IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 8 May 2024, accepted 17 June 2024, date of publication 26 June 2024, date of current version 3 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3419065

RESEARCH ARTICLE

Adaptive VNF Placement Considering Overall Latency and 5G Wireless Channel Reliability in Industry 4.0: A Reinforcement Learning Based Approach

NAUMAN SAQIB[1], NOR FADZILAH ABDULLAH[1], (Member, IEEE), ASMA ABU-SAMAH[1], (Member, IEEE), HAIDER A. H. ALOBAIDY[1,2], (Member, IEEE), AND ROSDIADEE NORDIN[3], (Senior Member, IEEE)

[1]Wireless Research@UKM, Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[2]Department of Information and Communications Engineering, College of Information Engineering, Al-Nahrain University, Baghdad 64074, Iraq
[3]Department of Engineering, School of Engineering and Technology, Sunway University, Bandar Sunway, Selangor 47500, Malaysia

Corresponding author: Nor Fadzilah Abdullah (fadzilah.abdullah@ukm.edu.my)

This work was supported in part by Universiti Kebangsaan Malaysia under Grant DPK-2022-009 and Grant GUP-2021-023.

ABSTRACT Industry 4.0 incorporates the integration of cloud computing, Industrial Internet of Things (IIoT), and modern communication technologies within the industrial automation systems. Various devices with different network requirements of high reliability and low latency, rely on connectivity. The 5G and Beyond (B5G) software-defined architecture facilitates Network Function Virtualization (NFV), which is an essential solution for fulfilling these stringent demands. NFV allows for the implementation and control of Virtual Network Functions (VNFs) in dynamic network environments. VNF placement optimization has been extensively studied in the 5G perspective outside the industry environment with a focus on minimizing delay and cost, increasing VNF reliability, and increasing resource efficiency. However, the complex dynamics of the wireless channel in industrial environments have a considerable impact on the essential delay factors that are important for optimizing the deployment of VNFs. This study focuses on modeling a Wireless Sensor Network (WSN) based Industry 4.0 factory automation scenario at mmWave band, formulating an optimization problem to minimize overall delay while considering packet loss rate in the 5G industrial wireless channel. The optimization problem is formulated as a Markov Decision Process (MDP) and two Reinforcement Learning (RL) based algorithms AVP-Q and AVP-DQN are proposed for optimizing the VNF placement. The proposed algorithms are extensively evaluated against the Value Iteration algorithm which assumes a completely known MDP model and two other algorithms from the literature. The simulated results show that AVP-DQN outperforms existing algorithms for this scenario by 39% and 22.6% and the achieved performance is only close to that of the Value Iteration algorithm.

INDEX TERMS 5G, industry 4.0, deep reinforcement learning, VNF placement, URLLC, mmWave.

I. INTRODUCTION

The Internet of Things (IoT) is a network that connects humans and objects through information-sensing devices and actuators. The IoT has received significant attention from both academia and industry in recent years, primarily driven by the rapid increase of interconnected devices and the demand for more efficient approaches to managing large numbers of these devices [1]. The development objectives of the Industrial Internet of Things (IIoT) and Industry 4.0 encompass the use of highly distributed intelligent computing and networking advancements in industrial production

The associate editor coordinating the review of this manuscript and approving it for publication was Olutayo O. Oyerinde.

and manufacturing systems. These advancements aim to enhance automation, quality, and control [2], [3]. Ensuring reliable Quality of Service (QoS) for IoT applications deployed in harsh industrial environments is difficult. Under such circumstances, communication experiences a significant reduction in signal strength due to obstructions, hence increasing the complexity of the network design phase [4]. IIoT networks experience computationally demanding, diverse, and complex industrial service requests that are time-sensitive, particularly when it comes to real-time surveillance, computation, and collaboration among various intelligent agents [5], [6].

5G and beyond (B5G) communication enables Ultra-Reliable Low Latency Communication (URLLC) services, which is significant progress in facilitating new IIoT applications with wireless connectivity [7], [8]. The 5G network is built with a software-defined architecture, which enables flexible programming to create distinct layers for various applications [9]. The performance characteristics of 5G, including its exceptional reliability of up to 99.999%, extremely low latency of less than 1 ms, and minimal power consumption, effectively address the limitations of current communication technologies in the industry [10]. Edge and fog computing, as well as multi-access edge computing (MEC) [11], play a crucial role in facilitating these applications. The primary idea is to enhance conventional cloud computing by placing computational resources in closer proximity to clients and end devices. Through these methods, both end devices and central cloud servers can transfer computing workloads to resources located at the edge or the fog. This leads to decreased latency and increased reliability [12]. Careful management of the network resources included within a distributed and heterogeneous infrastructure is necessary to satisfy the stringent latency and reliability requirements of industrial applications. Interconnected by public and private 5G networks, the underlying infrastructure comprises cloud/edge resources [13] that provide execution environments for Virtual Network Functions (VNFs). The VNF is based on software applications that provide network functionalities, such as network segmentation and traffic management within a factory automation network, running on top of existing hardware infrastructure, virtualized on servers. It is important to note that certain 5G core functions, like the User Plane Function (UPF), have a strong connection with the application VNFs [14]. Therefore, it is crucial to consider the interaction between these functions while placing VNFs at certain nodes. It is necessary to locate both the services and the UPF at the MEC servers close to the IoT sensors. This would effectively reduce the latency that sensors experience.

Wireless sensor networks (WSN) are being used in the Factory of the Future (FoF) to monitor processes and their relevant parameters in an industrial environment. The monitoring of this environment is commonly conducted by utilizing diverse sensor technologies, such as microphones, CO_2 sensors, pressure sensors, humidity sensors, and thermometers. The data collected from the system is analyzed using machine learning techniques to identify any abnormal patterns or deviations. Smart factories have become self-sustaining, economically efficient, and automated by incorporating wireless communications into pre-existing private networks. WSN fully utilizes wireless technologies for constructing industrial network infrastructure [15]. Machine interference, signal attenuation from various materials, and ever-changing industrial layouts are just a few of the issues that wireless communication channels face in the industry.

In this context, resource orchestration poses a challenge as it strives to consistently identify the optimal arrangement of software components that provide the desired service. The existing work discussed in section II primarily focuses on optimizing VNFs with a strong emphasis on cost and resource efficiency. The objective is to achieve a balance between the inexpensive deployment of VNFs and the efficient utilization of network resources. Furthermore, the analysis has focused on the reliability of VNFs, which are often used as backup functions to guarantee uninterrupted operation in the event of breakdowns. However, in the environment of Industry 4.0, where wireless communication channels are integrated into production floors and industrial facilities, the influence of these channels on the optimized placement of VNFs has not been investigated thoroughly. The interference from equipment, signal attenuation, and the inherent unpredictability of the industrial environment all contribute to the wireless channel conditions, which raises issues beyond typical optimization concerns. Understanding and applying the wireless channel model into the optimization framework is essential for adjusting VNF deployments to dynamic situations. Using cutting-edge reinforcement learning methods, this study aims to connect the areas of wireless channel models and VNF location optimization.

The contributions of this paper are summarised as follows:

1) Modeling of a wireless sensor network (WSN) based industry 4.0 factory automation scenario, formulating an optimization problem to minimize the overall delay considering the reliability of 5G industrial wireless channel at mmWave band.
2) Formulation of the optimization problem as Markov Decision Process (MDP), and the development of two Reinforcement Learning (RL) based algorithms to optimize the VNF placement.
3) Comparative evaluation of the proposed algorithms, through extensive simulations, with the Value Iteration algorithm which assumes a fully known MDP model leading to the most optimal solution. Two other algorithms from existing literature are also used for the evaluation of the proposed algorithms.

The rest of the paper is organized as follows. Section II presents the existing work in 5G VNF placement optimization. Section III presents the detailed system model, the problem formulation, and the details of the proposed RL-based algorithms. Section IV presents the performance evaluation of the proposed algorithms against a benchmark on simulated network traffic. Section V draws the conclusion.

II. RELATED WORK

Network slicing has gained significant attention in both academics and industry in recent times. With the increasing prevalence of virtualization technologies, there has been a growing interest in the allocation of VNFs on physical resources. The introduction of 5G technology has made it easier to implement Network Function Virtualization (NFV) architecture. The authors of [16] provide a comprehensive survey of the challenges and solutions for VNF placement in 5G networks and discuss a network virtualization approach that changes traditional network functions that run on non-standard hardware into software-based virtual machines running on standard hardware. However, there is a need to place these virtual resources optimally otherwise it may impact the QoS. Existing research commonly approaches the placement of VNFs as an optimization problem, which is typically classified as NP-hard [17].

The study in [18] focuses on optimizing the positioning and implementation of VNFs considering both the Edge and Cloud infrastructures. They formulate the problem as mixed integer programming (MIP), minimizing the total delay experienced by the user and the operational cost incurred by the service providers. Similarly, the study in [19] uses dynamic multi-objective optimization. In each operational cycle, the framework receives and analyses network traffic, determines the most relevant objective functions based on traffic state, recomputes and then deploys the solutions. Authors in [20] also consider the deployment of VNFs in the central cloud and edge nodes aiming to solve minimization of latency and maximization of service availability. They propose a meta-heuristic genetic algorithm (GA) to solve this optimization problem. The experimental results demonstrate that their GA yields near-optimal solutions in less time than an exact algorithm by CPLEX [21]. CPLEX is an integer programming solver by IBM.

Extensive work has been done by formulating the optimization problem as an Integer Programming (IP) problem and using heuristic algorithms to find the solution. However, considering the stochastic network environment and dynamic workload, many researchers have explored the RL framework utilization in optimizing the virtual resources placement problem. In [22] two policy-based RL algorithms, Proximal Policy Optimisation (PPO2) and Advantage Actor-Critic (A2C), are suggested for dynamic Service Function Chain (SFC) placement. The objective is to minimize energy usage while considering the requisite availability levels specified by the customer and the Service Level Agreement (SLA). The model is defined as MDP in which SFC requests are handled sequentially. The RL algorithms outperform the greedy method in terms of energy usage and acceptance rate. Similarly, [23] introduces a Deep Deterministic Policy Gradient (DDPG) RL method which aims to address the dynamic placement of VNFs between edge and cloud networks. The suggested method offers superior VNF placement in terms of meeting SLA requirements, minimizing end-to-end latency, and optimizing network resources as compared to competing options. The dynamic nature of the network environment and regulations may not align with the algorithms that rely on prior information to provide optimal solutions. A novel approach is introduced in [24] to address the issue of non-stationary traffic situations. This approach utilizes hybrid Deep Reinforcement Learning (DRL)-heuristic algorithms to effectively handle changes in traffic. This system integrates Advantage Actor Critic (AAC) with a Graph Convolutional Network (GCN). The results demonstrate that in a practical non-stationary network environment, the proposed hybrid DRL-heuristic approach is more reliable than pure DRL.

VNF placement optimization has been extensively studied in the 5G perspective outside the industry environment. However, within the industrial environment, the latency and reliability requirements are more strict. Also, the consideration of reliability of wireless channel is important in the industrial environment. In the context of wireless channel modeling, authors in [25] discuss in detail the wireless channel propagation modeling and characterization for wireless IoT technologies. Similarly, authors in [26] considered two 5G spectrums of 3.5 GHz and 28 GHz in an industrial automation environment to investigate the reliability in terms of network density deployment and frequency diversity. Authors in [27] investigate the VNF placement in industrial edge systems by minimizing the overall cost of deployment. They formulate the problem as an IP problem and use a heuristic method to solve it. The study in [28] formulates the optimization problem considering the cost of computing resources, communication links cost, and VNF migration cost for vertical industries in B5G. The authors solve the problem using Mixed Integer Linear Programming (MILP) first and then propose two meta-heuristic algorithms based on the GA approach. The results show that GA-based algorithms reach the optimal solution with less computational complexity. Table 1 summarizes the approaches used in the literature for solving the complex problem of VNF placement. Based on the literature survey, this study focuses on the placement optimization of VNFs in an industrial automation environment. The goal is to minimize the overall delay in fulfilling the service requests generated by WSNs in various areas of a factory with consideration of the reliability of the wireless channel in the industrial environment.

III. ADAPTIVE VNF PLACEMENT SYSTEM MODEL

This section introduces the factory model that is being used in this study, as well as the proposed algorithms for solving the optimization problem at hand.

A. 5G FACTORY MODEL

This study focuses on analyzing a physical WSN that is implemented in an indoor factory environment at mmWave

TABLE 1. Summary of VNF placement work.

Ref	Aim/Focus of Study	Technique Used	Parameters Optimized	Limitations
Nemeth et al. [29]	Cost-minimizing VNF placement optimization to meet reliability and low latency requirements of mobile robotics in the industrial environment	Heuristic algorithm using bin packing approach	Delay, Radio coverage, Battery, Cost, Resource utilization	Shared VNFs and round trip delay not considered [30]
Behravesh et al. [31]	Joint user association, SFC placement and resource allocation problem is formulated and solved.	MILP and heuristic approach	Cost, Resource utilization, service interruption	No comparison to other algorithms
Domenico et al. [32]	Optimization of joint VNF deployment and resource allocation in 5G-hybrid C-RAN architecture	Integer Linear Programming (ILP)	Resource utilization	Cost of resources not considered
Zhang et al. [33]	VNF placement considering different QoS requirements in smart city	ILP based heuristic algorithm	QoS	Energy and cost of resources not considered
Kiran et al. [34]	Minimizing VNF placement and resource cost on MEC nodes	Genetic Algorithm based heuristic approach	Resource cost	Delay of services not considered
Golkarifard et al. [35]	Minimizing cost by joint decision of VNF placement, resource allocation and 5G traffic routing	MINLP based heuristic algorithm	Resource cost	Energy consumption not considered
Mu et al. [36]	Minimizing the total energy consumption considering the shared resources for VNFs	DRL based DDAP	Resource cost and energy	Latency of services not considered in optimization problem
Santos et al. [22]	RL based dynamic SFC placement minimizing energy consumption and maximizing availability	A2C, PPO	Energy, Availability	Cost of resources not considered
Kibalya et al. [37]	RL-based approach for placement of VNF considering reliability using backup VNF while minimizing resource cost.	RL	Cost, Reliability	Energy consumption and delay factor is not considered
Araújo et al. [38]	ML technique to reduce the number of network components and runtime of services at substrate network level	ILP, ML	Cost, resource, latency	Energy consumption not considered
Dalgkitsis et al. [23]	Automate VNF deployment between edge and cloud nodes minimizing latency	DDPG	Latency, resource utilization	Resource cost not optimized
Pei et al. [39]	DRL based VNF placement optimization considering VNF placement and running cost, and penalty of rejecting SFCs	Double DQN	VNF placement and running cost, Reject ratio of SFCs	End-to-end latency considered as a constraint on SFC but not optimized
Ali et al. [40]	Implement traffic forecasting model to learn an ML-based VNF scaling model and an RL-based VNF placement model	ILP, Supervised ML, PPO	Server resource capacity, Overall latency between VNFs	Latencies considered are maximum allowed latencies between VNFs and not the end-to-end latency of an event. Latencies of the wireless channel not considered

band and uses a private 5G network for the provision of services. The sub-6GHz band is not considered in this study to remove the limitations on data rates, congestion, and latency in simulation. 5G technology offers a significant increase in bandwidth capacity and the primary constraint of network resources lies in their computational capacity rather than in network transmission. Therefore, this work assumes that the physical network is fully connected. This assumption is made to simplify the complexities of routing and data transmission scheduling at the network layer. Meanwhile, the TCP traffic is assumed for the event transmission, and the congestion control is addressed by the migration of data to the optimized location of the serving node. The event is defined as any activity/traffic generated by the sensor. The main focus of

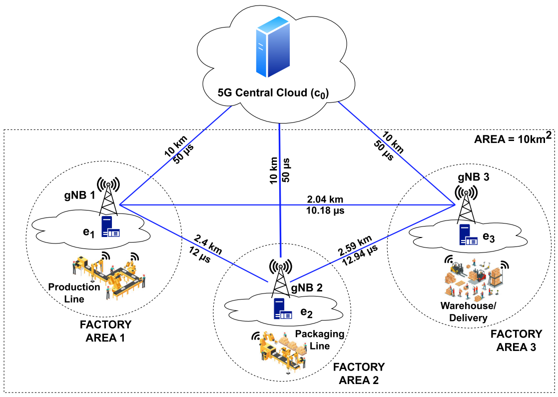

FIGURE 1. Industry 4.0 indoor environment.

the study is on solving the VNF placement problem at the application layer. In this scenario, the overall delay can be defined as the total of the computational time, transmission time, queuing time, and migration time. These parameters will be discussed in detail later in this section.

Fig. 1 shows an indoor factory setup in an area of 10 km^2, inspired by the TESLA GIGA Nevada [41]. The factory areas such as production, packaging, and warehouse have WSN deployed in their respective areas which are being served by private 5G gNBs. The network is represented as an undirected graph $G = (\mathcal{J} \subseteq \mathcal{K}, \mathcal{L})$ where $\mathcal{J} = \{1, 2, \ldots, J\}$, $\mathcal{K} = \{1, 2, \ldots, K\}$ and $\mathcal{L} = \{1, 2, \ldots, L\}$ denote a set of edge nodes, gNBs, and physical links respectively. The set of edge nodes is a subset of the set of gNBs. The establishment of connections, including wireless links between sensors and gNBs, interlinks between gNBs, and physical connections between 5G central cloud and gNBs, constitute the communication infrastructure. The 5G core network is assumed to be located at the central cloud and all the network functions communicate through a common interface and can be located at any edge node [42]. Each gNB can host an edge cloud, which is one of the essential requirements by the Next Generation Mobile Network Alliance to improve the 5G systems' flexibility [43]. Each gNB accommodates a maximum of one edge cloud. A network service request also referred to as a service function chain (SFC), consists of VNFs arranged in a specific order. Service request is modelled as $F = \left(N^V, L^V\right)$ where N^V is a set of VNFs and L^V is a set of links required for servicing the request. The wireless sensors within the factory area are grouped based on their gNBs, and this grouping is represented as $U = K$. Each traffic flow produced by a sensor of the group U at its local gNB is identified by its gNB identifier. The serving node refers to the central cloud or edge node that provides hosting for the VNF. An edge cloud can accommodate multiple VNFs, and a physical link can be associated with multiple virtual links. The latency of a connection is determined by the data transfer rate and the distance between the two nodes. A node can be either a gNB or the central cloud.

TABLE 2. Glossary of symbols.

Parameter	Description
F	Service request
N^V	Set of VNFs
L^V	Set of physical links
c_0	The 5G central cloud
e_i	The i-th edge node
e_x	The serving node
l_{ix}	The latency of wired connection between i-th edge node and the serving node
b_i	The i-th gNB
d_q	Queuing delay
d_r	Wireless link delay
$n_i(t)$	Number of events at i-th edge node
$m_i(t)$	Amount of data to be migrated from i-th edge node
$d_{comp}(t)$	The computational delay of VNF server
$d_{comm}(t)$	The communication delay of wired and wireless links
$d_{migr}(t)$	The migration delay of data transfer when VNF is shifted to different edge node

TABLE 3. Simulation parameters.

Parameter	Value
$l_{e_1-e_2}$	12 µs
$l_{e_1-e_3}$	10.18 µs
$l_{e_2-e_3}$	12.94 µs
$l_{c_0-e_1/e_2/e_3}$	50 µs
d_q	0.1 ms
d_r	0.138 ms
Data rate W	100 Mbps
Average number of Packets X	6000 Packets
Packet Size b	32 bits

The glossary of symbols used in the study is given in Table 2. The 5G central cloud is denoted as c_0. The i-th edge node is denoted as e_i and the serving node is represented by e_x. l_{ix} represents the latency between the i-th edge node and the serving node. n_i is the number of events occurring at the i-th edge node. The i-th gNB is represented as b_i.

Three edge nodes hosted by three different gNBs covering different areas of factory were modeled in a simplified scenario for this study with different events rates described as scenarios later in this section. Every edge node is assumed to have multiple CPU cores, with a total frequency of 10 GHz [44]. It is also assumed that these resources are sufficient to serve the events arriving at the edge node. The computational resources of the central cloud are assumed to be unlimited. The simulations were conducted on a computer equipped with an Intel Core i7 2.4 GHz CPU and 16.0 GB of RAM. MATLAB 2023a was used for the study. The important parameter values used in the simulations are summarized in Table 3.

The latency between edge nodes 1-2 ($l_{e_1-e_2}$) and 1-3 ($l_{e_1-e_3}$) is 12μs and 10.18μs respectively. The latency between edge nodes 2-3 ($l_{e_2-e_3}$) is 12.94μs. The latency between central cloud and edge nodes 1,2 and 3 $l_{c_0-e_1/e_2/e_3}$ is

50μs each. These wired links are based on optical fiber in this study and the associated delays of optical fiber links based on distance are discussed in [45]. The average queuing delay (d_q) for each event at the serving node is 0.1 ms [46]. As per 3GPP [47], the required data rate (W) for the links between nodes is 100 Mbps while the average packets generated per event (X) are 6000. The size of each packet (b) is 32 bits.

A simulated input dataset comprising the daily events arrival at each gNB over a span of 1000 days is created. Every day is divided into 1440 timesteps, with each timestep representing 1 minute in a real-world situation. The events generation from sensors follow a discrete Poisson distribution [48], [49] and multiple scenarios based on different event rates are created. The actual number of events fluctuates daily around the mean value mentioned in the scenarios.

- *Scenario 1 (s1)*: Each gNB receives 50 events per day to be processed by the serving node.
- *Scenario 2 (s2)*: gNB 1 receives 50 events per day to be processed while gNB 2 and gNB 3 receive 100 events per day each.
- *Scenario 3 (s3)*: gNB 1 receives 50 events per day,gNB 2 receives 100 events per day and gNB 3 receive 200 events per day.
- *Scenario 4 (s4)*: gNB 1 receives 50 events per day while gNB 2 and gNB 3 receive 250 events per day each.
- *Scenario 5 (s5)*: gNB 1 receives 50 events per day to be processed while gNB 2 and gNB 3 receive 500 events per day each.

B. PROBLEM FORMULATION FOR VNF PLACEMENT

Fig. 1 depicts a typical example of network slicing where a logical network is created as per 5G infrastructure [50]. This logical network processes the information received through sensors based on their generated events. This service requires a robust network that can provide sufficient quality of service (QoS) in all circumstances. Each sensor in a particular area can only be connected to one gNB through which its requests are served. In normal conditions, the delay requirements are satisfied by deploying VNFs in central cloud c_0. Due to some unscheduled occurrences i.e. detection of some malfunctioning in one of the assembly lines in a particular area of the factory, during a certain time of the day, the number of events generated by the sensors in that area may increase significantly. This increases the load on the network between central cloud c_0 and the affected gNB b_i, which results in delays greater than 5-10 ms and exceeds the Key Performance Indicator (KPI) recommendation by 3GPP [28], [47] for deployment of WSN in Industry 4.0.

To decrease the latency in the affected area, MEC VNFs might be deployed on the closest edge node, therefore creating an additional slice to process and respond to the events by the sensors in that area. To maintain the required QoS, it might be necessary to shift the MEC VNFs to other edge node based on the random occurrences of the events generated by the sensors. The challenges and issues in VNF placement and migration are discussed in detail in [51]. To prevent delays during deployment and migration operations, it is imperative to develop a VNF deployment policy that considers all pertinent elements, including the stochastic traffic load of the sensors. Because of the limited computational resources of an edge node, financial costs, and migration delays associated with the creation of VNFs, this study assumes that the network slice created for deployment of VNFs will be hosted at only one of the gNBs in a factory. The factors considered for selecting a deployment policy for VNF at either central cloud c_0 or one of the edge nodes e_i are discussed below.

1) DELAY MODEL

Deploying MEC VNF on the serving edge node e_x costs extra time while processing the events generated by the sensors due to the limited resource availability at edge nodes. Therefore, in any MEC system, the non-negligible task execution time must be considered. On the other hand, the processing and queuing time on the central cloud c_0 is considered to be negligible owing to its more powerful resources. Let the number of events at time instant t for i-th edge node be represented as $n_i(t)$, the computational resources that can be allocated to process one event as $f_i(t)$, the average number of packets transmitted per event as X and the size of the packet in bits as b then the processing time for one event at edge node e_x can be calculated as [44]:

$$d_{p,i}(t) = \frac{X \times b}{f_i(t)} \tag{1}$$

In addition, when the event generated at time t reaches edge node e_x, there can be incomplete events that are waiting to be processed because of the limited computing capabilities of the edge node. Hence, it is important to consider the waiting time in the queue. Let this queuing delay be represented as $d_{q,i}(t)$ then the total computational delay $d_{comp}(t)$ for the generated number of events $n_i(t)$ at time instant t is the sum of $d_{q,i}(t)$ and $d_{p,i}(t)$ and can be represented as [44]:

$$d_{comp}(t) = \sum_{i=1}^{J} \left[n_i(t) \times \left(d_{p,i}(t) + d_{q,i}(t) \right) \right] \tag{2}$$

Several links are involved in communication between sensors and VNFs. At first, the data generated by sensors is transmitted through a wireless channel to its local gNB. The characteristics of the wireless channel considered in this study are discussed in section III-B2. The radio link delay d_r [52] is given in Table 3. Secondly, based on the gNB that is hosting the VNFs, inter-gNB links and links between gNB and the central cloud may be involved. Let l_{ix} be the optical link delay between the i-th edge node and the serving node, and d_r be the wireless link delay then the communication delay $d_{comm}(t)$ is represented by [44]:

$$d_{comm}(t) = \sum_{i=1}^{J} \left[n_i(t) \times (l_{ix} + d_r) \right] \tag{3}$$

Deploying an MEC VNF in an edge node as proposed in [11] and moving the unprocessed sensor data from the previous edge node leads to extra delays. The migration methodology involves duplicating the MEC VNF and its sensor data to a new edge node, while the original node remains operational. Once a certain threshold is met, the original MEC VNF is halted, and the remaining sensor data is transferred. The events are then served with the new MEC VNF on the serving node. The study in [53] addresses the optimized methods of migrating the microservices in 5G architecture. Assuming the same computational power at all edge nodes, the only delay experienced by the sensors is the time during which data is being transferred from the previous serving node to the new serving node. Therefore, this study focuses mainly on this migration downtime for the migration delay calculation. The migration downtime d_{migr} is calculated as [44]:

$$d_{migr}(t) = d_{init} + \frac{1}{W}\sum_{i=1}^{J} m_i(t) \tag{4}$$

where d_{init} is the initialization time of the VNFs at the new serving cloud, $m_i(t)$ is the amount of data that needs to be migrated from i-th edge node to the serving node at time instant t and W is the data rate with which the data can be migrated.

2) PACKET LOSS MODEL

Packet losses in wireless channels have a direct impact on the reliability of a wireless channel since they might result in partial or incorrect delivery of information. The more frequent these packet losses occur, the less reliable the wireless channel becomes. In recent years, mmWave communications have become a potential contender to meet the growing need for throughput and latency in many use cases. mmWave allows for high data transfer rates critical for factory automation use cases such as machine vision for real-time inspection and quality control, high-definition video streaming for remote monitoring and maintenance, or downloading large software updates for industrial machines. mmWave also offers significantly lower latency when it comes to applications like real-time control of robotic arms and other automated machinery, or time-sensitive communication between machines and central control systems etc. However, the packet loss model assumption is most relevant due to its sensitivity to blockage and attenuation. This research employs mmWave at 60GHz band to meet strict industrial communication requirements. The packet loss is not considered at the wired connections with the assumption of reliable fiber links between the nodes. The study follows the stochastic channel model presented in [54] that is a Floating Intercept (FI) path loss model and is given as:

$$PL(d) = \delta + 10\beta \log_{10} d + X_\sigma \tag{5}$$

The unit of PL is dB, δ is the floating-point intercept in dB, β is the path loss dependence on distance, X_σ represents the large-scale PL fluctuation with a random Gaussian variable and d is the distance between transmitter and receiver in meters.

The range of Bit Error Rate (BER) is obtained against different Signal to Noise Ratio (SNR) values using the above path loss model and then the packet loss rate is determined from BER values [55]. Based on the packet loss rate, the packet losses are simulated for the timestep when there is any event according to the traffic generated for each scenario defined previously. The distance of sensors is considered to be within 10 m range. For the transmissions against each event generation, the BER in the range of 10^{-4} to 10^{-1} is adopted from [54] for the determination of packet loss rate.

C. MODELING ALGORITHMS

1) FORMULATION OF PROBLEM AS MDP

In the adaptive VNF placement scheme, the goal is to obtain an RL-based policy that places VNFs at the location that minimizes overall delay and considers the packet losses in the wireless channel in stochastic network traffic conditions over a time window T. Consideration of both latency and the reliability of wireless channel is suitable for URLLC application in factory automation. As minimizing overall delay (d_{Total}) and considering the wireless channel reliability using the Packet Loss Rate (P_{LR}) are considered simultaneously, this problem is formulated as a multi-objective integer programming optimization problem presented in (6).

$$\text{minimize}\left(\sum_{t=1}^{T} d_{Total}(t)\ ,\ P_{LR}(t)\right) \tag{6}$$

where

$$d_{Total}(t) = d_{comp}(t) + d_{comm}(t) + d_{migr}(t) \tag{7}$$

$$P_{LR}(t) = \frac{P_L(t)}{P_T(t)} \tag{8}$$

where P_L are the lost packets and P_T are the total transmitted packets. The weighted sum method can be used to transform the multi-objective into a single-objective optimization problem where ω_1 and ω_2 are the weighting factors such that $\omega_1 + \omega_2 = 1$. Normalization is also applied on d_{Total} to keep both objective functions on the same scale. The final optimization problem is given in (9)

$$\text{minimize}\left\{\sum_{t=1}^{T}\left(\omega_1\left(\frac{d_{Total}(t) - d_{min}}{d_{max} - d_{min}}\right) + \omega_2 P_{LR}(t)\right)\right\} \tag{9}$$

where d_{max} and d_{min} are the maximum and minimum overall delays.

This optimization problem is formulated as MDP framework in which state, action, and reward function are defined as follows.

- *State s*: Let the entire state space be S and the state $s(t) \in S$ at time instant t be defined to contain information about the serving node and number of events at all nodes.

Let there be K gNBs in the network then:

$$s(t) = \{e_x, n_1, n_2, \ldots, n_K\}$$

where $e_x = 0$ if VNFs are deployed at the central cloud and $e_x = i$ if VNFs are deployed at i-th edge node and $1 \leq i \leq K$. n_i refers to the number of events generated by the sensors at the i-th gNB.

- *Action a*: The action $a(t)$ at time instant t is defined to be the selection of a node for the deployment of VNFs. Let A be the entire action space then $a(t) \in A$ is given as

$$a(t) = \{i\}, \quad i \in \{0, 1, 2, \ldots, N\}$$

where $a(t) = 0$ means VNFs are deployed at the central cloud while $a(t) = 1, \ldots, N$ means VNFs are deployed at i-th node.

- *Reward R*: MDP models involve the optimization of an objective function, usually a discounted cumulative reward. The reward function $R(s, a)$ quantifies the immediate benefits, expenses, or impacts linked to certain decisions, contingent upon the state s and/or action a. Reinforcement learning algorithms typically aim to maximize cumulative rewards. The goal of adaptive VNF placement is to minimize the overall delay along with minimization of packet loss rate as formulated in (9). Hence the reward function is developed in (10) by taking the negative of overall delay and packet loss rate to formulate a maximization objective. By defining the reward function as the negative value of the metric that needs to be minimized, the algorithm aims to maximize this negative value, hence achieving the objective of minimizing the original metric. This approach simplifies the optimization process and maintains consistency with the standard practice in reinforcement learning.

$$R = -\sum_{t=1}^{T} \left(\omega_1 \left(\frac{d_{Total}(t) - d_{min}}{d_{max} - d_{min}} \right) + \omega_2 P_{LR}(t) \right) \tag{10}$$

2) VALUE ITERATION

Value Iteration is a Dynamic Programming (DP) approach used to solve MDPs and find the best policies. It works by iteratively updating the values of states until they converge to the optimal values. The algorithm uses the Bellman Equation [56] to show how the value of a state affects the expected total reward, considering the transition probabilities associated with each state. Complete knowledge of the MDP model, including transition probabilities, is an essential requirement for Value Iteration.

To implement the Value Iteration algorithm for this study, the transition probability function is adopted from [57] and the events arrival rate is assumed to be known. Because of the complexity of the model, transition probabilities are based on the events arrival rate only, and the packet loss rate is not incorporated. As the MDP model is fully known, the algorithm converges to the best optimal policy.

The computational complexity of Value Iteration is $O(|S|^2 \cdot |A|)$ [58], where $|S|$ is the total number of states of MDP and $|A|$ is the total number of actions. Though computationally intensive, the approach provides an optimal policy and maximizes the cumulative reward. However, in real-world applications, it is often complicated to obtain a complete and accurate MDP model.

3) Q-AGENT BASED ADAPTIVE VNF PLACEMENT

Q-learning is a type of reinforcement learning that does not require a model and is used to determine the optimal action to take based on the present state of the agent. The complete model of the MDP e.g. transition probabilities, is not required to be known.

In this study, one of the proposed RL based algorithms is Q-Agent based Adaptive VNF placement (AVP-Q) presented as Algorithm 1 in this section. Lines 5-13 form a loop in which the agent does an action in the current state s to transition to the next state s'. This transition results in an immediate reward R, which is then used to update the Q-table. In the Q-table, $Q(s, a)$ reflects the expected long-term reward of choosing action a in state s. Line 7 determines the actions for state s based on the ϵ-greedy policy.

The learning rate (α) controls the extent of the modifications made by the agent to its Q-values in response to new information. While a lower learning rate offers more stability but slower learning, a higher learning rate enables faster adaptation to new experiences but may also cause instability. The discount factor (γ) determines the agent's evaluation of future benefits. In situations such as VNF placement where the implications could be long-term, the agent is more likely to prioritize long-term reward when the γ value is higher. This leads to more strategic decision-making. A lower γ value, on the other hand, places more emphasis on immediate reward, which results in more short-term decisions. The exploration rate ϵ balances between exploiting the action with the highest Q value and exploring additional actions randomly.

The AVP-Q can dynamically change the placement scheme in real time based on the current event status and packet loss conditions. The upper bound on the computational complexity of Q-Learning is $O(n^2)$ [59].

4) DQN-AGENT BASED ADAPTIVE VNF PLACEMENT

The Deep Q-Network (DQN) is a reinforcement learning technique that enhances the conventional Q-learning method by integrating deep neural networks to estimate the Q-values. This enables the algorithm to handle complex and high-dimensional state spaces, making it particularly effective in scenarios with continuous observation spaces. The ability to handle continuous observation spaces makes DQN Agent more flexible and versatile, suitable for a variety of real-world applications.

The second proposed RL based algorithm in this study is the DQN-based Adaptive VNF Placement (AVP-DQN),

Algorithm 1 Q-Agent Based Adaptive VNF Placement (AVP-Q)

1: Initialize Q-table as per $|S|$ and $|A|$
2: Set hyperparameters: learning rate (α), discount factor (γ), exploration rate (ϵ)
3: **for all** episodes **do**
4: Start at initial state
5: **for** each step **do**
6: Read current events conditions and packet losses at each gNB
7: Explore or exploit based on ϵ
8: Execute action, observe next state
9: Compute reward R using (10)
10: Update Q-value using Bellman Equation:

$$Q(s,a) = Q(s,a) + \alpha \cdot [R + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

11: Transition to the next state
12: **end for**
13: Adjust ϵ
14: **end for**
15: The resulting Q-table signifies learned optimal VNF placements

presented as Algorithm 2, which allows for adaptability and flexibility in handling larger network state spaces. In AVP-DQN, the agent interacts with the environment according to an ϵ-greedy strategy. During each interaction, the agent observes network events and packet losses. The agent chooses actions and stores experiences in an experience replay memory (D). Periodically, a mini-batch is sampled from this memory, and the Q-network is trained using a loss function which is computed as the mean squared difference between predicted Q-values and target Q-values. The trained Q-network, after the convergence criteria are met, contains the optimal VNF placement strategy. The DQN agent uses Double Q-Learning [60] to overcome the overestimation of action values.

AVP-DQN uses a neural network with three fully connected hidden layers. The input layer, which receives state observations, consists of $J+1$ units where J is the number of edge nodes, followed by hidden layers of 64,128 and 64 units respectively. The output layer of N units represents the Q-values associated with each potential action. Rectified Linear Unit (ReLU) activation function is used for non-linearity. The neural network architecture was carefully selected after a thorough assessment of several topologies to maximize reward outcomes. The thorough testing procedure enabled the choice of an architecture that regularly provided excellent performance in improving overall rewards. The neural network is trained with a mean squared error loss function and an Adam optimizer.

The selection of batch size and update frequency in DQN significantly influences the learning dynamics and

Algorithm 2 DQN-Agent Based Adaptive VNF Placement (AVP-DQN)

1: Initialize replay memory (D), Q-network weights (θ), target network weights (θ')
2: Set hyperparameters: learning rate (α), discount factor (γ), exploration rate (ϵ), batch size, and update frequency
3: Initialize state (S) and action (A) spaces
4: Initialize Q-network $Q(s,a,\theta)$ with weights θ and target Q-network $Q(s,a,\theta')$ with weights θ'
5: **for all** episodes **do**
6: Initialize state s.
7: **for** each step **do**
8: Read current events conditions and packet losses at each gNB
9: Select action a using ϵ-greedy strategy from $Q(s,a,\theta)$
10: Deploy VNF at the selected placement position, observe new state s'
11: Compute reward R using (10)
12: Save experience tuple (s,a,R,s') in replay memory D
13: Compute target Q-values, update θ using loss from mini-batch
14: Periodically update target Q-network weights θ'
15: **end for**
16: Adjust ϵ
17: **end for**
18: The trained Q-network represents optimal VNF placements

efficiency of the agent. The batch size specifies the number of experiences taken from the replay buffer for updating the neural network's weights in each training iteration. Increasing the batch size can yield more consistent and precise updates, but it may require greater computational resources. On the other hand, reducing the batch size may increase the amount of variation, but it can result in quicker learning with lower computational requirements. Similarly, the update frequency determines the frequency at which the target neural network is updated with experiences from the replay memory. Increasing the frequency of updates enables the agent to promptly adjust to new information, but it may also result in increased correlation among experiences. Reducing the frequency of updates can help reduce correlation, but it may lead to slower learning.

The computational complexity of DQN-based algorithms is determined by several parameters, such as the neural network's size, the state space's dimensionality, the environment's complexity, and the number of training episodes. The complexity increases with the increase in the number of hidden layers. If the number of layers of DQN neural network is represented by L, and the number of neurons in layer $l \in L$ is represented by u_l, the time complexity of DQN based algorithm is $O(MT\sum_{l=0}^{L-1} u_l u_{l+1})$ [61] where M is the number of iterations in each episode and T is the number of episodes.

IV. PERFORMANCE EVALUATION

For the evaluation of AVP-Q and AVP-DQN, median daily reward is chosen as the evaluation metric. The training process consists of 1000 episodes with each day representing one episode. The median reward of last 140 days (episodes) was used to evaluate the performance of the algorithms. The median reward is chosen rather than the average because the median value is more resistant to outliers that could occur in the data than the average value. The event rate scenarios were simulated using the Monte Carlo approach, with events created using the Poisson distribution over various random seeds. Based on the results of these Monte Carlo trials, the reinforcement learning agents were trained to achieve the reported outcomes.

AVP-Q and AVP-DQN were compared with the Value Iteration algorithm presented in section III-C2 and two other algorithms namely One Step Look Ahead and Random Location described in [44]. One step look ahead algorithm evaluates the anticipated results of each placement selection and chooses the placement that maximizes the network's performance in the next time step. This algorithm is like RL algorithms but with short-term reward only. The random location algorithm chooses a VNF placement randomly at a timestep in which there is an event. It retains that location until the events on that gNB are served completely.

A. RL HYPERPARAMETERS TUNING

The optimal performance of RL agents is dependent on the selection of suitable hyperparameters. Properly selected hyperparameters have a substantial influence on the convergence rate, stability, and overall performance of these agents. The selected hyperparameters for AVP-Q and AVP-DQN are summarized in Table 4. AVP-Q was implemented utilizing a Q-table structure. The learning rate of 0.01 resulted in better performance when compared to other rates, as depicted in Fig. 2. Other hyperparameters were selected after a methodical assessment, enhancing the overall optimization of the model's performance. The discount factor was set to 0.99, giving priority to future rewards. The starting exploration rate was set at 0.9 and decreased progressively (Epsilon Decay) to 0.1 over 1050 episodes. This approach allowed for adequate exploration in the early stages while also encouraging the convergence towards optimal policies.

In AVP-DQN, the DQN agent is trained on a learning rate of 0.0001 selected among the multiple candidates as shown in Fig. 3. Although 0.00001 produces a slightly greater reward, the chosen learning rate exhibits a more efficient convergence by achieving its highest reward value in fewer episodes. To improve stability, a replay memory buffer of size 10,000 was implemented. This allows the agent to learn from a wide range of past experiences. To address the issue of frequent changes in target values causing instability, a target network was also implemented and updated using the smoothing factor (τ) of 0.001.

TABLE 4. Hyperparameters for AVP-Q and AVP-DQN.

Parameter	AVP-Q	AVP-DQN
Learning Rate (α)	0.01	0.0001
Discount Factor (γ)	0.99	0.99
Exploration Rate (ϵ)	0.9	0.9
Epsilon Decay	0.1	0.01
Replay Buffer Size	-	10000
Batch Size	-	64
Smoothing Factor (τ)	-	0.001

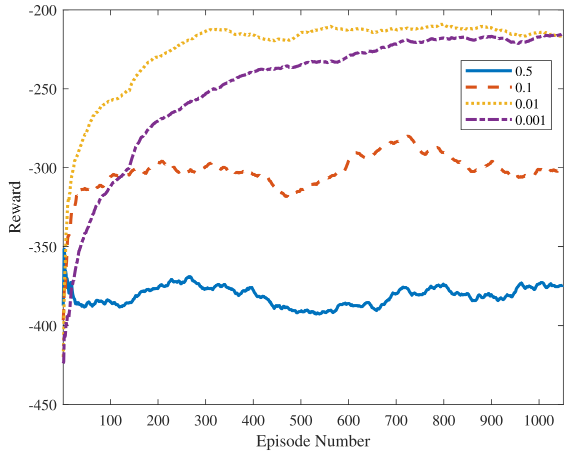

FIGURE 2. Learning rate comparison for AVP-Q.

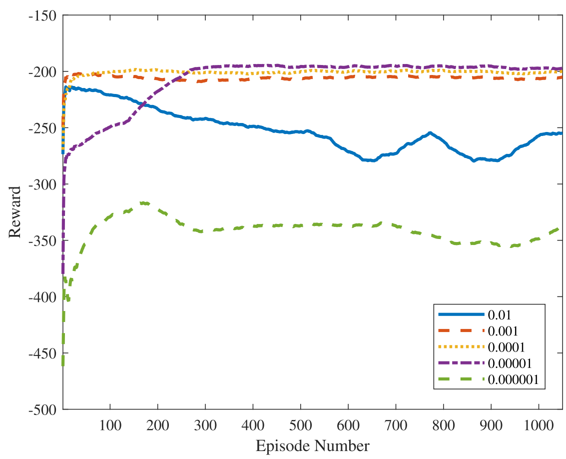

FIGURE 3. Learning rate comparison for AVP-DQN.

B. OPTIMIZATION BASED ON OVERALL DELAY ONLY

In this section AVP-Q and AVP-DQN are compared with Value Iteration, One Step Look Ahead, and Random Location algorithms while considering overall delay only.

Fig. 4 shows that the value iteration performs best in all scenarios. This is because the MDP model is fully known in this case and the reward converges to its best possible value.

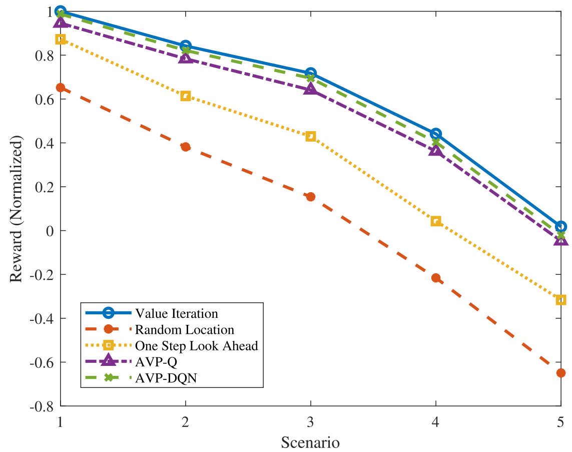

FIGURE 4. Delay minimization comparison.

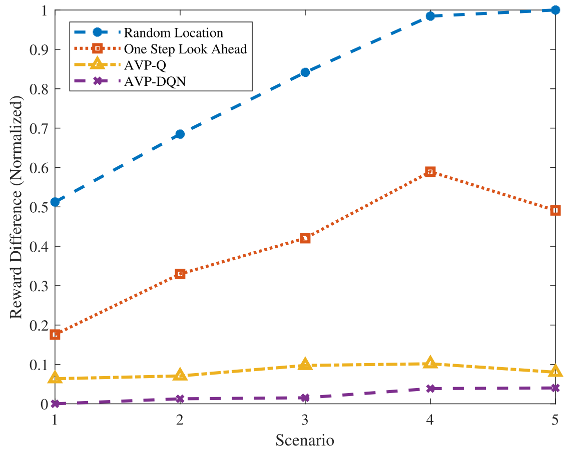

FIGURE 5. Absolute reward difference with value iteration algorithm.

AVP-Q and AVP-DQN provide comparable performance to the value iteration algorithm for all considered scenarios. On average over all scenarios, Value Iteration shows an advantage of 9.64% and 4.05% only over AVP-Q and AVP-DQN respectively. These results show that both algorithms are equally effective in addressing the optimization problem.

Similarly, on average over all scenarios, AVP-Q outperforms Random Location and One Step Look Ahead algorithms by 43.2% and 23.83% respectively. AVP-DQN shows even more superiority with 47.41% and 29.8% average improvement over Random Location and One Step Look Ahead algorithms respectively.

It can also be observed from Fig. 4 that as the number of events per day increased, AVP-Q and AVP-DQN adapted to this change and kept their performance close to the Value Iteration algorithm while Random Location and One Step Look Ahead resulted in lower reward values. Fig. 5 shows the normalized absolute reward differences between Value Iteration and every other algorithm across all scenarios. It can be observed that as the rate of events increases Random Location and One Step Look Ahead move away from the Value Iteration while AVP-Q and AVP-DQN maintain their performance.

C. OPTIMIZATION BASED ON OVERALL DELAY AND 5G WIRELESS CHANNEL RELIABILITY

Fig. 6 shows the results when channel reliability factor is added to the reward function. The ω_1 and ω_2 which are the weight factors in (10) are set to $\omega_1 = \omega_2 = 0.5$. Because of the packet losses, the retransmissions occur and the delay increases, hence reducing the reward value. Given the complexity of the MDP model, especially when accounting for channel reliability, the Value Iteration algorithm is excluded from consideration due to its potential inability to produce the most optimal solution with a partial MDP model.

AVP-Q surpasses Random Location and One Step Look Ahead algorithms on average over all scenarios by 34.67% and 16.85% respectively. Similarly, AVP-DQN performs

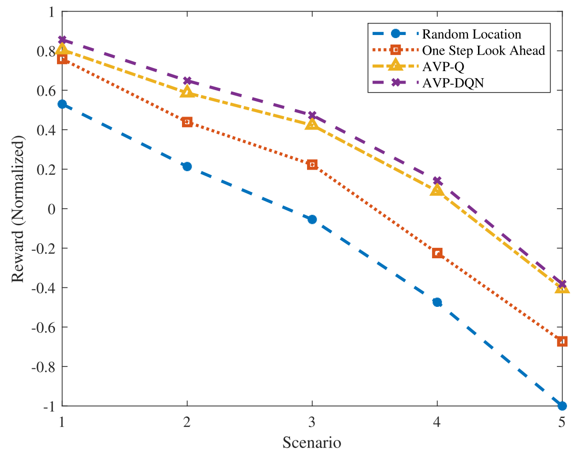

FIGURE 6. Integrated performance: minimizing delay and packet loss rate.

better than Random Location and One Step Look Ahead algorithms by 39% and 22.6% respectively. The results show that even in the complex scenario while considering the wireless channel reliability in terms of packet losses, AVP-Q and AVP-DQN adapt to the environment to maximize the reward and get the optimal solution among the contending algorithms.

Fig. 7 shows the boxplot of normalized rewards obtained from 1050 episodes for each algorithms considering scenario 3. The normalized Inter Quartile Range (IQR) values are 0.21, 0.23, 0.18 and 0.12 for Random Location, One Step Look Ahead, AVP-Q and AVP-DQN respectively. The lower interquartile range shows that throughout 1050 episodes, the episode reward values are more stable and consistent in the case of AVP-DQN.

In summary, for the considered VNF placement scenario, both AVP-Q and AVP-DQN perform better than Random Location and One Step Look Ahead algorithms. Although AVP-Q almost reaches the same optimal solution as AVP-DQN but from Fig. 2 and 3 for the selected learning rates it can be concluded that AVP-Q converges to its optimal

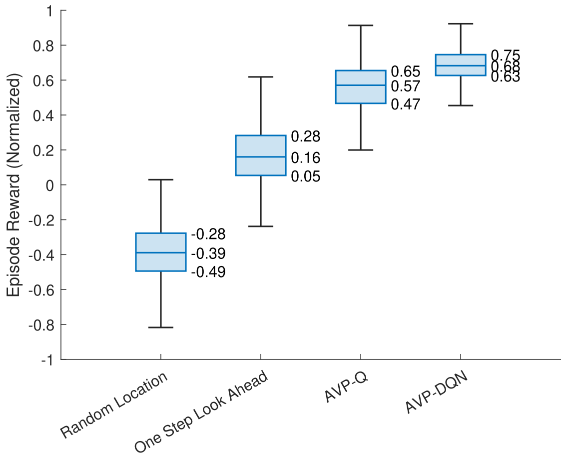

FIGURE 7. Algorithm performance distribution across 1050 episodes.

solution in fewer episodes than AVP-DQN. Also, AVP-DQN has the advantage of continuous state space representation in its model so in case of a complex network model where state space is large, AVP-DQN is preferred over AVP-Q.

V. CONCLUSION

This study focused on the hard problem of optimizing the placement of VNFs in 5G networks, specifically focusing on the issues present in the Industry 4.0 environment. The study formulated an optimization problem that focused on reducing overall delay and considered the reliability of 5G industrial wireless channel at mmWave band using packet loss rate as the reliability parameter. Two RL-based algorithms were presented that can adapt to changing wireless conditions. These algorithms were rigorously evaluated using extensive simulations. The results show that the Value Iteration algorithm has an advantage of 9.64% and 4.05% only over AVP-Q and AVP-DQN respectively while considering the overall delay only. In minimizing the overall delay while considering the effects of the wireless channel as well, AVP-DQN performed better than AVP-Q and surpassed the Random Location and One Step look Ahead algorithms by 39% and 22.6% respectively. AVP-DQN was also shown to be more consistent and converged to the optimal solution in fewer episodes than AVP-Q. The thorough evaluation of these algorithms through extensive simulations, benchmarking against the optimal Value Iteration algorithm, and comparison with algorithms from existing literature emphasized the importance of taking into account wireless channel dynamics to achieve robust and efficient VNF placements in Industry 4.0 settings.

The future work involves exploring the performance of AVP-DQN in complex industrial environments such as large network sizes, increased femtocells in a factory area, and evolving resource availability. The current study is focused on a controlled simulated environment. Future research will examine the use of AVP-DQN in an actual industrial environment to evaluate its scalability, adaptability, and real-time performance. In addition, investigating advanced RL agents such as Dueling DQN, Proximal Policy Optimization (PPO) and Deep Deterministic Policy Gradient (DDPG) shows potential for better performance in more complex scenarios. Moreover, it is worth exploring the integration of VNF reliability into the placement strategy to guarantee the quality of services.

ACKNOWLEDGMENT

The authors extend their gratitude to the author of [57] for the provision of partial source codes.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] W. Ayoub, A. E. Samhat, F. Nouvel, M. Mroue, and J.-C. Prévotet, "Internet of Mobile things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and supported mobility," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1561–1581, 2nd Quart., 2019.

[2] M. W. Condry and C. B. Nelson, "Using smart edge IoT devices for safer, rapid response with industry IoT control operations," *Proc. IEEE*, vol. 104, no. 5, pp. 938–946, May 2016.

[3] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial Internet of Things: A cyber-physical systems perspective," *IEEE Access*, vol. 6, pp. 78238–78259, 2018.

[4] S. Singh, P. Alam, P. Kumar, and S. Kaur, "Internet of Things for precision agriculture applications," in *Proc. IEEE Int. Conf. Image Inf. Process.*, Nov. 2019, pp. 420–424.

[5] T. Ma, Y. Wang, W. Hu, D. El-Banna, and K. Zhang, "Evaluation of flexilink as unified real-time protocol for industrial networks," in *Proc. 13th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, May 2018, pp. 123–128.

[6] C. Sauer, E. Lyczkowski, M. Sliskovic, and M. Schmidt, "Real-time alarm dissemination in mobile industrial networks," in *Proc. 22nd IEEE Int. Conf. Ind. Technol. (ICIT)*, vol. 1, Mar. 2021, pp. 1152–1156.

[7] 5G-ACIA, "5G for industrial Internet of Things (IIoT): Capabilities, features, and potential," 5G Alliance for Connected Industries Autom. (5G-ACIA), Frankfurt, Germany, Tech. Rep., Nov. 2021. [Online]. Available: https://5g-acia.org/whitepapers/5g-for-industrial-internet-of-things/

[8] M. S. Jit Singh, W. S. Wan Saleh, A. T. Abed, and M. A. Fauzi, "A review on massive MIMO antennas for 5G communication systems on challenges and limitations," *Jurnal Kejuruteraan*, vol. 35, no. 1, pp. 95–103, Jan. 2023.

[9] S. K. Rao and R. Prasad, "Impact of 5G technologies on industry 4.0," *Wireless Pers. Commun.*, vol. 100, no. 1, pp. 145–159, May 2018.

[10] Z. Albataineh, A. Andrawes, N. Abdullah, and R. Nordin, "Energy-efficient beyond 5G multiple access technique with simultaneous wireless information and power transfer for the factory of the future," *Energies*, vol. 15, no. 16, p. 6059, Aug. 2022.

[11] S. Kekki, W. Featherstone, Y. Fang, P. Kuure, A. Li, A. Ranjan, D. Purkayastha, F. Jiangping, D. Frydman, G. Verin, K.-W. Wen, K. Kim, R. Arora, A. Odgers, L. M. Contreras, and S. Scarpina, "MEC in 5G networks," Eur. Telecommun. Standards Inst. (ETSI), ETSI, Sophia Antipolis, France, White Paper, no. 28, Jun. 2018.

[12] E. N. Pencheva, I. I. Atanasov, and V. G. Vladislavov, "Mission critical messaging using multi-access edge computing," *Cybern. Inf. Technol.*, vol. 19, no. 4, pp. 73–89, Nov. 2019.

[13] 5G-ACIA, "5G non-public networks for industrial scenarios," 5G Alliance for Connected Industries Autom. (5G-ACIA), Frankfurt, Germany, Tech. Rep., Jul. 2019. [Online]. Available: https://5g-acia.org/whitepapers/5g-non-public-networks-for-industrial-scenarios/

[14] D. Harutyunyan, R. Behravesh, and N. Slamnik-Kriještorac, "Cost-efficient placement and scaling of 5G core network and MEC-enabled application VNFs," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, May 2021, pp. 241–249.

[15] M. Majid, S. Habib, A. R. Javed, M. Rizwan, G. Srivastava, T. R. Gadekallu, and J. C.-W. Lin, "Applications of wireless sensor networks and Internet of Things frameworks in the industry revolution 4.0: A systematic literature review," *Sensors*, vol. 22, no. 6, p. 2087, Mar. 2022.

[16] W. Attaoui, E. Sabir, H. Elbiaze, and M. Guizani, "VNF and CNF placement in 5G: Recent advances and future trends," *IEEE Trans. Netw. Service Manage.*, vol. 20, no. 4, pp. 4698–4733, Aug. 2023.

[17] B. Sonkoly, J. Czentye, M. Szalay, B. Németh, and L. Toka, "Survey on placement methods in the edge and beyond," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2590–2629, 4th Quart., 2021.

[18] A. Leivadeas, G. Kesidis, M. Ibnkahla, and I. Lambadaris, "VNF placement optimization at the edge and cloud," *Future Internet*, vol. 11, no. 3, p. 69, 2019.

[19] A. Ocampo, N. Tapia, and D. P. Pinto-Roa, "Virtual network function chaining placement based on dynamic multi-objective optimization and multi-criteria decision making," in *Proc. Latin Amer. Netw. Conf.* New York, NY, USA: Association for Computing Machinery, 2022, pp. 2–9.

[20] L. Yala, P. A. Frangoudis, and A. Ksentini, "Latency and availability driven VNF placement in a MEC-NFV environment," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.

[21] IBM. *IBM Ilog CPLEX Optimizer*. Accessed: Mar. 23, 2024. [Online]. Available: https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-optimizer

[22] G. L. Santos, T. Lynn, J. Kelner, and P. T. Endo, "Availability-aware and energy-aware dynamic SFC placement using reinforcement learning," *J. Supercomput.*, vol. 77, no. 11, pp. 12711–12740, Nov. 2021.

[23] A. Dalgkitsis, P.-V. Mekikis, A. Antonopoulos, G. Kormentzas, and C. Verikoukis, "Dynamic resource aware VNF placement with deep reinforcement learning for 5G networks," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[24] J. J. A. Esteves, A. Boubendir, F. Guillemin, and P. Sens, "DRL-based slice placement under non-stationary conditions," in *Proc. 17th Int. Conf. Netw. Service Manage. (CNSM)*, Oct. 2021, pp. 225–233.

[25] H. A. H. Alobaidy, M. Jit Singh, M. Behjati, R. Nordin, and N. F. Abdullah, "Wireless transmissions, propagation and channel modelling for IoT technologies: Applications and challenges," *IEEE Access*, vol. 10, pp. 24095–24131, 2022.

[26] A. M. Ramly, N. F. Abdullah, and R. Nordin, "Cross-layer design and performance analysis for ultra-reliable factory of the future based on 5G mobile networks," *IEEE Access*, vol. 9, pp. 68161–68175, 2021.

[27] M. Liu and S. B. Alias, "Cost-efficient virtual network function placement in an industrial edge system: A proposed method," *IEEE Syst., Man, Cybern. Mag.*, vol. 9, no. 1, pp. 10–17, Jan. 2023.

[28] R. Fedrizzi, R. Behravesh, C. Costa, and F. Granelli, "An adaptive GA-based slice provisioning method for vertical industries in 5G and beyond networks," *Comput. Netw.*, vol. 218, Dec. 2022, Art. no. 109397.

[29] B. Németh, N. Molner, J. Martín-Pérez, C. J. Bernardos, A. de la Oliva, and B. Sonkoly, "Delay and reliability-constrained VNF placement on mobile and volatile 5G infrastructure," *IEEE Trans. Mobile Comput.*, vol. 21, no. 9, pp. 3150–3162, Sep. 2022.

[30] I. Dimolitsas, D. Dechouniotis, and S. Papavassiliou, "Time-efficient distributed virtual network embedding for round-trip delay minimization," *J. Netw. Comput. Appl.*, vol. 217, Aug. 2023, Art. no. 103691.

[31] R. Behravesh, D. Harutyunyan, E. Coronado, and R. Riggio, "Time-sensitive mobile user association and SFC placement in MEC-enabled 5G networks," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 3, pp. 3006–3020, Sep. 2021.

[32] A. De Domenico, Y.-F. Liu, and W. Yu, "Optimal virtual network function deployment for 5G network slicing in a hybrid cloud infrastructure," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7942–7956, Dec. 2020.

[33] Z. Zhang, G. Wu, and H. Ren, "Multi-attribute-based QoS-aware virtual network function placement and service chaining algorithms in smart cities," *Comput. Electr. Eng.*, vol. 96, Dec. 2021, Art. no. 107465.

[34] N. Kiran, X. Liu, S. Wang, and C. Yin, "VNF placement and resource allocation in SDN/NFV-enabled MEC networks," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–6.

[35] M. Golkarifard, C. F. Chiasserini, F. Malandrino, and A. Movaghar, "Dynamic VNF placement, resource allocation and traffic routing in 5G," *Comput. Netw.*, vol. 188, Apr. 2021, Art. no. 107830.

[36] Y. Mu, L. Wang, and J. Zhao, "Energy-efficient and interference-aware VNF placement with deep reinforcement learning," in *Proc. IFIP Netw. Conf. (IFIP Networking)*, Jun. 2021, pp. 1–9.

[37] G. Kibalya, J. Serrat, J.-L. Gorricho, D. G. Bujjingo, J. Sserugunda, and P. Zhang, "A reinforcement learning approach for placement of stateful virtualized network functions," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, May 2021, pp. 672–676.

[38] S. M. A. Araújo, F. S. H. de Souza, and G. R. Mateus, "A hybrid optimization-machine learning approach for the VNF placement and chaining problem," *Comput. Netw.*, vol. 199, Nov. 2021, Art. no. 108474.

[39] J. Pei, P. Hong, M. Pan, J. Liu, and J. Zhou, "Optimal VNF placement via deep reinforcement learning in SDN/NFV-enabled networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 2, pp. 263–278, Feb. 2020.

[40] K. Ali and M. Jammal, "Proactive VNF scaling and placement in 5G O-RAN using ML," *IEEE Trans. Netw. Service Manage.*, vol. 21, no. 1, pp. 174–186, Sep. 2024.

[41] *Continuing Our Investment in Nevada*. Accessed: Mar. 23, 2024. [Online]. Available: https://www.tesla.com/blog/continuing-our-investment-nevada

[42] 3Gpp. *5G System Overview*. Accessed: May 8, 2024. [Online]. Available: https://www.3gpp.org/technologies/5g-system-overview

[43] NGMN-Alliance, "5G white paper 2," NGMN Alliance, Düsseldorf, Germany, White Paper NGMN 5G White Paper 2, Jul. 2020. [Online]. Available: https://www.ngmn.org/publications/5g-white-paper-2.html

[44] S. Wang, Y. Guo, N. Zhang, P. Yang, A. Zhou, and X. Shen, "Delay-aware microservice coordination in mobile edge computing: A reinforcement learning approach," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 939–951, Mar. 2021.

[45] J. Coffey, "Latency in optical fiber systems," COMMSCOPE, Claremont, NC, USA, White Paper No. WP-111432-EN (02/17), 2017. [Online]. Available: https://www.commscope.com/globalassets/digizuite/2799-latency-in-optical-fiber-systems-wp-111432-en.pdf

[46] Q. Ye, W. Zhuang, X. Li, and J. Rao, "End-to-end delay modeling for embedded VNF chains in 5G core networks," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 692–704, Feb. 2019.

[47] 3GPP, "3rd generation partnership project; technical specification group services and system aspects; study on communication for automation in vertical domains (release 16)," Tech. Rep. TR 22.804, v16.3.0, Jul. 2020.

[48] K. Atefi, S. Yahya, A. Rezaei, and A. Erfanian, "Traffic behavior of local area network based on M/M/1 queuing model using Poisson and exponential distribution," in *Proc. IEEE Region 10 Symp. (TENSYMP)*, May 2016, pp. 19–23.

[49] M. A. Siddiqi, H. Yu, and J. Joung, "5G ultra-reliable low-latency communication implementation challenges and operational issues with IoT devices," *Electronics*, vol. 8, no. 9, p. 981, Sep. 2019.

[50] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, "5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges," *Comput. Netw.*, vol. 167, Feb. 2020, Art. no. 106984.

[51] A. Laghrissi and T. Taleb, "A survey on the placement of virtual resources and virtual network functions," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1409–1434, 2nd Quart., 2019.

[52] M. T. Kawser and Z. N. Ahmed, "Dependency of indoor delay spread and path loss on 5G and 6G frequencies and its simple formulation," in *Proc. Future Technol. Conf.* Vancouver, BC, Canada: Springer, 2022, pp. 92–105. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-18458-1_7

[53] S. Ramanathan, K. Kondepu, M. Razo, M. Tacca, L. Valcarenghi, and A. Fumagalli, "Live migration of virtual machine and container based mobile core network components: A comprehensive study," *IEEE Access*, vol. 9, pp. 105082–105100, 2021.

[54] J. Osa, N. Björsell, I. Val, and M. Mendicute, "Measurement based stochastic channel model for 60 GHz mmWave industrial communications," *IEEE Open J. Ind. Electron. Soc.*, vol. 4, pp. 603–617, 2023.

[55] J. Speiran and E. M. Shakshuki, "Understanding the effect of physical parameters on packet loss in veins VANET simulator," *Procedia Comput. Sci.*, vol. 201, pp. 359–367, 2022.

[56] W. Uther, *Markov Decision Processes*. Boston, MA, MA: Springer, 2010, pp. 642–646.

[57] A. Othman, N. A. Nayan, and S. N. H. S. Abdullah, "Automated deployment of virtual network function in 5G network slicing using deep reinforcement learning," *IEEE Access*, vol. 10, pp. 61065–61079, 2022.

[58] Y. Chen and M. Wang, "Lower bound on the computational complexity of discounted Markov decision problems," 2017, *arXiv:1705.07312*.

[59] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[60] H. V. Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. 30th Conf. Artif. Intell. (AAAI)*, Mar. 2016, pp. 2094–2100.

[61] J. Tan and W. Guan, "Resource allocation of fog radio access network based on deep reinforcement learning," *Eng. Rep.*, vol. 4, no. 5, May 2022, Art. no. e12497.

NAUMAN SAQIB received the B.Sc. degree in electrical engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2009, and the M.Sc. degree in communication/networks from the University of Engineering and Technology (UET), Taxila, Pakistan, in 2013. He is currently pursuing the Ph.D. degree with the Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia (UKM).

Since 2012, he has been with the Center of Excellence for Science and Technology, Pakistan, as a Research and Development Engineer with major interests in the field of wireless communications, signal processing, and artificial intelligence.

NOR FADZILAH ABDULLAH (Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from Universiti Teknologi Malaysia, in 2001, the M.Sc. degree (Hons.) in communications engineering from The University of Manchester, U.K., in 2003, and the Ph.D. degree in electrical and electronic engineering from the University of Bristol, U.K., in 2012. She is currently an Associate Professor with Universiti Kebangsaan Malaysia, Selangor, Malaysia. Her research interests include 5G, millimeter wave, LTE-A, vehicular networks, massive MIMO, space-time coding, fountain codes, and channel propagation modeling and estimation.

ASMA ABU-SAMAH (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical, electronics, automation systems and signal processing from Universite de Joseph Fourier, Grenoble, France, in 2008 and 2010, respectively, and the Ph.D. degree in automated control and production systems from Université Grenoble Alpes, France, in 2016. She is currently a Senior Lecturer with Universiti Kebangsaan Malaysia (UKM). Her main research interests include the application of machine learning in wireless communications and the Internet of Things (IoT).

HAIDER A. H. ALOBAIDY (Member, IEEE) received the M.Sc. degree from Al-Mustansiriyah University, Iraq, in 2016, and the Ph.D. degree in electrical, electronics, and system engineering from UKM, in 2022. He is currently a Postdoctoral Researcher with the Department of Electrical, Electronic and Systems Engineering, UKM. He is also a Lecturer with the Department of Information and Communications Engineering, College of Information Engineering, Al-Nahrain University, Baghdad, Iraq. His research interests include wireless communication, wireless sensor networks, the IoT, machine learning, and channel propagation modeling and estimation.

ROSDIADEE NORDIN (Senior Member, IEEE) received the B.Eng. degree from Universiti Kebangsaan Malaysia (UKM), in 2001, and the Ph.D. degree from the University of Bristol, U.K., in 2011. He is currently a Professor with the Department of Engineering, School of Engineering and Technology, Sunway University, Malaysia. His research interests include beyond 5G wireless communications, specifically focusing on advanced wireless transmission techniques and channel modeling, aerial wireless communications, and wireless communications for the Internet of Things applications. He was a recipient of the Leadership in Innovation Fellowship, the Technopreneur Program under the Royal Academy of Engineering, U.K., in 2021, the Top Research Scientists Malaysia (TRSM), and the Prestigious Award under the Academy of Science Malaysia, in 2020.