k} \in \{0,1\}$, $P_{d,d_0,k} \in [0, \hat{P}_i]$, and $h_{d,d_0,k}$ denote the control variable for resource allocation, the transmit power and the channel gain of link (d, d_0) on resource block (RB) k, respectively. Δf_μ and N_0 denote the sub-carrier bandwidth of RB k of the numerology μ and the thermal noise spectral power, respectively.

Accordingly, the achievable per resource data rate for link (d, d_0) on RB k is given by:

$$r_{d,d_0,k} = g_{\phi,c}(\Gamma_{d,d_0,k}), \quad \forall k \in \mathcal{K}, \quad (4)$$

where $g_{\phi,c}(\cdot)$ is a monotonic non-decreasing function of SINR that maps SINR values to the modulation according to transmission mode ϕ and channel quality indicator (CQI) [4]. Hence, the achievable data rate for link (d, d_0) is:

$$R_{d,d_0} = \sum_{k \in \mathcal{K}} r_{d,d_0,k}, \quad \forall (d, d_0) \in \mathcal{D}_s \times \mathcal{D}_c. \quad (5)$$

Next, we assume a block fading channel where the fading process is approximately constant in time domain for 12 sub-channels of an RB. Then, we have the symbol error rate (SER) and block error rate (BLER) for link (d, d_0) on resource k as:

$$\epsilon_{d,d_0,k} = \rho_{\phi,m}(\Gamma_{d,d_0,k}), \quad \forall k \in \mathcal{K}, \quad (6)$$

$$\varepsilon_{d,d_0,k} = 1 - (1 - \Gamma_{d,d_0,k})^{N_s}, \quad \forall (d, d_0) \in \mathcal{D}_s \times \mathcal{D}_c, \quad (7)$$

where $\rho_{\phi,m}(\cdot)$ is a strictly decreasing function of SINR that maps SINR values to the SER based on the transmission mode ϕ and $m \in \{2, 4, 16, 64, 256\}$ is the modulation order for PSK (phase shift keying) and m-QAM (quadrature amplitude modulation) [5].

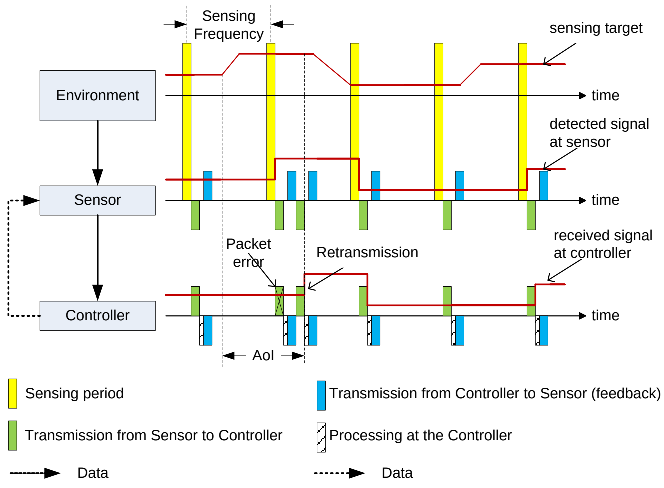

Fig. 2: Age of Information in industrial IoT

3. Problem Formulation

Objective function: We choose to minimize the AoI which consists of two parts of latency. δ_{sensing} is governed by the sensing period τ, and δ_j is governed by the communication system configuration $\boldsymbol{x} = (\mathbf{y}, \mathbf{z}, \mathbf{P}, \phi, c, m)$. Hence, we have individual objective function as follows:

$$u_{d,d_0}(\boldsymbol{\tau}, \boldsymbol{x}) = y_j(\delta_{\text{sensing}}(\tau_d) + \delta_j(\boldsymbol{x})), \quad \forall d \in \mathcal{D}_s \quad (8)$$

We formulate the minimization of AoI (Min-AoI) as an optimization problem given as:

$$\text{minimize:} \quad U(\boldsymbol{\tau}, \boldsymbol{x}) = \frac{1}{|\mathcal{D}_s|} \sum_{d \in \mathcal{D}_s} u_{d,d_0}(\boldsymbol{\tau}, \boldsymbol{x}) \quad (9)$$

$$\text{subject to:} \quad \check{\tau} \le \tau_d \le \hat{\tau}, \quad (10)$$

$$R_{d,d_0} \ge y_j \hat{R}_j, \ \forall j \in \mathcal{J}_i, \quad (11)$$

$$\varepsilon_{d,d_0} \le \hat{\varepsilon}_j, \ \forall j \in \mathcal{J}_i, \quad (12)$$

$$\delta_j \le \hat{\delta}_j, \ \forall j \in \mathcal{J}_i, \quad (13)$$

$$\sum_{k \in \mathcal{K}} P_{d,d_0,k} \le \hat{P}_d, \ \forall d \in \mathcal{D}_s, \quad (14)$$

$$\sum_{k \in \mathcal{K}} z_{d,d_0,k} \le |\mathcal{K}|, \ \forall d \in \mathcal{D}_s \quad (15)$$

$$\sum_{(d,d_0)} z_{d,d_0,k} \le 1, \ \forall k \in \mathcal{K}. \quad (16)$$

$$y_j, z_{d,d_0,k} \in \{0,1\}, \ P_{d,d_0,k} \in [0, \hat{P}_i]$$

In (9), we minimize the expected value of each AoI for all device. Constraint (10) set the lower and upper limits of the sensing period for each device d. Constraints (11)-(13) are three QoS constraints for user traffic demand, 1) guaranteed bit rate (GBR), $\hat{R}_j$, 2) reliability, $\hat{\varepsilon}_j$, and 3) latency, $\hat{\delta}_j$ [6]. (14) and (15) are budget constraints for power and resource for all the devices. (16) ensures orthogonal allocation of available resource to prevent interference. (Min-AoI) is a combinatorial optimization problem.

4. Markov Approximation Framework

We use Markov approximation framework [7] to solve (**Min-AoI**). Let $f = \{\boldsymbol{\tau}, \boldsymbol{x}\}$, $f \in \mathcal{F}$ be a network configuration where $\mathcal{F}$ is the set of all feasible configurations which satisfy the constraints (10)-(16). Further, for ease of presentation, let $U_f = U(\boldsymbol{\tau}, \boldsymbol{x})$. Then, by [7],

$$\max_{f \in \mathcal{F}} U_f \quad \Rightarrow \quad \begin{array}{ll} \max\limits_{\boldsymbol{\pi} \geq 0} & \sum_{f \in \mathcal{F}} \pi_f U_f \\ \text{s.t.} & \sum_{f \in \mathcal{F}} \pi_f = 1 \end{array} \tag{17}$$

We then apply log-sum-exponential approximation to (17) with the following differentiable function [7], [8]:

$$\max_{f \in \mathcal{F}} U_f \approx g_\beta(U_f) \triangleq \frac{1}{\beta} \log\left[\sum_{f \in \mathcal{F}} \exp(\beta U_f)\right] \tag{18}$$

The upper bound of the approximation gap is $\frac{1}{\beta}\log|\mathcal{F}|$ [8]. (18) is the same as the optimal value of the following problem:

$$\begin{array}{ll} \max\limits_{\boldsymbol{\pi} \geq 0} & \sum_{f \in \mathcal{F}} \pi_f U_f - \frac{1}{\beta} \sum_{f \in \mathcal{F}} \pi_f \log \pi_f \\ \text{s.t.} & \sum_{f \in \mathcal{F}} \pi_f = 1, \end{array} \tag{19}$$

where π_f is the proportion of time the configuration f is in use and β is a positive constant. By solving the Karush-Kuhn-Tucker (KKT) conditions [8] of (16),

$$\pi_f^*(U_f) = \frac{\exp(\beta U_f)}{\sum_{f' \in \mathcal{F}} \exp(\beta U_{f'})}, \ \forall f \in \mathcal{F}. \tag{20}$$

5. Numerical Analysis

We perform simulations with python to evaluate how Markov Approximation framework can solve (**Min-AoI**). We created a 5G network model with one deployed SBS. Next, we created a dynamic network scenario where IoT sensor devices are created using a Poisson Point Process. The environment changes happening at each IoT sensor also follows a Poisson Process. We observe the AoI as the key performance metric. We vary the two key parameters of the Markov Approximation framework, namely, the consolidation rate, β, and the learning rate, ω. First, we run the simulation using fixed $\beta = 1$, and $\omega = 0.5$. For comparison, we then run the simulations using dynamic values of β and ω. The results are plotted in Fig. 3 and Fig. 4, which show that Markov approximation framework achieves near-optimal solution.

6. Conclusion

In this paper, we formulate a joint status sampling and updating process as a minimization of Age of Information (AoI) at the IoT controller. The formulated problem is a combinatorial problem and NP-hard. We then apply a learning algorithm, i.e., Markov approximation to solve the formulated (**Min-AoI**) problem. The simulation results show that our formulation is viable and our proposed learning solution achieves near-optimal solution.

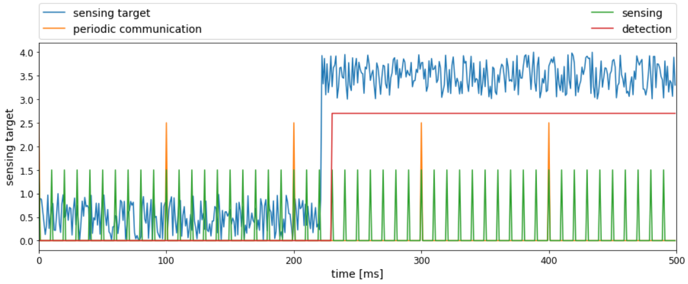

Fig. 3: Time line showing sensing frequency and detection latency.

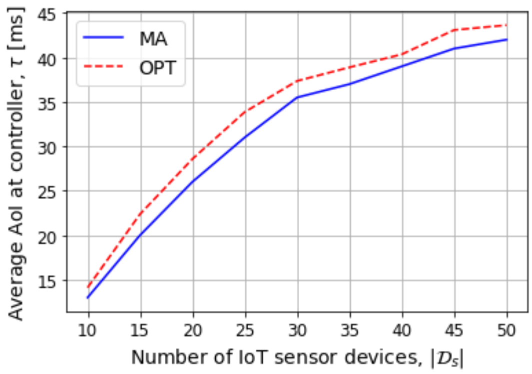

Fig. 4: Average AoI at the controller versus number of IoT sensor devices in the network.

Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2015-0-00557, Resilient/Fault-Tolerant Autonomic Networking Based on Physicality, Relationship and Service Semantic of IoT Devices) Dr. C. S. Hong is the corresponding author.

References

[1] L. Atzori, A. Iera and G. Morabito, "The Internet of Things: A survey," *Computer Networks,* vol. 54, no. 15, pp. 2787-2805, 2010.

[2] S. Kaul, R. Yates and M. Gruteser, "Real-time status: How often should one update?," in *Proceedings of IEEE INFOCOM*, Orlando, FL, 2012.

[3] B. Zhou and W. Saad, "Joint Status Sampling and Updating for Minimizing Age of Information in the Internet of Things," *arXiv:1807.04356.*

[4] 3GPP 36.213, *Technical Specification Group RAN; E-UTRA; Physical layer procedures (R-15) v15.3.0,* 3GPP, 2018-09.

[5] 3GPP 38.211, *Technical Specification Group RAN; NR; Physical channels and modulation (R-15) v15.3.0,* 3GPP, 2018-09.

[6] 3GPP 23.501, *Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2 (R-15) v15.3.0,* 3GPP, 2018-09.

[7] M. Chen, S. Liew, Z. Shao and C. Kai, "Markov Approximation for Combinatorial Network Optimization," *IEEE Trans. Inf. Theory,* vol. 59, no. 10, pp. 6301-6327, 2013.

[8] S. Boyd, and L. Vandenberhe, Convex Optimization, Cambridge University Press, 2004.