AA-DRL: AoI-Aware Deep Reinforcement Learning Approach for D2D-Assisted Industrial IoT

Hossam Farag*, Mohamed Ragab[†], and Čedomir Stefanović*

* Department of Electronic Systems, Aalborg University, Denmark

[†] Centre for Frontier AI Research, Agency for Science Technology and Research (A*STAR), Singapore

Email: hmf@es.aau.dk, mohamedr002@e.ntu.edu.sg, cs@es.aau.dk

Abstract-In real-time Industrial Internet of Things (IIoT), e.g., monitoring and control scenarios, the freshness of data is crucial to maintain the system functionality and stability. In this paper, we propose an AoI-Aware Deep Reinforcement Learning (AA-DRL) approach to minimize the Peak Age of Information (PAoI) in D2D-assisted IIoT networks. Particularly, we analyzed the success probability and the average PAoI via stochastic geometry, and formulate an optimization problem with the objective to find the optimal scheduling policy that minimizes PAoI. In order to solve the non-convex scheduling problem, we develop a Neural Network (NN) structure that exploits the Geographic Location Information (GLI) along with feedback stages to perform unsupervised learning over randomly deployed networks. Our motivation is based on the observation that in various transmission contexts, the wireless channel intensity is mainly influenced by distance-dependant path loss, which could be calculated using the GLI of each link. The performance of the AA-DRL method is evaluated via numerical results that demonstrate the effectiveness of our proposed method to improve the PAoI performance compared to a recent benchmark while maintains lower complexity against the conventional iterative optimization method.

Index Terms-Industrial IoT, neural networks, age of information

I. INTRODUCTION

The Internet of Things (IoT) technology evolves rapidly as a worldwide network of interconnected intelligent devices that are capable of sensing, communicating, and processing to support a variety of applications, such as industrial monitoring, health monitoring and vehicular networks [1]. Different from consumer IoT, Industrial IoT (IIoT) networks are characterized by strict communication requirements to maintain production efficiency and avoid safety-critical situation [2]. IIoT networks are evolving from the typical star network configuration to Device-to-Device (D2D) communications where sensoractuator pairs in propinquity communicate directly without evolving a central node (e.g., access point or a Base Station (BS) [3]. In outband D2D communication [3], the D2D pairs operate in full frequency reuse model where the they communication over the same frequency band in uncoordinated fashion. For IIoT scenarios that comprise enormous D2D pairs, this incurs significant interference when D2D links are activated at the same time. Such uncoordinated access would cause noticeable degradation in the network performance in terms of delay, throughput and most importantly the information freshness. Information freshness is crucial for a typical real-time control and monitoring scenario as it affects the derived intelligent and autonomous decisions and the system stability as well. For instance, in oil refineries, valve actuators should acquire timely monitoring of oil level to avoid oil tank spillage [4]. Information freshness is quantified by Age-of-Information (AoI) [5], a process-level metric which, from the receiver perspective, counts the time elapsed since the latest received information was generated. Efficient D2D links scheduling has been proved to effectively enhance various performance metrics [3]. However, this approach is not necessarily effective to improve and optimize AoI. In addition, scheduling problems for wireless networks with complex interference are usually non-convex and NP-hard [6]. The conventional method for link scheduling first involves estimating the interference of channels and then optimizing the schedule using these estimates [7]. Estimating Channel State Information (CSI) in densely deployed networks can be costly, and even achieving a near-optimal solution to the resulting optimization problem can be intricate. For instance, in network consisting of N D2D links, N^2 CSI are required in the pathloss matrix within each coherence block, which corresponds to a computational complexity of at least $O(N^2)$. To that end, machine learning and artificial intelligence techniques are employed to find optimal schedules in dense networks [8].

In this work, we develop an AoI-Aware Deep Reinforcement Learning (AA-DRL) approach to minimize Peak AoI (PAoI) in dense D2D-assisted IIoT networks. The learning approach is based on the collected Geographic Location Information (GLI) to select the optimal scheduling policy. Our motivation is based on the fact that in various transmission contexts, the wireless channel intensity is mainly influenced by distancedependant path loss. Moreover, the pattern of interference within a network largely depends on how transmitters and receivers are positioned relative to one another. We derive the successful transmission probability and the average PAoI considering a preemptive queuing policy and spatial coupled interference between the D2D pairs. Then, we formulate the PAoI optimization problem under stationary randomized policy. We solve the non-convex scheduling problem via a proposed Neural Network (NN) structure where the scheduling policy is mapped to the GLI while explicit CSI is not required. The NN is trained via an unsupervised training process by utilizing the GLI along with feedback stages to obtain the optimal scheduling policy. Our obtained numerical results show that the proposed AA-DRL approach achieves improved PAoI performance compared to a recent benchmark [9] while maintains lower complexity against the conventional iterative optimization method.

The remainder of the paper is organized as follows. Related work is presented in Section II. Section III describes the system model. In Section IV, we introduce our proposed DRL algorithm and the NN structure. Performance evaluations and conclusion are presented in Sections V and Section VI, respectively.

II. RELATED WORK

Several research works have been conducted with the goal of minimizing AoI in wireless networks. The authors in [10] studied the average AoI under different queue management schemes which concluded that packet replacement can promote reduced AoI compared to the conventional First Come First Served (FCFS) approach. The authors in [11]-[13] introduced centralized scheduling methods to minimize the average AoI. However, the centralized scheduling approach is inefficient in D2D-enabled IIoT networks as it would incur high overhead and extended delay. Efficient scheduling in D2D networks was tackled by different works that focus on the analysis and optimization of resource allocation [14], traffic density [15], or user fairness [16], while less attention was paid to the minimization of AoI. Moreover, other works apply stochastic geometry to model the spatial relationship of D2D devices, and adopting AoI-aware decentralized scheduling with the assumption of having full CSI [17]. However, it is shown that it is very challenging to obtain a global CSI in D2D networks [18]. A backlog-aware protocol was presented in [19] to minimize the average AoI subject to a delay constraint. The work in [9] presented a locally-adaptive slotted-ALOHA protocol where a link-wise access probability is dynamically selected to minimize the AoI considering a unit-size buffers and Last Come First Served (LCFS) queuing discipline. Both [19] and [9] require the exchange of queuestatus information between neighboring nodes to find the optimal access probability, which implies significant overhead and complexity in dense IIoT networks. Different from these works, our proposed AA-DRL has the potential to optimize the AoI performance of IIoT networks while maintaining low complexity with no explicit need for CSI.

III. SYSTEM MODEL

We consider an IIoT network with a three-layer hierarchy as depicted in Fig. 1. The first layer is the field network that consists of N D2D pairs transmitting status updates via a shared time-slotted channel, where each transmission attempt fits within the duration of one time slot. The second layer includes edge computing nodes with storage, computing, communication, and other resources that are used for computingintensive tasks. The third layer is the industrial cloud where historical data about the field and edge nodes is stored in the cloud for long-term data analysis. The considered model is common in most remote data acquisition and distributed

control applications, and is shown to have promising advantages in supporting efficient resource management [20]. In the field network layer, each transmitter generates status update following a Poisson process with average rate λ packets/slot. Such arrival model captures the scenario where such traffic can be triggered by the occurrence of some random incident [21]. The D2D transmitters have a single-occupancy backlog, i.e., the output buffer can accommodate only one packet. Moreover, we consider a preemptive queuing policy where an arriving packet can preempt the one currently in the service (if there is such). We consider a stationary randomized scheduling policy [22], in which the D2D transmitters are activated to transmit in each time slot with a given slot-access probability that is subject to optimization. Let $\pi = \{p_1, p_2, ..., p_N\}$ denotes the the randomized scheduling policy and $\boldsymbol{\Pi}$ is the class that represents all the possible policies where $\pi \in \Pi$. Particularly, for a given scheduling policy π , the D2D transmitter n_i , $i \in \{1, 2, ..., N\}$, is activated with a probability p_i across all the time slots.

A. Successful Update Probability

The collisions among simultaneously active D2D pairs are not necessarily destructive, due to the capture effect [23]. In that sense, a packet is decoded successfully at the end of a time slot when the Signal-to-Interference plus-Noise ratio (SINR) at the corresponding receiver exceeds the capture ratio β . The SINR at the AP in a time slot t given a set \mathcal{V} of interfering links can be written as [17]

$$\operatorname{SINR}_{i}(t) = \frac{P_{t}|h_{i,i}|^{2}d_{i,i}^{-\alpha}}{\sigma^{2} + \sum_{j \in \mathcal{V} \setminus \{i\}} P_{t}|h_{j,i}|^{2}d_{j,i}^{-\alpha}}, \qquad (1)$$

where P_t is the transmission power (assumed fixed for all D2D pairs), $h_{j,i}$ is the random variable that represents the Rayleigh fading of the channel between transmitter of link j and the receiver of link i with $h_{j,i} \sim \exp(1)$, $d_{j,i}$ is the distance between between transmitter of link j and the receiver of link i, σ^2 denotes the power of AWGN, and α is the path loss exponent. According to the considered channel model, for an arbitrary node n_i , the conditional successful decoding probability in timeslot t given a certain scheduling policy π can be obtained as

$$\begin{split} \phi_i(t|\pi) &= \Pr(\operatorname{SINR}_i > \beta | \pi) \\ &= \Pr\left(\frac{P_t | h_{i,i} |^2 d_i^{-\alpha}}{\sigma^2 + \sum_{j \in \mathcal{V} \setminus \{i\}} P_t | h_{j,i} |^2 d_j^{-\alpha} a_j(t)} > \beta | \pi\right) \\ &= \exp\left(-\frac{\beta \sigma^2}{P_t d_i^{-\alpha}}\right) \prod_{j \in \mathcal{V} \setminus \{i\}} \left(\frac{1}{1 + a_j(t) \frac{\beta d_j^{-\alpha}}{d_i^{-\alpha}}}\right). \end{split}$$
(2)

Using (2) and considering the adopted randomized scheduling policy, the average successful probability $\overline{\phi_i(t)}$ is given as

$$\overline{\phi_i(t)} = p_i \exp\left(-\frac{\beta\sigma^2}{P_t d_i^{-\alpha}}\right) \prod_{j \in \mathcal{V} \setminus \{i\}} \left(1 - \frac{p_j}{1 + \frac{d_i^{-\alpha}}{\beta d_j^{-\alpha}}}\right).$$
(3)



Fig. 1. The hierarchy of the considered IIoT network.

B. Analysis of the Peak Aol

The evaluation of the PAoI is statistically identical for all the D2D pairs, hence, in the following we focus on the derivation of the PAoI of an arbitrary D2D pair. Fig. 2 shows an example evolution of the AoI of the considered preemptive queuing scheme. Let t_j , j = 1, 2, 3, ..., the generation time of the *j*th update. We denote X_j as the random variable that represents the interarrival time between consecutive updates, $X_j = t_{j+1} - t_j$, which follows an exponentially distribution with mean $1/\lambda$. Note that an update may not be received correctly by the D2D receiver due to transmission failures or preemption. Hereafter, we use a different index $i \leq j$ to refer to the successfully received updates. Let d_i refers to the departure time of *i*th update that is successfully received by the corresponding D2D receiver, and S_i is its corresponding service time. We denote g_i as the generation time of the first generated update after d_{i-1} , and is given as

$$g_i \triangleq \min\{t_j \mid t_j > d_{i-1}\}. \tag{4}$$

Therefore, we can see that the indices i and j in general do not refer to the same update. For instance, in Fig. 2, the generated update at t_3 is not received, and the successfully received update at d_2 is the one generated at t_4 . We define $W_i = g_i - d_{i-1}$ as the interval between the reception of the (i-1)th update until the generation of the next update. We define the interval $T_i = d_i - g_i$, which represents the interval from g_i until the next update received successfully. Note that T_i spans the generation instants of failed updates. We also define the interdeparture time between two consecutive successfully received status updates $Y_i = d_i - d_{i-1}$. From the definition of W_i and T_i , we have $Y_i = T_i + W_i$. Therefore, the



Fig. 2. Evaluation of the AoI of an arbitrary node n_i with D = 6.

PAoI, denoted as A_{P_i} (the value of AoI just before receiving the update at d_i), can be given as

$$A_{P_i} = Y_i + S_{i-1}, (5)$$

where S_{i-1} represents the service time of the update received before d_i . For instance, in Fig. 2, the PAoI at d_2 is equal to $Y_2 + S_1$, where S_1 is the service time of the update received before d_2 , which is d_1 . The average PAoI ($\mathbb{E}[A_{P_i}]$) is given by

$$\mathbb{E}[A_{P_i}] = \mathbb{E}[Y] + \mathbb{E}[S] = \mathbb{E}[T] + \mathbb{E}[W] + \mathbb{E}[S], \quad (6)$$

where $\mathbb{E}[Y] = \mathbb{E}[T] + \mathbb{E}[W]$. We denote γ as the probability that an update is preempted. The value of γ is given as a function of the arrival intensity and the service time as follows

$$\gamma = 1 - e^{(-\mu\lambda)},\tag{7}$$

where μ is the duration of one time slot, which represents the deterministic service time S of an update.

For the considered queuing system, when a packet departs, it leaves the system empty, hence, W will follow the same distribution as the interarrival time, i.e., $\mathbb{E}[W] = 1/\lambda$. Moreover, we have $\mathbb{E}[S] = \mu$. Therefore, the term $\mathbb{E}[T]$ can be evaluated using the following recursive method [24] as

Π

$$\mathbb{E}[T] = \underbrace{(1-\gamma)(1-\alpha)\mathbb{E}[S]}_{R_1} + \underbrace{(1-\gamma)\alpha\left(\mathbb{E}[S] + \mathbb{E}[W] + \mathbb{E}[\hat{T}]\right)}_{R_2} + \underbrace{\gamma(\mathbb{E}[X \mid X < S] + \mathbb{E}[\hat{T}])}_{R_3}, \tag{8}$$

where $\alpha = 1 - \phi_i(t)$ represents the failed transmission probability at time slot t. The first term R_1 in (8) denotes the case when the first update (generated at g_i) is not preempted $(1-\gamma)$ by other updates and is received successfully. The term R_2 refers to the case when the first update is not preempted, but the transmission fails. In this case, the system spends the service time S for the first update, then waits for the period W until the next update is generated. The evaluation of \hat{T} is the same as T, hence $\mathbb{E}[\hat{T}] = \mathbb{E}[T]$. The term R_3 represents the case that the first generated update is preempted by a new update. In that case, the effective generation interval, i.e., the generation interval given that a packet, is preempted can be given as

$$\mathbb{E}[X \mid X < S] = \frac{\int_0^{\mu} s\lambda e^{-s\lambda} \, ds}{1 - e^{(-\mu\lambda)}} = \frac{1}{\lambda} + \mu \left(1 - \frac{1}{\gamma}\right). \tag{9}$$

Using (9) in (8) and substituting $\mathbb{E}[S]$ and $\mathbb{E}[W]$, $\mathbb{E}[T]$ can be obtained as

$$\mathbb{E}[T] = \frac{\gamma + \alpha - \gamma \alpha}{\lambda (1 - \beta)(1 - \alpha)}.$$
(10)

Then, $\mathbb{E}[Y]$ becomes

$$\mathbb{E}[Y] = \mathbb{E}[T] + \mathbb{E}[W]$$

= $\frac{\gamma + \alpha - \gamma \alpha}{\lambda(1 - \gamma)(1 - \alpha)} + \frac{1}{\lambda} = \frac{1}{\lambda(1 - \gamma)(1 - \alpha)}.$ (11)

Based on (6) and (11), we obtain the average PAoI for the PR scheme $\mathbb{E}[A_{P_i}]$ as

$$\mathbb{E}[A_{P_i}] = \frac{1}{\lambda(1-\gamma)(1-\alpha)} + \mu.$$
(12)

For a given λ and μ , $\mathbb{E}[A_{P_i}]$ is mainly influenced by $\phi_i(t)$, which depends on p_i . Our goal in this work is to find the optimal scheduling policy π that minimizes the average PAoI $\mathbb{E}[A_{P_i}]$. Therefore, our optimization problem can be formulated as follows

$$\min_{\pi \in \Pi} \mathbb{E}[A_{P_i}]$$
s.t. $0 < p_i \le 1, \ \forall i \in \{1, 2, ..., N\}.$

$$(13)$$

Deriving a closed for expression for the solution of (13) a hard problem due to the spatio-temporal correlation of link states. In the following section, we introduce a DRL algorithm to solve optimization problem in (13).

IV. THE GLI-BASED DEEP LEARNING APPROACH

In this section, we present the design structure of the GLIbased NN to approximate the solution of (13) by mapping the GLI to the scheduling policy. The conventional method to solve the optimization problem in (13) requires a full information about the CSI, which incurs $O(N^2)$ computational complexity for a network of N D2D pairs. The CSI could be mapped using GLI, which is considered as a function of CSI that captures the main feature of the wireless channels (the path loss and shadowing of a wireless link are mostly functions of distance and location). In that sense, we use the GLI as input to the NN to acquire the optimal scheduling policy π^* . The structure of NN is depicted in Fig. 3 and is illustrated in more details as follows.



Fig. 3. The structure of considered NN.

A. D2D Density Grid

First, we construct grid matrices to quantize the continuous form of the locations of the transmitters and receivers as shown in Fig. 4. We assume that the considered network is distributed in a square-shaped area with a side length of L. The whole layout is then partitioned into square cells, where the GLI information is represented as the tuple $\{(x_i^{tx}, y_i^{tx}), (x_i^{rx}, y_i^{rx})\}_{i=1}^N$, where (x_i, y_i) is the index of the cell and the coordinate values ranges from 0 to L. We consider two sub-grids, G^{Tx} and G^{Rx} that represent the activation state of the transmitters and receivers, respectively. For a grid size of $R \times R$, the transmitter matrix G_i^{Tx} of link i is defined as

$$G_i^{Tx} = \begin{cases} 1 & \text{if}(x, y) = \left\lceil (x_i^{tx}, y_i^{tx}) * R/L \right\rceil \\ 0 & \text{otherwise.} \end{cases}$$
(14)

Based on the activation probability p_i , we have

$$G^{Tx} = \sum_{j=i}^{N} p_j G_j^{Tx}.$$
 (15)

Note that the same applies to G^{Rx} . In this case, the matrices G^{Tx} and G^{Rx} represent GLI information to solve the problem in (13) where probability p_i reflects the interference to other links, hence they could be regarded as feature matrices to the convolution layer presented in the next subsection.

B. The 3-Layer Convolution Stage

The two matrices G^{Tx} and G^{Rx} are processed via three connected convolution layers and the output is a set of extracted feature after each layer. Each entry in the resulting matrix comes from a unique convolution positioned at the corresponding index of the input matrix using the convolution filter. When the index represents the receiver's location, then the convolution essentially extracts features from all the transmitters in proximity to this receiver, based on the size of the convolution filter. Both G^{Tx} and G^{Rx} undergo the convolution phase concurrently, producing three matrices each. Then, every link on the D2D plane gleans a total of six features from these matrices, based on the index of its individual receiver or transmitter.



Fig. 4. Construction of the transmitter and receiver density grids.

C. The Fully Connected Stage

The second stage is fully connected stage that comprises two hidden layers. We consider the Rectified Linear Unit (ReLU) as the activation function of each neuron where a sigmoid non-linearity is used at the output node to produce the activation probability p_i . For the considered D2D configuration with N links, the feature vectors for each link are processed through the fully connected layer. This results in a collection of activation probability vectors π . In order to consider the Tx-Rx features in the learning process, we include the p_i from the previous iteration and the distance between the Tx and Rx of the link as two features into the fully connected stage. Particularly, the output probabilities p_i are used as inputs to get new GLI using (15). This in turn helps to improve the convergence of the NN and enhance the training process in general.

D. The NN Training Process

We train the NN using a randomly generated set of D2D layouts to minimize the AoI via gradient descent on the convolutional filter weights and the NN weight parameters. Specifically, the locations of N transmitters are first generated uniformly within the region $L \times L$, and then the locations of the corresponding receivers are generated following a uniform distribution within a pairwise distances of $\{d_{min}, d_{max}\}$. Although the training stage would require the channel gains, it will no longer be required after the network is well-trained and only N GLI is required to obtain the optimal schedule. This way, for each iteration, the NN tends to improve the scheduling policy of the previous iteration.

V. PERFORMANCE EVALUATIONS

In this section, we present the setup of the training and test processes of the considered NN, and evaluate the performance of the proposed AA-DRL via numerical results with the relevant parameters listed in Table I.

A. Network Setup and Training Process

We consider a D2D-assisted IIoT network where N D2D pairs are distributed in 600 meters \times 600 meters region. The

TABLE I EVALUATION PARAMETERS

Parameter	Value
Deployment area	$600 \text{ meters} \times 600 \text{ meters}$
Hidden layer	30 neurons
GLI grid length R	150
Size of convolution filter	10×10
Path loss exponent (α)	3
SINR threshold (β)	0 dB
Noise power (σ^2)	-90 dBm
Transmission power	100 mW



Fig. 5. Comparison of the average PAoI under different values of λ with N = 100 and N = 400.



Fig. 6. CDF of AA-DRL and benchmark with N = 100 and $\lambda = 0.2$ of 1000 different network layouts.

D2D transmitters are uniformly positioned within the deployment area while the locations of the corresponding receivers are generated following a uniform distribution within the pairwise distances { $d_{min} = 2$ meters and $d_{max} = 80$ meters}. We generate 10000 sets of such layout to train the NN and 5000 sets for testing. The NN consists of 3 convolution filters with each filter of size 20×20 . Each hidden layer in the fully connected stage comprises 30 neurons utilizing ReLU and sigmoid functions.



Fig. 7. Log scale values of the computation time required to optimize a single layout under varying N.

B. Numerical Results

The following results show the performance of the proposed AA-DRL scheme in terms of the average PAoI and the computational complexity. Moreover, we use the work in [9] as a benchmark to prove the effectiveness of our proposed approach. Fig.5 compares the average PAoI of AA-DRL and the benchmark in [9] under varying λ . The figure shows superior performance of the proposed AA-DRL over the benchmark, especially when the network size increases from N = 100 to N = 400. For instance, while AA-DRL achieves 55% reduction in the AoI compared to the benchmark at N = 100 and $\lambda = 0.71$, this percentage increases to 75% at N = 400 and $\lambda = 0.81$. The effectiveness of our proposed AA-DRL is attained through incorporating and mapping of the GLI in the scheduling policy, while the benchmark work is based only on local observations of the backlog status of the users, which would be inefficient in high interference regimes.

To further emphasize the effectiveness of the proposed AA-DRL considering different layouts, we plot the Cumulative Distribution Function (CDF) of the average PAoI in Fig. 6 with N = 100 and $\lambda = 0.2$. Although different PAoI values are obtained under different layouts, we can observe that the CDF curves of both approaches have the same trend, hence the performance improvements of AA-DRL could be guaranteed under different network distributions (i.e., different spatial locations).

One concern regarding the proposed AA-DRL would be its corresponding computational complexity. In the following, we roughly analyze the computational complexity of AA-DRL and compare it with an optimal iterative approach. The iterative approach uses a convex optimization solver (e.g., CVX) to obtain a local optimal probability p_i . The access probability p_i is updated in each iteration towards a minimum PAoI based on the obtained p_i from the previous iteration. This way, N one-variable problems are solved in a single iteration. For each iteration, the problem requires the collection of N^2 elements of the CSI matrix, leading to at least $O(N^2)$ computational complexity of this algorithm. For our proposed AA-DRL algorithm, the overall computation is $f \times [R^2 \times (c_1^2 + c_2^2 + c_3^2) + N \times 10 \times h_1 \times h_2]$, where f is the number of feedback rounds, c_i is the size of the convolution filter of stage i and h_1 and h_2 are the number of neurons in each hidden layer. Therefore, for a given D2D layout, the time complexity of AA-DRL scales as O(N). Fig. 7 shows the computation time in log scale of AA-DRL and the iterative method under different values of N. For the sake of reasonable comparison, we chose hardware configurations that are optimally compatible with each algorithm. For AA-DRL, we used Nvidia GPU GeForce GTX 1080Ti, while for the iterative algorithm we used Intel CPU Core i7-8700K @ 3.70GHz. The design of the NN in the proposed AA-DRL is inherently amenable to parallel processing, gaining substantial advantages from the parallel computational capabilities of GPUs. Conversely, the iterative algorithm exhibits inherently sequential computation patterns, making it more suited to CPUs, which offer higher clock speeds. As it is demonstrated by Fig. 7, our proposed AA-DRL shows considerable computational advantages over the iterative approach in large-scale deployments while achieving optimal AoI performance. It is also worth mentioning that the considered IIoT architecture shown in Fig. 1 provides computational advantage where the complexity of AA-DRL would be further relaxed taking advantage of the computational capabilities of the edge node.

VI. CONCLUSION

In this paper, we proposed a DRL-based approach to optimize AoI in D2D-assisted IIoT networks. We formulated a scheduling problem to minimize the average PAoI and developed a NN that maps the GLI to the optimal scheduling policy. The obtained results showed that our proposed approach achieves improved PAoI compared to a recent benchmark, all while exhibiting reduced computational complexity in contrast to the traditional iterative minimization algorithm.

ACKNOWLEDGEMENT

This paper has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 883315.

REFERENCES

- J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, 2017.
- [2] H. Farag, E. Sisinni, M. Gidlund, and P. Österberg, "Priority-aware wireless fieldbus protocol for mixed-criticality industrial wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2767–2780, 2019.
- [3] L. Liu and W. Yu, "A d2d-based protocol for ultra-reliable wireless communications for industrial automation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5045–5058, 2018.
- [4] P. Gil, A. Santos, and A. Cardoso, "Dealing with outliers in wireless sensor networks: An oil refinery application," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1589–1596, 2014.
- [5] Y. Sun et al., "Update or wait: How to keep your data fresh," IEEE Trans. Inf. Theory, vol. 63, no. 11, pp. 7492–7508, 2017.
- [6] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 5381–5394, 2018.

- [7] J. Peng, H. Qiu, J. Cai, W. Xu, and J. Wang, "D2d-assisted multi-user cooperative partial offloading, transmission scheduling and computation allocating for mec," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4858–4873, 2021.
- [8] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1248–1261, 2019.
- [9] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor, "Spatiotemporal analysis for age of information in random access networks under lastcome first-serve with replacement protocol," *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2813–2829, 2022.
- [10] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Transactions* on *Information Theory*, vol. 62, no. 4, pp. 1897–1910, 2016.
- [11] R. Talak, S. Karaman, and E. Modiano, "Improving age of information in wireless networks with perfect channel state information," *IEEE/ACM Transactions on Networking*, vol. 28, no. 4, pp. 1765–1778, 2020.
- [12] C. Li, Q. Liu, S. Li, Y. Chen, Y. T. Hou, W. Lou, and S. Kompella, "Scheduling with age of information guarantee," *IEEE/ACM Transactions on Networking*, pp. 1–14, 2022.
- [13] H. Farag, M. Gidlund, and Č. Stefanović, "A deep reinforcement learning approach for improving age of information in mission-critical iot," in 2021 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAI0T), 2021, pp. 14–18.
- [14] M. Elnourani, S. Deshmukh, and B. Beferull-Lozano, "Distributed resource allocation in underlay multicast d2d communications," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3409–3422, 2021.
- [15] G. Chisci, H. Elsawy, A. Conti, M.-S. Alouini, and M. Z. Win, "Uncoordinated massive wireless networks: Spatiotemporal models and multiaccess strategies," *IEEE/ACM Transactions on Networking*, vol. 27, no. 3, pp. 918–931, 2019.
- [16] M. Liu and L. Zhang, "Resource allocation for d2d underlay communications with proportional fairness using iterative-based approach," *IEEE Access*, vol. 8, pp. 143787–143801, 2020.
- [17] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor, "Optimizing information freshness in wireless networks: A stochastic geometry approach," *IEEE Transactions on Mobile Computing*, vol. 20, no. 6, pp. 2269–2280, 2021.
- [18] M. Li, C. Chen, H. Wu, X. Guan, and X. Shen, "Age-of-information aware scheduling for edge-assisted industrial wireless networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5562–5571, 2021.
- [19] H. Farag, C. Stefanovic, and M. Gidlund, "Distributed backlog-aware protocol for heterogeneous d2d communication-assisted wireless sensor networks," *IEEE Transactions on Mobile Computing*, pp. 1–12, 2023.
- [20] C. Chen, L. Lyu, S. Zhu, and X. Guan, "On-demand transmission for edge-assisted remote control in industrial network systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4842–4854, 2020.
- [21] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid arq under a resource constraint," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1900–1913, 2019.
- [22] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," *IEEE/ACM Transactions on Networking*, vol. 28, no. 1, pp. 15–28, 2020.
- [23] M. Zorzi and R. Rao, "Capture and retransmission control in mobile radio," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 8, pp. 1289–1298, 1994.
- [24] Y. Gu, H. Chen, Y. Zhou, Y. Li, and B. Vucetic, "Timely status update in internet of things monitoring systems: An age-energy tradeoff," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5324–5335, 2019.