

A Semantic-Aware Multiple Access Scheme for Distributed, Dynamic 6G-Based Applications

Hamidreza Mazandarani¹, Masoud Shokrnezhad², and Tarik Taleb^{1,2}

¹ Ruhr University Bochum, Bochum, Germany; hr.mazandarani@ieee.org, tarik.taleb@rub.de

² Oulu University, Oulu, Finland; {masoud.shokrnezhad, tarik.taleb}@oulu.fi

Abstract—The emergence of the semantic-aware paradigm presents opportunities for innovative services, especially in the context of 6G-based applications. Although significant progress has been made in semantic extraction techniques, the incorporation of semantic information into resource allocation decision-making is still in its early stages, lacking consideration of the requirements and characteristics of future systems. In response, this paper introduces a novel formulation for the problem of multiple access to the wireless spectrum. It aims to optimize the utilization-fairness trade-off, using the α -fairness metric, while accounting for user data correlation by introducing the concepts of self- and assisted throughputs. Initially, the problem is analyzed to identify its optimal solution. Subsequently, a Semantic-Aware Multi-Agent Double and Dueling Deep Q-Learning (SAMA-D3QL) technique is proposed. This method is grounded in Model-free Multi-Agent Deep Reinforcement Learning (MADRL), enabling the user equipment to autonomously make decisions regarding wireless spectrum access based solely on their local individual observations. The efficiency of the proposed technique is evaluated through two scenarios: single-channel and multi-channel. The findings illustrate that, across a spectrum of α values, association matrices, and channels, SAMA-D3QL consistently outperforms alternative approaches. This establishes it as a promising candidate for facilitating the realization of future federated, dynamically evolving applications.

Index Terms—6G, Semantic-awareness, Resource Allocation, Multiple Access, Medium Access Control (MAC), Wireless Spectrum, Utilization, Fairness, Throughput, Deep Q-Learning, Reinforcement Learning, Dynamic, Distributed.

I. INTRODUCTION

Communication systems are undergoing a transformation from the traditional *bit-oriented* model, focused solely on transmitting data bits, to a *semantic-aware* paradigm. In this new approach, the semantic value of various system components (such as messages, users, and resources) is quantified and subsequently integrated into the service provisioning process [1]. This paradigm shift opens up significant opportunities for innovative services, particularly in the context of the 6G-based Metaverse, where seamless integration of physical and virtual environments is a key goal [2]. To support this transition, versatile data segmentation methods like the Segment Anything Model (SAM) [3] and its lightweight version FastSAM [4] play a pivotal role in capturing meaningful understanding from user data. For instance, consider a scenario involving a group of nearby users engaged in a holographic meeting with a partially shared background and environmental sounds. Through data segmentation and semantic extraction, it becomes feasible to improve network efficiency and effectiveness



Fig. 1. An illustration of similar segments.

and accommodate a greater number of users and requests with stringent end-to-end requirements. Fig. 1 demonstrates this concept, showcasing segmented images of a desk from different viewpoints, revealing shared similarities across multiple segments.

The incorporation of semantics into decision-making processes at various network layers is still in its nascent stages of development. Our knowledge management framework for semantic-aware systems, known as KB MANO [1], in conjunction with the introduction of the knowledge space concept within the Meta-Networking framework [5] and the reasoning plane discussed in [6], represents some of the initial endeavors to conceptualize network-wide semantic awareness. Particularly, in the data link layer, a semantic-aware Medium Access Control (MAC) protocol can optimize the utilization of limited radio resources by eliminating redundant data transmission. The Model Division Multiple Access (MDMA) is an example of such an approach [7], dedicating a portion of resources to transmit shared information among users. In a similar vein, Kassab *et al.* [8] conducted research focusing on a scenario involving a collective of Internet of Things (IoT) devices engaged in the concurrent monitoring of correlated on-off events, such as variations in temperature levels. Another line of research, exemplified by DeepSC-IR/MT [9] and DeepMA [10], explores the joint semantic encoding of data from multiple users (i.e., multi-user semantic communications). While these methods are effective in specific contexts, they do not account for the distributed and dynamic nature of future applications, making them less scalable and potentially unsuitable for 6G systems.

When dealing with an environment containing ever-changing multiple distributed actors, Multi-Agent Deep Reinforcement Learning (MADRL) stands out as an effective approach. Its adaptability to handle non-stationarity, stemming

from the co-evolution of agents' policies, makes it a compelling choice [11]. In MADRL, agents base their transmission decisions on historical observations, such as success or failure in task fulfillment, while coordination among agents occurs through policy sharing, message passing, or centralized training. MADRL-based multiple access control has been discussed in various papers. For instance, Naparstek *et al.* [12] proposed an efficient Nash equilibrium for multi-agent MAC where all agents adopt the same policy. Sohaib *et al.* [13] extended MADRL to MAC by using transmission policies over consecutive time slots as decision variables (as opposed to per-time-slot decision-making), ensuring short-term fairness under varying user conditions. Guo *et al.* [14] introduced the QMIX-advanced LBT (QLBT) MAC protocol, building upon the QMIX algorithm for MADRL problems [15]. QLBT incorporates two reward signals, one for maximizing network utilization and another for maintaining fairness. However, despite these advancements, none of the examined MADRL-based MAC protocols explicitly address the growing need for semantic awareness, a critical aspect of evolving communication systems.

In response to the existing gap in the current literature, we propose a novel formulation for the multiple-access problem to optimize the trade-off between utilization and fairness, taking into account the inherent correlation in users' data. This departure from the conventional bit-oriented paradigm underscores our paradigm shift towards semantic awareness. Notably, we demonstrate the intractable nature of this problem and present a solution approach employing model-free MADRL. Our contribution, the Semantic-Aware Multi-Agent Double and Dueling Deep Q-Learning (SAMA-D3QL) method, expanding upon the principles established in our prior research [16]–[19]. SAMA-D3QL operates under the premise that users possess dynamic associations with various semantic segments, some of which may be shared among them. Consequently, when a user transmits a packet, it benefits all users with whom she shares pertinent segments.

The subsequent sections of this paper are organized as follows. Section II introduces the system model and provides a comprehensive problem formulation. The optimal problem solution and the details of the SAMA-D3QL technique are elucidated in Section III. Section IV presents the numerical results, while Section V concludes our work with closing remarks.

II. PROBLEM FORMULATION

In this study, a configuration comprising a Small Base Station (SBS) and a collection of \mathcal{N} intelligent User Equipment (UE) is examined, each uniquely labeled as u_i , where $i \in \mathbb{N} = \{1, \dots, \mathcal{N}\}$. These UEs are engaged in contention for access to \mathcal{C} perfectly time-slotted communication channels designated for transmitting their data to the SBS. It is essential to note that simultaneous transmissions over a single channel inevitably lead to collision. Consequently, in the conventional bit-oriented framework, the highest achievable network utilization is capped at \mathcal{C} packets per time slot.

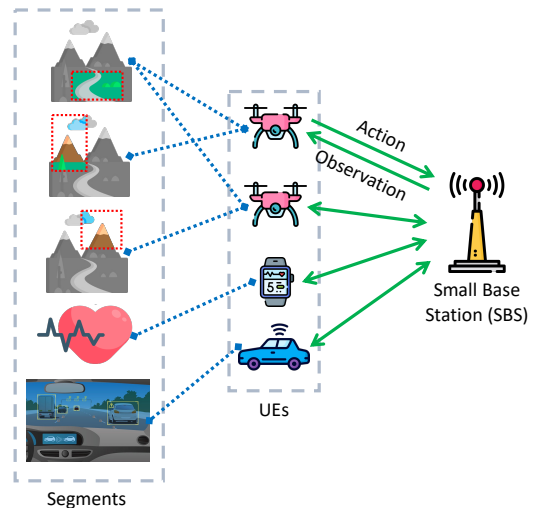


Fig. 2. A sample network consisting of four UEs ($\mathcal{N} = 4$) and five segments ($\mathcal{K} = 5$). As illustrated, one separate segment is associated with each user, in addition to one shared segment among the first and second UEs.

It is presumed that within the network, there exists shared semantic information among subsets of UEs. For instance, in a holographic meeting, nearby UEs may share some environmental data while retaining their individual face and voice data. This scenario is represented by \mathcal{K} distinct semantic entities, termed *segments*, denoted as $k \in \mathbb{K} = \{1, \dots, \mathcal{K}\}$. At each time slot $t \in \{0, \dots, \mathcal{T}\}$, each UE i is associated/disassociated with a segment k , signified by a binary variable $a_{i,k}^t \in \{0, 1\}$, forming a predefined binary association matrix $\mathbb{A} = [a_{i,k}^t]_{\mathcal{N} \times \mathcal{K} \times \mathcal{T}}$. The matrix serves as the input to the system, and an orchestration entity, such as KB MANO as discussed by Shokrnezhad *et al.* [1], is responsible for generating it. In an illustrative example, shown in Fig. 2, the association matrix at time slot t is given as $\mathbb{A}^t = [[1, 1, 0, 0, 0], [1, 0, 1, 0, 0], [0, 0, 0, 1, 0], [0, 0, 0, 0, 1]]$.

To evaluate the semantic efficiency, we introduce a segment transmission indicator, denoted as y_k^t . This indicator operates under the assumption that, at any given time slot, all UE-associated segments are combined into a single packet. Furthermore, receiving a specific segment by the SBS multiple times within a time slot does not enhance throughput. y_k^t is defined as:

$$y_k^t \triangleq \min \left\{ 1, \sum_{j \in \mathbb{N}} a_{j,k}^t \cdot z_j^t \right\}, \quad (1)$$

where $z_j^t \in \{0, 1\}$ signifies whether UE j has successfully transmitted a packet during time slot t . In Appendix A, a more generalized form is presented that does not rely on the assumption that receiving multiple copies of a segment by the SBS is unproductive.

Based on (1), the normalized throughput of each UE can be determined. Considering that $\mathbb{W}^t = \{t - \mathcal{H}, \dots, t\}$ represents a sliding time window of length \mathcal{H} , and $\Omega_i^t = 1 / (\sum_{k, t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t)$ acts as the normalization factor, the normalized throughput of u_i is as follows:

$$x_i^t \triangleq \Omega_i^t \cdot \sum_{k, t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t \cdot y_k^t \quad (2)$$

By isolating u_i from other UEs in (2), we can break down x_i^t into two components: *self throughput* (\hat{x}_i^t) and *assisted throughput* (\ddot{x}_i^t), defined in (3). The assisted throughput for UE i represents the contribution of other UEs in transmitting shared segments between them and u_i to the SBS.

$$\begin{aligned} x_i^t &= \hat{x}_i^t + \ddot{x}_i^t \\ \hat{x}_i^t &= \Omega_i^t \cdot \sum_{k,i,t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t \cdot z_i^t \\ \ddot{x}_i^t &= \Omega_i^t \cdot \sum_{k,i,t \in \mathbb{K}, \mathbb{W}^t} \left(a_{i,k}^t \cdot (1 - z_i^t) \cdot y_k^{t,-i} \right) \end{aligned} \quad (3)$$

Here, $y_k^{t,-i} = \min \{1, \sum_{j \in \mathbb{N} \setminus \{i\}} a_{j,k}^t \cdot z_j^t\}$.

The problem can now be formally defined. Our objective is to maximize network utilization across the network's lifetime while maintaining fairness among UEs. We adopt the α -fairness metric as the objective function, as expressed in C1. With $\alpha = 0$ and $\alpha = \infty$ as two extreme cases: sum throughput maximization and max-min fairness maximization, respectively. In our time-slotted multi-channel network configuration, the successful transmission of u_i at time slot t (i.e., z_i^t) is contingent upon u_i transmitting a packet on a channel that all other UEs are not utilizing. C2 and C3 formalize this condition, where $m_{i,c}^t$ denotes whether u_i is transmitting on channel c at time slot t . C2 ensures that each UE transmits on at most one channel during each time slot, and C3 guarantees that z_i^t is equal to 1 if u_i selects a channel that is not already chosen by other users. Consequently, the transmission matrix $\mathbb{M} = [m_{i,c}^t]_{\mathcal{N} \times \mathcal{C} \times \mathcal{T}}$ determines $\mathbb{Z} = [z_i^t]_{\mathcal{N} \times \mathcal{T}}$, which influences UE throughputs $\mathbb{X} = [x_i^t]_{\mathcal{N} \times \mathcal{T}}$ as per (3). Finally, we calculate $\mathbf{U}_\alpha(\mathbb{X})$. In symbolic terms, we have a causal relationship denoted as $\mathbb{M} \rightarrow \mathbb{Z} \rightarrow \mathbb{X} \rightarrow \mathbf{U}_\alpha(\mathbb{X})$, where the arrow indicates the causal flow.

$$\max_{\mathbb{M}} \sum_{t \in \{0, \dots, \mathcal{T}\}} \mathbf{U}_\alpha(\mathbb{X}^t) \quad \text{s.t.} \quad (4)$$

$$\mathbf{U}_\alpha(\mathbb{X}^t) = \begin{cases} \sum_{i \in \mathbb{N}} \log(x_i^t) & \text{if } \alpha = 1 \\ (1 - \alpha)^{-1} \cdot \sum_{i \in \mathbb{N}} (x_i^t)^{1-\alpha} & \text{if } \alpha \neq 1 \end{cases} \quad (\text{C1})$$

$$\sum_{c \in \{0, \dots, \mathcal{C}\}} m_{i,c}^t \leq 1 \quad (\text{C2})$$

$$z_i^t = \sum_{c \in \{0, \dots, \mathcal{C}\}} m_{i,c}^t \cdot \prod_{j \neq i} (1 - m_{j,c}^t) \quad (\text{C3})$$

$$x_i^t = \hat{x}_i^t + \ddot{x}_i^t \quad \text{according to (3)} \quad (\text{C4})$$

Note that C2, C3, and C4 are applicable to all values of $i \in \mathbb{N}$ and $t \in \{0, \dots, \mathcal{T}\}$.

III. PROPOSED METHOD

A. Optimal Solutions

In a semantic-oblivious system (where each UE has its segments and no assisted throughputs exist), it is proved that the optimal point for (4) is attained when all UEs receive the same channel allocation ($x_i^t = \min\{\mathcal{C}/\mathcal{N}, 1\}$ for all i and t)

TABLE I
THE OPTIMAL ALLOCATIONS FOR THE INSTANCE DEPICTED IN FIG. 2.

α	Self Throughputs	Normalized Throughputs
0	[0.5, 0.5, 0, 0]	[0.75, 0.75, 0, 0]
0.5	[0.3, 0.3, 0.2, 0.2]	[0.45, 0.45, 0.2, 0.2]
1	[0.25, 0.25, 0.25, 0.25]	[0.375, 0.375, 0.25, 0.25]
∞	[0.2, 0.2, 0.3, 0.3]	[0.3, 0.3, 0.3, 0.3]

[12]. However, in a semantic-aware system, due to the interdependency of UEs' throughputs, achieving optimality is not as straightforward as allocating resources equally. Returning to the modest example of Fig. 2, optimal allocations for different α values are provided in Table I based on an exhaustive search over the solution space (assuming $\mathcal{C} = 1$ and stationarity over time). The first row, representing no fairness ($\alpha = 0$), allocates the entire channel to the first two UEs who share each other's information, resulting in an optimal objective value of 1.5, a 50% improvement over the bit-oriented scenario. As α increases, the last two UEs gain a larger share of the channel as there is no assisted throughput for them.

In more intricate and realistic settings, characterized by an increased number of UEs, channels, or segments, the optimization of the nonlinear problem presented in (4) becomes challenging, if not infeasible. Furthermore, it is essential to acknowledge that, in real-world scenarios, we do not have access to the information of future time slots. Additionally, in distributed scenarios where UEs are unaware of each other's decisions or associations, we face a fundamental information gap that hinders optimal problem-solving. Consequently, we turn to model-free MADRL approaches, which will be elaborated upon in the following subsection.

B. Learning Mechanism

In what follows, we will elucidate the fundamental components of our MADRL-based approach, denoted as Semantic-Aware Multi-Agent Double and Dueling Deep Q-Learning (SAMA-D3QL). This exposition will encompass discussions on action and state spaces, reward mechanisms, and the architectural framework of the learning process.

1) *Action Space*: The action space for all UEs is specified in (5). In this set, the action of u_i can take one of two forms: $\rho_i : (0, c)$, representing the act of sensing channel c , or $\rho_i : (1, c)$, signifying the transmission of a packet on channel c .

$$\rho = \left\{ \left\{ \rho_i : (\phi, c) \mid \phi \in \{0, 1\}, c \in \{1, \dots, \mathcal{C}\} \right\} \mid i \in \mathbb{N} \right\} \quad (5)$$

2) *State Space*: The state space consists of four distinct components. To provide UEs with real-time fairness information, we incorporate the *delay to last successful transmission* (D2LT) metric [14] represented as v_i . This metric signifies the number of time slots elapsed since the last successful transmission by u_i . Specifically, we include the normalized D2LT (i.e., $\bar{v}_i = v_i / \sum_{j \in \mathbb{N}} v_j$) in the state space of UEs. The following component relates to the UE's observation. When sensing channel c , the observation set for each UE is denoted as $\dot{o}_i = \{B: \text{Busy}, I: \text{Idle}\}$, while in the case of packet transmission, it is $\dot{o}_i = \{S: \text{Success}, C: \text{Collision}\}$.

Furthermore, we introduce another indicator called "assisted transmission," denoted as $\ddot{o}_i \in [0, 1]$. At each time slot, the assisted transmission of u_i is simply the ratio of its segments transmitted by other UEs. Given that the final element within the state space pertains to UE actions, the state of each UE comprises the most recent \mathcal{H} (D2LT, observation, assisted transmission, action) tuples. Thus, the global system state at time slot t is defined as follows:

$$\mathbf{S}^t = \left\{ s_i^t : \left\{ (\bar{v}_i^t, \dot{o}_i^t, \ddot{o}_i^t, \rho_i^t) \mid t \in \mathbb{W}^t \right\} \mid i \in \mathbb{N} \right\} \quad (6)$$

3) *Reward*: Since we consider a cooperative scenario in which all UEs aim to maximize α -fairness of the network, the system reward is empirically defined as the weighted average of successful transmissions, as outlined in (7). These weights are determined by the D2LT values, encouraging UEs that have experienced longer delays to access the channel.

$$\mathbf{R}^t = \frac{1}{\mathcal{N}} \cdot \sum_{i \in \mathbb{N}} \bar{v}_i^t \cdot \mathbb{1}(\dot{o}_i^t = S) \quad (7)$$

4) *Training Process*: Our proposed methodology is founded on the Value Decomposition Network (VDN) framework [20], which adheres to the Centralized Training and Decentralized Execution (CTDE) paradigm. This strategy operates on the premise that during the training phase, full access to the global system state is available, while user decisions are made based on their local information. Similarly, the training process of our proposed method is centralized within the SBS, where UE policies are being trained. Given that the SBS possesses knowledge of UE experiences, including their actions and observations, there is no need for agents to transmit their experiences to the SBS. However, it is assumed that the SBS periodically updates and disseminates policies to the UEs via high-bandwidth and low-latency dedicated communication channels. UEs autonomously select their actions based on their respective policies and individual observations without relying on the actions or states of other UEs. Utilizing Q-learning, the summation of individual user Q-values is employed to compute the total Q-value within VDN, precisely calculated using the following equation:

$$Q_{tot}(\mathbf{S}^t, \boldsymbol{\rho}^t) = \sum_{i \in \mathbb{N}} Q_i(s_i^t, \rho_i^t) \quad (8)$$

With the above definition, the policy parameter update function for all agents, represented as $\mathcal{W} = \{\mathcal{W}_i \mid i \in \mathbb{N}\}$, can be expressed as follows:

$$\begin{aligned} \mathcal{W}^{t+1} = & \mathcal{W}^t + \\ & \sigma[Y_{D3QL}^t - Q_{tot}(\mathbf{S}_t, \boldsymbol{\rho}_t; \mathcal{W}^t)] \nabla_{\mathcal{W}^t} \cdot \\ & Q_{tot}(\mathbf{S}_t, \boldsymbol{\rho}_t; \mathcal{W}^t), \end{aligned} \quad (9)$$

where the target value (Y_{D3QL}^t) is determined using the D3QL algorithm, which is an extension of the original DQL algorithm that incorporates both dueling and double mechanisms [16]. Our approach is outlined in Algorithm 1, wherein lines 4-15 pertain to the decentralized execution of actions by UEs, while lines 17-20 depict the centralized training conducted at the SBS.

Algorithm 1: SAMA-D3QL

Input: \mathcal{T} , ϵ' , and $\tilde{\epsilon}$

```

1  $\mathcal{W} \leftarrow \mathbf{0}$ ,  $\mathcal{W}^- \leftarrow \mathbf{0}$ ,  $\epsilon \leftarrow 1$ ,  $memory \leftarrow \{\}$ 
2 foreach  $t$  in  $\{0, \dots, \mathcal{T}\}$  do
3   * Decentralized Execution (4-15)
4   foreach  $i$  in  $\mathbb{N}$  do
5      $\zeta \leftarrow$  generate a random number from  $[0 : 1]$ 
6     if  $\zeta > \epsilon$  then
7        $(\phi, c) \leftarrow \operatorname{argmax}_{\rho \in \rho_i} Q(s_i^t, \rho, \mathcal{W}_i)$ 
8     else
9       select a random  $(\phi, c)$  from  $\rho_i$ 
10    if  $\phi = 1$  then
11      transmit packet on channel  $c$ 
12    else
13      sense channel  $c$ 
14      receive  $\dot{o}_i^t$  and  $\ddot{o}_i^t$  from the SBS
15      construct  $s_i^{t+1}$ 
16  * Centralized Training (17-20)
17  calculate  $\mathbf{R}^t$  according to (7)
18   $memory \leftarrow \{\mathbf{R}^t\} \cup \{(s_i^t, (\phi, c), s_i^{t+1}) \mid i \in \mathbb{N}\}$ 
19  choose a batch of samples from  $memory$ 
20  train the agent according to (9)
21  if  $\epsilon > \tilde{\epsilon}$  then
22     $\epsilon \leftarrow \epsilon \cdot \epsilon'$ 

```

IV. EVALUATION

In this section, a series of experiments is conducted to perform a numerical evaluation of SAMA-D3QL. Each experiment involves a comparison between SAMA-D3QL and its semantic-oblivious counterpart (in terms of both network and algorithmic design), referred to as MA-D3QL. Specifically, MA-D3QL does not incorporate assisted transmission into its decision-making process and operates under the assumption that the SBS lacks awareness of the correlation among UEs' source data, thereby decoding only the self-throughputs. Additionally, we compare the results with those obtained using random agents in the semantic-aware network, serving as a benchmark. To facilitate algorithm comparison, we vary the alpha value in the α -fairness function, considering four settings: 0, 0.5, 1, and 2, representing different degrees of fairness. In the first scenario, we establish two modest association matrices, for which optimal values can be extracted, to delve into the effectiveness of our strategy in detail. Subsequently, in the second scenario, we assess the performance of our scheme in a multi-channel setting involving a higher number of nodes and segments. Table II details the simulation's hyperparameters and configurations. It is noteworthy that we have opted for LSTM networks as the feature extractor, owing to their capability to process sequential data.

A. Scenario 1: Single-Channel Setting

In this scenario, we conduct a comparative analysis of the algorithms using two predefined UE-segment association matrices. The first matrix (Fig. 3-A) is the one depicted in Fig. 2

TABLE II
TRAINING CONFIGURATIONS.

Parameter	Value
State size (\mathcal{H})	4 experiences
Capacity of experience memory	500 experiences
Batch size	32
Discount Factor (γ)	0.9
Learning rate	0.001
Exploration parameters $\tilde{\epsilon}, \epsilon'$	0.005, 0.995
Approximator model	LSTM with 64 units + fully connected with 64 and 32 units
Target network update frequency	Every 50 steps

with optimal points stated in Table. I. The second matrix (Fig. 3-B) associates two segments with each pair of four users, represented as $A = [[1, 1, 0, 0], [1, 1, 0, 0], [0, 0, 1, 1], [0, 0, 1, 1]]$. Fig. 3 illustrates the evolution of objective functions over time on the left side and the average UE throughputs over all-time on the right side. The lighter portion of each UE throughput, if present, indicates the assisted throughput of that UE. Across different values of α , SAMA-D3QL consistently and significantly outperforms MA-D3QL in both network configurations converging to the optimal solution (as denoted by the dotted line). It's worth noting that in the single-channel setup, optimal values are obtained through exhaustive searches over self-throughputs. However, this calculation becomes infeasible in the multi-channel network, where UEs sharing segments may transmit simultaneously.

B. Scenario 2: Multi-channel Setting

In this scenario, we have conducted three experiments with the objective of assessing the impact of fairness levels (quantified by the parameter α) and the sparsity of association matrices on algorithm effectiveness. The experimental setup entails six UEs contending for access to three channels, with varying trade-offs between utilization and fairness. In addition, three different association matrices, each characterized by a distinct level of sparsity, are utilized. The results depicted in Fig. 4 highlight the superior performance of SAMA-D3QL agents across all experiments. However, the extent of this performance advantage varies depending on the specific association matrix and the chosen fairness level. Intriguingly, random agents achieve comparable results to MA-D3QL agents, underscoring the potential of semantic awareness in enhancing the efficiency and efficacy of communication systems. Notably, in this particular scenario with multiple channels, random agents exhibit better performance compared to the previous scenario, likely due to the increased channel diversity. Lastly, the enhanced performance of SAMA-D3QL is particularly notable in scenarios with denser association matrices, such as in topology B. This implies that semantic-aware multiple access could provide increased advantages in situations with more correlated traffic.

V. CONCLUSION

This paper introduced a novel semantic-aware problem formulation and algorithmic approach for addressing the challenges of multiple access in distributed and dynamic 6G-based applications. The core objective of this research was to strike a

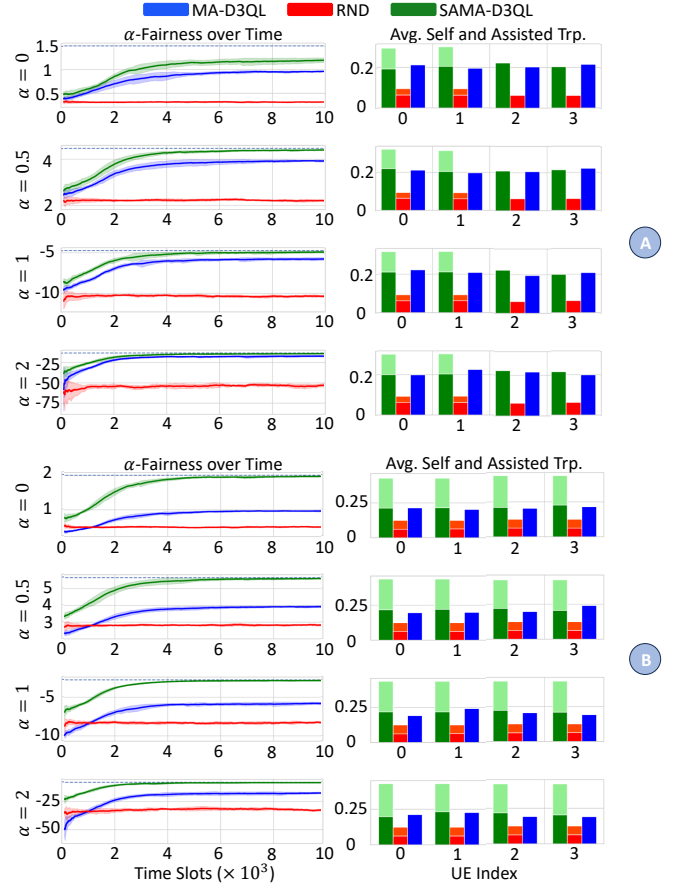


Fig. 3. Objective functions over time and all-time average UE throughputs in the first scenario, encompassing experiments A and B, comparing the SAMA-D3QL, MA-D3QL, and RND (random agents) algorithms. The reported results are based on averaging the outcomes of 10 rounds of simulations. It's noteworthy that, when present, the lighter segment of each UE throughput signifies the assisted throughput for that particular UE.

balance between maximizing network utilization and ensuring fairness, all while considering the inherent data correlations among users. To achieve this goal, we devised a structured and adaptable formal framework that leverages the semantic relationships between users. Additionally, we introduced the SAMA-D3QL algorithm, which demonstrates a remarkable capacity to significantly enhance the α -fairness of the system compared to semantic-oblivious approaches. Importantly, this approach relies on decentralized multi-agent Q-learning, enabling UEs to make autonomous decisions regarding wireless spectrum access without requiring global knowledge of the system. This advancement holds the potential to bring future federated, dynamically evolving applications one step closer to real-world deployment.

Our prospective future endeavors will involve expanding this framework to encompass more realistic and practical scenarios, including those with time-varying and non-binary UE-segment association matrices, as well as variable packet-size scenarios. Implementing more advanced MADRL algorithms with reduced communication overhead would also enhance the practicality of our proposed scheme. In addition, it would be advantageous to assess the framework's performance in real-

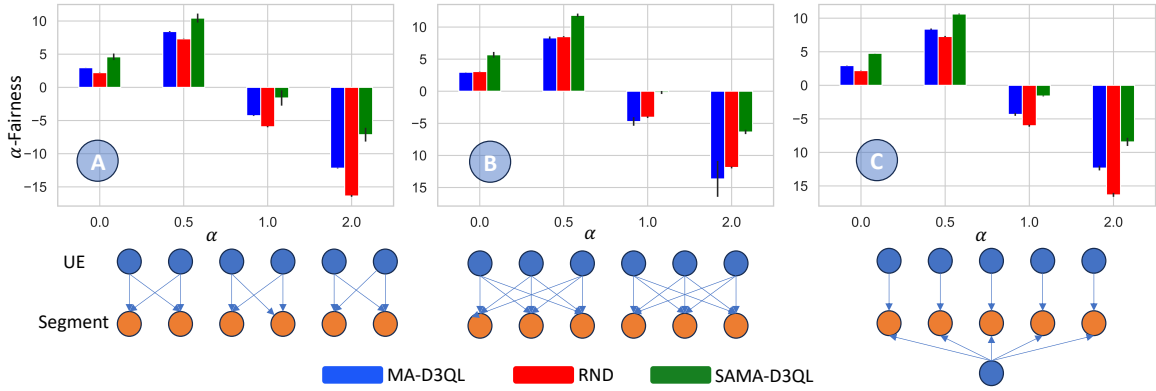


Fig. 4. The average value of objective functions for the last 1000 time slots (of 10k time slots) in experiments A, B, and C. The UE-segment association matrix for each experiment is illustrated. Results are the average of 10 rounds of simulations.

world settings where approximate semantic correlations are prevalent.

APPENDIX A GENERAL SETTING OF THE PROBLEM

Throughout this paper, we have operated under the assumption that the reception of multiple duplicates of a segment by the SBS does not enhance throughput. As a result, we have employed the minimum function in (1) and subsequently dissected user throughput into self and assisted components. In a more generalized context, where the association matrix \mathbb{A} is non-binary or the channel exhibits imperfections, we must reformulate (1) in the following manner:

$$y_k^t \triangleq \mathbf{g}\left(\{a_{j,k}^t \cdot z_j^t \mid j \in \mathbb{N}\}\right), \quad (10)$$

where $\mathbf{g}(\cdot)$ represents a function that determines the successful reception of a segment by the SBS based on all received signals from the UEs. Consequently, we should employ (2) (instead of (3)) to formulate C4 in (4).

ACKNOWLEDGMENT

This research work was also conducted at ICTFICIAL Oy. It is partially supported by the European Union's Horizon 2020 Research and Innovation Program through the aerOS project under Grant No. 101069732; the Business Finland 6Bridge 6Core project under Grant No. 8410/31/2022; the European Union's HE research and innovation programme HORIZON-JUSNS-2022 under the 6GSandbox project (Grant No. 101096328); and and the Research Council of Finland 6G Flagship Programme under Grant No. 346208. The paper reflects only the authors' views. The Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

- [1] M. Shokrnezhad, H. Mazandarani, and T. Taleb, "Semantic Revolution from Communications to Orchestration for 6G: Challenges, Possibilities, and Research Directions," in *IEEE Network*, 2023, (Under Review).
- [2] H. Yu, M. Shokrnezhad *et al.*, "Towards 6G-Based Metaverse: Supporting Highly-Dynamic Deterministic Multi-User Extended Reality Services," in *IEEE Network*. IEEE, 2023.
- [3] A. Kirillov, E. Mintun *et al.*, "Segment Anything," *arXiv preprint arXiv:2304.02643*, 2023.
- [4] X. Zhao, W. Ding *et al.*, "Fast Segment Anything," *arXiv preprint arXiv:2306.12156*, 2023.
- [5] Y. Lin, C. Wu *et al.*, "Meta-Networking: Beyond the Shannon Limit with Multi-faceted Information, Semantic Communication, IoV, Multi-Faceted Information," *IEEE Network*, vol. 37, no. 4, pp. 256–264, 2023.
- [6] C. Chaccour, W. Saad *et al.*, "Less Data, More Knowledge: Building Next Generation Semantic Communication Networks," *arXiv preprint arXiv:2211.14343*, 2022.
- [7] P. Zhang, X. Xu *et al.*, "Model Division Multiple Access for Semantic Communications," *Frontiers of Information Technology & Electronic Engineering*, pp. 1–12, 2023.
- [8] R. Kassab, A. Destounis *et al.*, "Multi-Agent Deep Stochastic Policy Gradient for Event Based Dynamic Spectrum Access," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2020, pp. 1–6.
- [9] H. Xie, Z. Qin *et al.*, "Task-Oriented Multi-User Semantic Communications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2584–2597, 2022.
- [10] W. Zhang, K. Bai *et al.*, "Deepma: End-to-end Deep Multiple Access for Wireless Image Transmission in Semantic Communication," *IEEE Transactions on Cognitive Communications and Networking*, 2023.
- [11] S. Gronauer and K. Diepold, "Multi-Agent Deep Reinforcement Learning: A Survey," *Artificial Intelligence Review*, pp. 1–49, 2022.
- [12] O. Naparstek and K. Cohen, "Deep Multi-User Reinforcement Learning for Distributed Dynamic Spectrum Access," *IEEE Transactions on wireless communications*, vol. 18, no. 1, pp. 310–323, 2018.
- [13] M. Sohaib, J. Jeong, and S.-W. Jeon, "Dynamic Multi-Channel Access via Multi-Agent Reinforcement Learning: Throughput and Fairness Guarantees," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 3994–4008, 2021.
- [14] Z. Guo, Z. Chen *et al.*, "Multi-Agent Reinforcement Learning-Based Distributed Channel Access for Next Generation Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1587–1599, 2022.
- [15] T. Rashid, M. Samvelyan *et al.*, "Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning," *The Journal of Machine Learning Research*, vol. 21, no. 1, pp. 7234–7284, 2020.
- [16] H. Mazandarani, M. Shokrnezhad, and T. Taleb, "Self-Sustaining Multiple Access with Symmetry-aware Continual Deep Reinforcement Learning for Dynamic Metaverse Applications," in *IEEE Transactions on Mobile Computing (TMC)*, 2023, (Under Review).
- [17] H. Mazandarani, M. Shokrnezhad, T. Taleb, and R. Li, "Self-Sustaining Multiple Access with Continual Deep Reinforcement Learning for Dynamic Metaverse Applications," *IEEE MetaCom, Kyoto*, 2023.
- [18] M. Farhoudi, M. Shokrnezhad, and T. Taleb, "QoS-Aware Service Prediction and Orchestration in Cloud-Network Integrated Beyond 5G," *arXiv preprint arXiv:2309.10185*, 2023.
- [19] M. Shokrnezhad, T. Taleb, and P. Dazzi, "Double Deep Q-Learning-based Path Selection and Service Placement for Latency-Sensitive Beyond 5G Applications," *IEEE Transactions on Mobile Computing*, pp. 1–14, 2023.
- [20] P. Sunehag, G. Lever *et al.*, "Value-Decomposition Networks for Cooperative Multi-Sgent Learning," *arXiv preprint arXiv:1706.05296*, 2017.