

Hybrid Deep Reinforcement Learning and Northern Goshawk Optimization-Based Scheduling Framework for Industrial WSANs

K. Eswaramoorthy³

³Professor, Department of Electrical and Electronics Engineering,
Saveetha Engineering College, Thandalam, Kancheepuram,
Chennai, Tamil Nadu, India. eswaramoorthyk@saveetha.ac.in

Corresponding Author E-mail

ABSTRACT: Industrial Wireless Sensor-Actuator Networks (WSANs) are significant in modern industrial automation for real-time observation, communication, and management of industrial procedures. However, dynamic network conditions, congestion, limited available energy, variable quality of channels, and dissimilar requirements of various applications render scheduling tasks and resources a challenging issue. This paper introduces a smart scheduling system that combines Northern Goshawk Optimization (NGO) and Deep Reinforcement Learning (DRL) for efficient network resource management. The proposed system collects industrial data such as temperature, pressure, vibration and machine status as well as monitors network indicators including remaining energy, queue length, quality of channels, load of traffic, and the importance of packets. NGO is utilized for producing schedules, fitness evaluation, increase resource allocation and scheduling. The scheduling operations received the improvement from the DRL module that deals with observing the state of environment, estimating the reward, implementing the policy and selecting adaptive decision. Subsequently, an intelligent dynamic scheduling engine is employed and implements avoidance of congestion, transmission based on priorities, energy-efficient scheduling and real-time resource allocation. The framework provides scheduling actuators, sensors and emergency alerts in the process of industrial data transmission and controlling operations. OMNeT++ 5.6 is used for simulating the proposed system to test its efficacy in various environments related to industrial networks and various parameters are used to evaluate the proposed system. The DRL-NGO framework achieves a scheduling ratio of 95%, a network life of 1.9 s, and an average time of transmission of 30 ms, thus showing improvement in the scheduling capabilities in dynamic industrial WSAN environments.

Keywords: Industrial WSANs, Northern Goshawk Optimization, Deep Reinforcement Learning.

1. Introduction

Industrial WSANs are used to interconnect sensors and actuators in industrial premises such as production factories, steel manufacturing and oil refining plants [1]. Cost reduction and limited spectrum availability are the principal reasons for the growing trend in WSAN to deploy multiple control loops on one network, as well as for enabling several monitoring and control solutions.

In a WirelessHART, ISA-100 or WIA-PA based industrial WSAN, the control plane is represented by a central network manager controlling the network and control loops [2-4]. For each feedback control loop that monitors process parameters and configurations, the sensors transmit process-related variables to the controlling unit through the network [5]. The controlling unit sends control instructions to the

actuators, which operate the control and security devices [6]. The architecture of WSAN is depicted in figure 1.

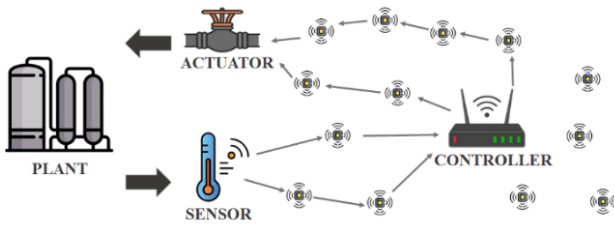


Figure 1: *WSAN architecture*

For instance, the widespread suppliers of WirelessHART networks, Emerson Process Management explored more than 54,835 installations of WirelessHART network technology and collected data over 19.7 billion operating hours [7-8]. More than ten years of use and implementation of WSAN networks proved the possibility of application of the IEEE802.15.4 technology in real business. Nevertheless, configuration of WSAN for its application specification based on Quality of Service (QoS) characteristics is a diligence which implies theoretical calculations, simulation and practical tests among other things [9-11]. Moreover, the need for energy-efficient scheduling is crucial in WSANs as many of the sensors and actuators work with limited energy resources and placed in inaccessible locations. Hence, a proper scheduling mechanism is implemented to perform communication requirements while adhering to QoS standards.

In the literature, a considerable efforts is dedicated to create a theoretical model of low-energy networks and optimization of WSAN operation by changing various MAC layer parameters [12]. Particle swarm optimization (PSO) [13] converges quickly and performs well in scheduling the WSAN resources; but it experience premature convergence and sensitivity to changing parameters. Genetic algorithm (GA) [14] contains a powerful capability of global search. However, its genetic operations lead to high computational complexity. Least-Laxity- First with Remaining

Conflicts (LLF-RC) scheduling algorithm [15] prioritizes time-critical WSAN functions based on their remaining laxities while captivating scheduling conflicts and thus provide low complexity performance; however, it shows limited ability to changing conditions in WSAN including energy conditions, congestions, amount of traffic and channel conditions.

1.1 Research gap

The PSO system is susceptible to premature convergence including sensitivity to parameter selection. The GA method contains high computational processes. The LLF-RC ensures limitations regarding adaptation to changes of energy levels, traffic congestion, queue condition and channel quality. Such limitations indicate that there is a necessity of developing a new scheduling method allowing for optimizing network resources and adaptable to changes occurring in industrial WSAN system. To address this gap, the proposed work develops a Hybrid DRL-NGO system for impending, congestion competent and priority-wise scheduling in industrial WSANs.

1.2 Contributions

The main contributions of this research includes,

- The NGO module generate candidate schedules, fitness evaluation, update population, best schedule selection and resource allocation for finding an optimal solution to scheduling difficulties in industrial WSAN.
- The DRL module implements state observation, reward calculation, policy learning and adaptive decision making to select appropriate scheduling actions as per the network alterations dynamically.
- The hybrid combination of DRL and NGO improves resource allocation mechanisms is WSANs, as well as congestion avoidance, priority-based transmission, and energy efficiency.

2. Proposed Methodology

The Hybrid DRL and NGO-Based scheduling method is presented in this paper for industrial WSANs. The block diagram of the proposed system is implemented in figure 2. The method begins with data collection from industrial sensors

and actuators, continues with network state monitoring that captures important parameters such as residual energy nodes, lengths of queues, traffic loads, conditions of the channel and priority of packets.

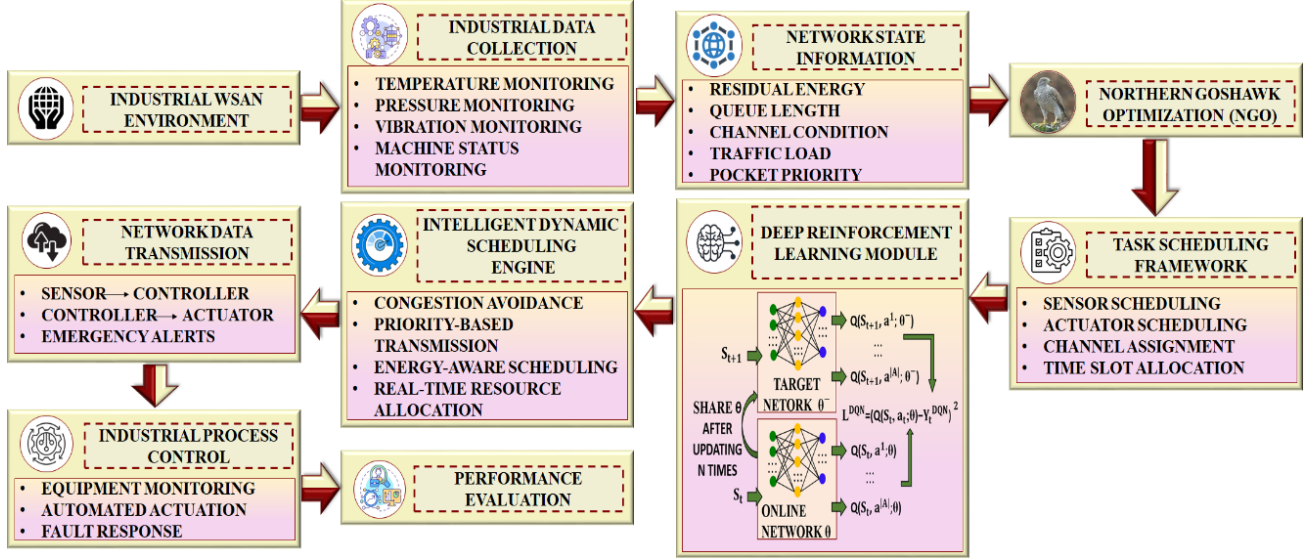


Figure 2: Proposed framework

After network state monitoring, the NGO module performs candidate schedules, fitness evaluation, update populations, selection of best schedule and resource allocation to identify an efficient scheduling solution. After that, the DRL module is provided with the resulting scheduling data which produce an improved scheduling decision by performing the state observation, action selection, reward counting, policy learning and adaptive decision making. The produced decision is processed by the intelligent dynamic scheduling engine that ensures effective and safe transmission of data from the industrial WSAN. Performance evaluation of the system is accomplished through different parameters.

2.1 Data collection and network state monitoring

The scheduling framework is responsible for effective and dependable transferring information in industrial WSANs. Sensors gather industrial data, including temperature, pressure, vibration,

and state of machines. The acquired information is analyzed based on network characteristics, such as remaining energy, queue size, channel conditions, traffic load and packet priority. The scheduler distributes transmission opportunities by employing these metrics and providing significance to important industrial information while selecting energy-saving and reliable communication routes. Proper scheduling lowers the amount of data in the network, packet loss, transmission time, and energy consumption while enhancing communication accuracy and network performance. Next, the generated data is led to the NGO algorithm to optimize the results.

2.2 Northern Goshawk Optimization

The scheduling methodology proposed by NGO is used an effective mission for industrial WSANs. Each feasible solution to the scheduling problem is denoted as a position of goshawk. The array N for potential solutions is defined as,

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N \times m} \quad (1)$$

Here, X indicates the population matrix, X_i refers to the i^{th} candidate solution, N indicates all the potential solutions and m stands for the number of scheduling variables. The potential solutions are evaluated through the fitness functions as,

$$F(X) = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N \times 1} \quad (2)$$

Where, $F(X)$ represents the fitness function array, $F(X_i)$ indicates the i^{th} fitness value of the candidate solution. The flowchart of NGO is illustrated in Figure 3.

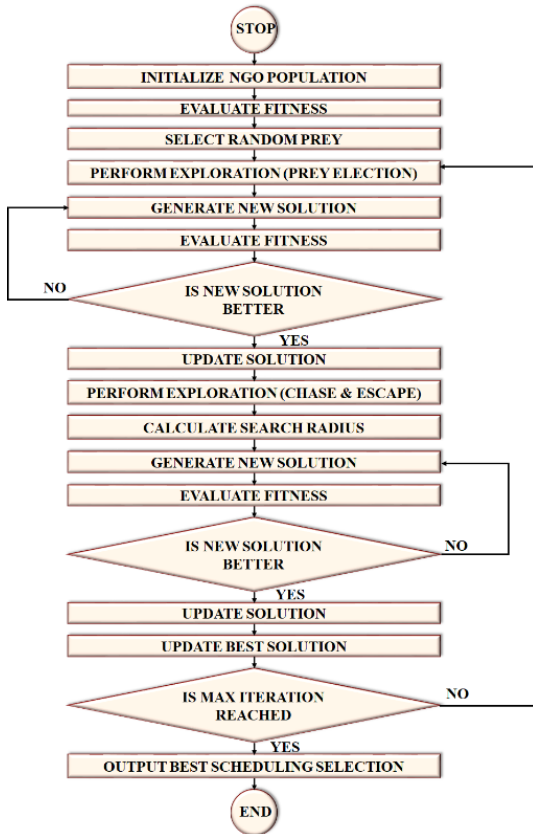


Figure 3: Flowchart of NGO

For minimization problems, less fitness function indicates better scheduling approach. The optimization is conducted in two phases.

Phase 1: Prey Detection (Exploration) assumes that each goshawk randomly selects another population as prey.

$$P_i = X_k \quad (3)$$

Where, P_i denotes the position of the prey by the i^{th} goshawk, X_k is the population founded by random means, k is the index of the candidate solution and i is the current index. The new position is generated as,

$$x_{i,j}^{new,P1} = \begin{cases} x_{i,j} + r(P_{i,j} - I x_{i,j}), & F_{Pi} < F_i \\ x_{i,j} + r(x_{i,j} - P_{i,j}), & F_{Pi} \geq F_i \end{cases} \quad (4)$$

Where, $x_{i,j}$ is the j^{th} scheduling variable of the i^{th} solution, $P_{i,j}$ is the j^{th} variable of chosen prey, r is a random number which is defined in $[0,1]$. The new solution is accepted if it provides better fitness than the previous solution,

$$X_i = \begin{cases} x_i^{new,P1}, & F_i^{new,P1} < F_i \\ x_{i,j}, & F_i^{new,P1} \geq F_i \end{cases} \quad (5)$$

Where, $x_i^{new,P1}$ is the solution obtained through the respective phase, while $F_i^{new,P1}$ is the fitness of the new solution. The generation process covers the required scheduling operations.

Phase 2: Chase and Escape (Exploitation): In this stage the hawk is performing local searching around its efficient solution.

$$x_{i,j}^{new,P2} = x_{i,j} + R(2r - 1)x_{i,j} \quad (6)$$

Here, $x_{i,j}^{new,P2}$ is the updated j^{th} variable, and R is the search radius. The search radius is defined as,

$$R = 0.02 \left(1 - \frac{t}{T}\right) \quad (7)$$

Where, R is the decreasing search radius, 0.02 is the initial scaling factor. As t increases, R is decreased to enable more accurate search near the effective scheduling solution and the new solution is acquired by,

$$X_i = \begin{cases} x_i^{new,P2}, & F_i^{new,P2} < F_i \\ x_{i,j}, & F_i^{new,P2} \geq F_i \end{cases} \quad (8)$$

Here, $F_t^{new,P2}$ is the fitness value. After completion of both phases, the best fitness value and the respective scheduling solution are maintained. The whole process is repeated during T iterations while the population is modified continuously through exploration and exploitation. At the end of optimization process, the best obtained solution is considered as a quasi-optimal scheduling solution for the industrial WSN. The computational complexity of NGO is stated as,

$$O(N.(1 + 2T.m)) \quad (9)$$

Where, N indicates the number of goshawk solutions and m indicates the number of scheduling variables. The final optimized solution obtained from NGO is provided to task scheduling framework, which uses in various processes such as sensor scheduling, actuator scheduling, channel assignment and time slot allocation to execute the task efficiently and reliably in industrial WSN.

2.3 Deep Reinforcement Learning

DRL is used for the intelligent and flexible scheduling of industrial sensors. In value-based DRL approaches, the network state, s_t or state-action pair (s_t, a_t) acts as input to the neural network that approximates the corresponding Q-function or value function. The regression loss when computing the Q-function is given by,

$$L^Q = (Q(s_t, a_t; \theta) - Y_t^Q)^2 \quad (10)$$

Here, Y_t^Q shows the target Q-value and θ is the parameter of the neural network. In Deep Q-Network (DQN) method, the use of the neural network allows one to estimate the Q-values for all possible scheduling actions. The target Q-value is calculated by Bellman's equation as,

$$Y_t^{DQN} = r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta) \quad (11)$$

Here, r_{t+1} is the reward received and γ is the discount factor. To improve training stability during DQN training, an online network and a

target network are used. The target Q-value is calculated as,

$$Y_t^{DQN} = r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta^-) \quad (12)$$

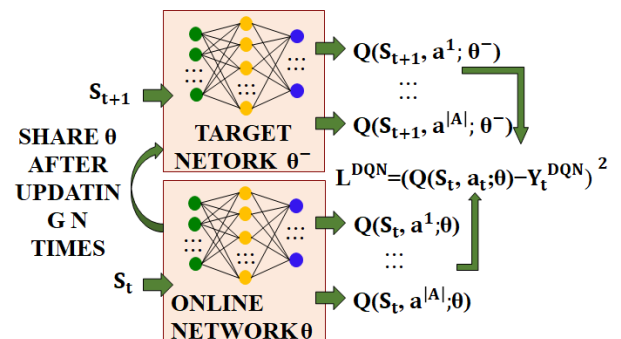
Here, θ^- represents the parameter of the target network. The online network parameters are updated by minimizing the DQN loss through the gradient-based update which is given by the following equation,

$$\theta \leftarrow \theta + \alpha \left(Y_t^{DQN} - Q(s_t, a_t; \theta) \right) \nabla_{\theta} Q(s_t, a_t; \theta) \quad (13)$$

Where, α is the learning rate and the replay technique is employed to save the experience for future sampling randomly, hence reducing correlation in observations and increasing stability of learning process. To overcome the problem of Q-value estimation in the DQN, Double DQN (DDQN) employs the online network for selection and the target network is responsible for action evaluation. The DDQN target value is computed using

$$Y_t^{DDQN} = r_{t+1} + \gamma Q\left(s_{t+1}, \arg \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta); \theta^-\right) \quad (14)$$

The corresponding loss in DDQN is calculated from the Q-function loss replacing Y_t^Q with Y_t^{DDQN} . The diagram of DQN and DDQN is illustrated in figure 4 and 5.

Figure 4: *DQN*

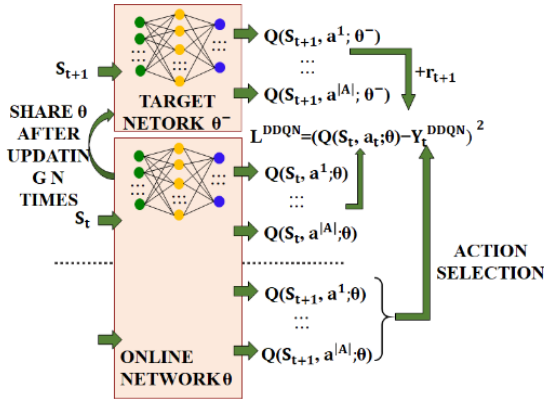


Figure 5: DDQN

Hence, both DQN and DDQN help to train the intelligent dynamic scheduling engine in learning appropriate scheduling decisions based on experience. Learned Q-values are used for congestion avoidance, priority-based transmissions, energy-efficient scheduling and reliable real-time data transmission in the industrial WSAN. Subsequently, the scheduling information is led to industrial data transmission,

where communication is performed between sensor to controller and controller to actuator. To meet the latency and reliability requirements of industrial applications, the scheduling system prioritizes emergency alerts and time-critical control information.

3. Result and Discussion

The dataset is an IIoT network model which is generated by a simulation created in OMNet++ 5.6. This model contains 20 sensor nodes, 8 actuator nodes and a centralized controller which is controlled by a hybrid DRL-NGO scheduler. Sensor data such as temperature, pressure and energy is generated for 1000 seconds in the simulation. The purpose of this simulated environment is to facilitate the evaluation of DRL and NGO optimization conducted for dynamic resource management purposes.

Table 1: Dataset Description

Parameter	Description	Value
Dataset Type	Synthetic Simulation Dataset	Generated
Simulation Tool	OMNeT++ 5.6 Discrete Event Simulator	OMNeT++
Number of Sensor Nodes	Sensing Nodes (Sensors)	20
Number of Actuator Nodes	Actuation Nodes	8
Number of Controller Nodes	Centralized Control Hub	1
Gateway / Scheduler	Hybrid DRL-NGO Scheduling Module	1
Total Device Nodes	Total IIoT Devices	29
Sampling Interval	Sensor Reading / Data Generation Interval	1 s (or your value)
Simulation Time	Total Runtime	1000 s (or your value)
Communication Protocol	Custom OMNeT++ Message Packets	UDP-like / Custom Packets
Traffic Type	Periodic Sensor Data & Actuator Commands	Periodic & Event-Driven
Optimization Algorithms	DRL & NGO Optimizer	Enabled
Fault Scenarios	Node Failure, Congestion, Energy Depletion	Enabled

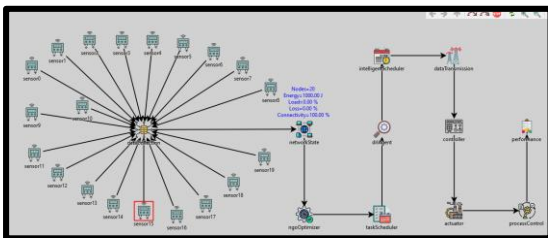


Figure 6: DRL-NGO based industrial WSAN

Figure 6 depicts a WSAN designed for industry based on DRL principles where many sensors send information to a central coordinator. The intelligent scheduler enables effective planning

and distribution of resources, which ensures effective communication among the controller, actuator and industrial process.

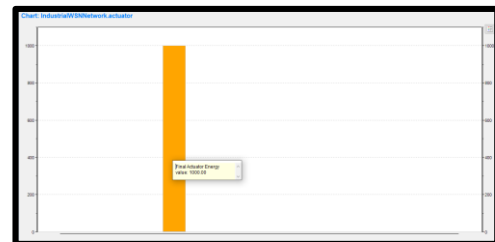


Figure 7: Industrial actuator energy consumption

Figure 7 depicts the final actuator energy for the industrial process with the value of 1000 units. The single bar shows that the reported actuator reaches 1000, which refers to the end-result of energy consumption.

Figure 8 illustrates the performance of different machines of industrial WSN controller using varying parameters. The greatest value of controller energy is 1000 while the values of final temperature and final processing delay are 20. All other parameters require lower values of 10 which shows lower activity in resource usage.

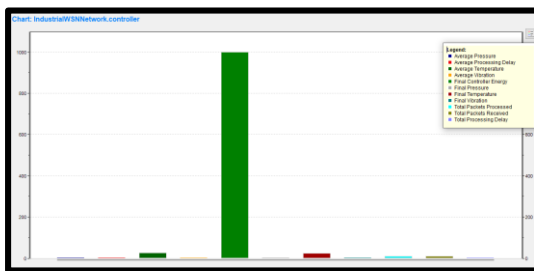


Figure 8: *Controller node performance*

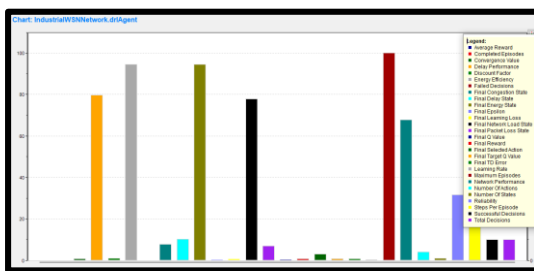


Figure 9: DRL learning metrics

Figure 9 depicts the metrics of agent performance. The final reward and success rate are around 100, while other three values of completed episodes and final savings are 95. The values of energy efficiency are 78, whereas the final network loss state is 68. Other metrics register lower values of less than 35.

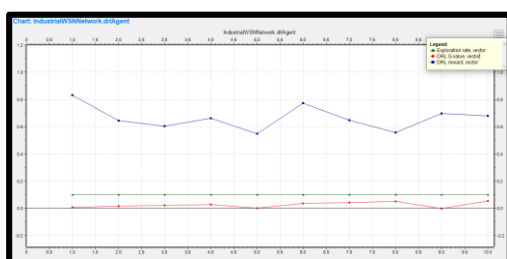


Figure 10: *DRL training and vector analysis*

Figure 10 displays that exploration fluctuates high but both DRL reward and vector increase. Exploration appears lower and stabilizes around 0.6–0.7 at the end.

Figure 11 indicates that NGO best fitness remains virtually unchanged at 0.21 with a slight decrease after the 4th iteration. Overall, the optimization results remain stable.

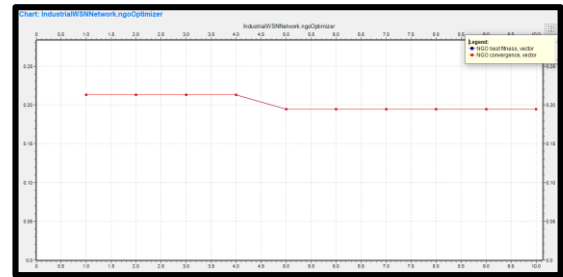


Figure 11: *NGO optimizer convergence*

In Figure 12, the proposed method achieves scheduling ratio of 95% with the total number of 10 nodes while after reaching 80 nodes, the scheduling ratio for the proposed method reduced to 5%, outperforming other methods.

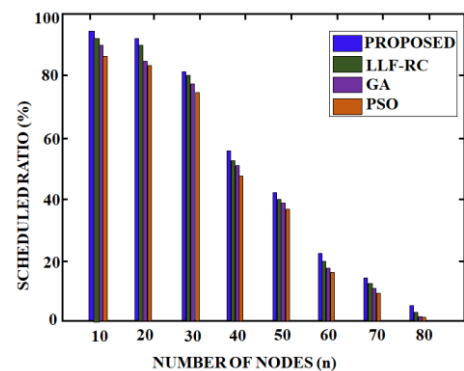


Figure 12: Scheduling ratio

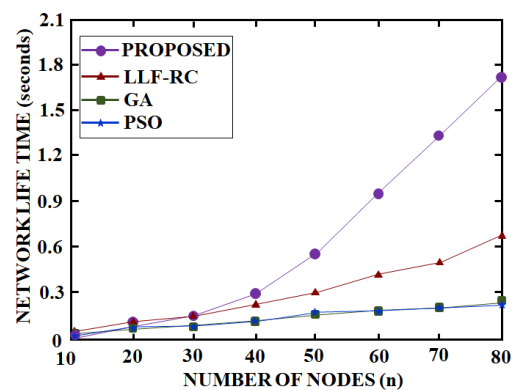


Figure 13: *Network lifetime*

The proposed method in Figure 13 provides the highest lifetime of the network as it is increased with the number of nodes from 0.1 s for 10 nodes to 1.9 s for 80 nodes.

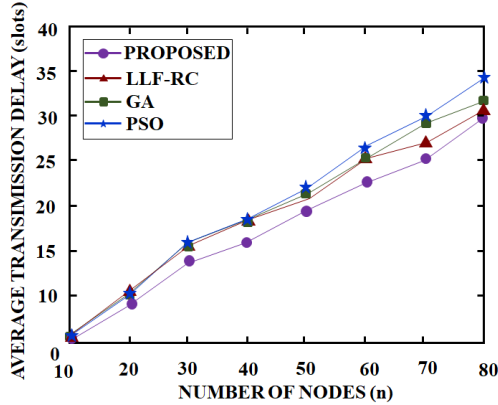


Figure 14: Average transmission delay

The proposed method in Figure 14 relatively provide low transmission delay of 30 ms compared to other methods such as LLF-RC [15], GA [14] and DSO [13].

4. Conclusion

The hybrid DRL and NGO scheduling framework effectively assessed for the industrial WSANs. The data collection phase generated important information about the industry and network such as temperature, pressure, vibration, machine working state, energy, queue length, channel conditions, and packet priorities. The NGO stage analyzed potential schedules for effective allocation of resources and optimal schedule creation. The DRL function is performed to monitor changes in network conditions, define rewards and create appropriate scheduling selections. The proposed approach effectively prevent congestion, prioritize transmissions of packets, save energy and allocate resources in real time. The overall framework is implemented based on OMNeT++ 5.6, and achieves 95% of scheduling time, 1.9 ms of network lifetime and 30 ms of transmission delay, which demonstrates the effectiveness of algorithm for monitoring the flow of sensor, actuator and emergency data in a constantly changing industrial WSANs. Future

work focus on practical implementations, large scale industrial deployments of WSANs and better security methods.

References

1. Mohamed E. ElMenshawy, Year: 2025, "A data-driven decision support system to enhance planning and scheduling in industrial construction projects,"
2. Eden, Shereen Ismail Hunde; Razi Iqbal, Year: 2026, "Simulation and performance analysis of the MBMRF protocol for TSCH-based WSANs", Journal on Wireless Communications and Networking.
3. Venkata Prashant Modekurthy; Mahbubur Rahman; Abusayeed Saifullah, Year: 2023, "Towards mixed criticality industrial wireless sensor-actuator network", In Proceedings of the 24th International Conference on Distributed Computing and Networking, pp. 425-430.
4. Armaghan Darbandi; Myung-Kyun Kim, Year: 2023, "Autonomous scheduling for reliable transmissions in industrial wireless sensor networks", Energies, Vol: 16, No: 20, pp. 7039.
5. Marco Pettorali; Francesca Righetti; Carlo Vallati, Sajal K. Das; Giuseppe Anastasi, Year: 2024, "Mobility management in TSCH-based industrial wireless networks", IEEE Transactions on Mobile Computing, Vol: 23, No: 9, pp. 8710-8728.
6. Xindan Liang; Ying Qiu, Year: 2025, "Fine-Grained Real-Time Flow Scheduling in Wireless TSCH Networks", KSII Trans. Internet Inf. Syst., Vol: 19, No: 7, pp. 2369-2392.
7. Xia Cheng; Mo Sha, Year: 2024, "MERA: Meta-Learning Based Runtime Adaptation for Industrial Wireless Sensor-Actuator Networks", ACM Transactions on Sensor Networks, Vol: 20, No: 4, pp. 1-24.

8. Xuan Zhang; Yong-Long Wang; Heejung Byun, Year: 2025, "Divisive hierarchical clustering for energy saving and latency reduction in UAV-assisted WSANs", *EURASIP Journal on Wireless Communications and Networking*, Vol: 2025, No: 1, pp. 2.
9. Mehbub Alam; Nurzaman Ahmed; Rakesh Matam, Mithun Mukherjee; Ferdous Ahmed Barbhuiya, Year: 2023, "Sdn-based reconfigurable edge network architecture for industrial internet of things", *IEEE Internet of Things Journal*, Vol: 10, No: 18, pp. 16494-16503.
10. Eden Teshome Hunde; Shereen Ismail, Year: 2025, "MBMRF: A modified bidirectional IPv6 multicast protocol with mixed upward and downward forwarding for TSCH-enabled WSANs", *Simulation Modelling Practice and Theory*, Vol: 144, pp. 103172.
11. Lei Guo; Feiya Lv, Year: 2025, "A two-stage deep-learning framework for industrial anomaly detection: integrating small-sample semantic segmentation and knowledge distillation", *Machines*, Vol: 13, No: 8, pp. 712.
12. Paboth Kraikritayakul; Admir Barolli; Shinji Sakamoto; Shunya Higashi; Phudit Ampririt; Leonard Barolli, Year: 2026, "Actor Placement Optimization in WSANs by the PSO-HC-DGA Hybrid System for Two-Zone Industrial Environments", *Sensors*, Vol: 26, No: 5, pp. 1471.
13. Shun-Chi Yu, Year: 2026, "A two-stage fuzzy flexible flow shop scheduling problem with no-wait and re-entrant constraints using adaptive particle swarm optimization", *Scientific Reports*.
14. Akshay Goyal; Aron Alander; Jonathan Hjalmarsson; John Moberg; Ömer Faruk Acar; Ioanna Aslanidou, Year: 2026, "Simulation-Validated Genetic algorithm scheduling for industrial production systems", *Computers & Industrial Engineering*, pp. 111915.
15. Dingwen Yuan; Luis F. Abanto-Leon; Matthias Hollick, Year: 2025, "Scheduling mechanisms in wireless sensor-actuator networks for multirate periodic control in industry 4.0", *IEEE Sensors Journal*, Vol: 25, No. 12, pp. 23037-23055.