

A Semantic-Aware Multiple Access Scheme Leveraging Spatial Redundancy for Uplink-Dominant Network Services

Hamidreza Mazandarani¹, Masoud Shokrnezhad², and Tarik Taleb¹

¹ *Ruhr University Bochum, Bochum, Germany; {hamidreza.mazandarani, tarik.taleb}@rub.de*

² *ICTFicial Oy, Espoo, Finland; masoud.shokrnezhad@ictficial.com*

Abstract—The transition toward semantic-aware communication offers a paradigm shift for next-generation mobile networks, promising to decouple information significance from raw data transmission. Despite advances in semantic extraction, the integration of semantic intelligence into the Medium Access Control (MAC) layer remains underexplored, particularly in exploiting spatial correlations among users. To address this, we introduce a novel multiple access scheme designed for uplink-dominant network services. This framework optimizes the trade-off between spectrum utilization and sustainability by formulating variable-packet-length access as distinct α -fairness and energy efficiency problems. A key innovation of our approach is the quantification of spatial redundancies through novel metrics of self-throughput and assisted-throughput, which account for the semantic correlation of data across user equipment. We analyze these formulations to identify optimal bounds before proposing PRISM (Protocol for Redundancy Identification in Semantic Multiple-access). Grounded in Model-free Multi-Agent Deep Reinforcement Learning (MADRL), PRISM enables devices to autonomously govern spectrum access using only local observations. Extensive evaluations demonstrate that PRISM successfully leverages redundancies to outperform semantic-oblivious schemes, achieving up to 90% of the centralized optimal benchmark and improving both objectives by up to $2\times$ across diverse user-semantic association matrices. These results validate PRISM as a viable candidate for future distributed mobile network applications, complemented by orthogonal Multiple Access Schemes where signals are multiplexed in the semantic domain.

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

I. INTRODUCTION

Communication systems are transitioning from a traditional *bit-oriented* mindset, which focuses solely on transmitting the maximum number of data bits by either expanding resources or enhancing resource efficiency, to a *semantic-aware* paradigm. In this novel approach, the semantic value of various system components (including messages, users, and resources) is quantified and subsequently integrated into the service provisioning process to enhance semantic efficiency (Fig. 1) [1]. This paradigm shift opens up significant opportunities for innovative services, particularly within the context of the 6G-based Metaverse, where the seamless integration of physical and virtual environments is a primary objective [2]. Furthermore, such a minimalist strategy, based on the design philosophy of *Less data, more knowledge* [3], promotes the development of sustainable networks [4] by prioritizing the efficient and equitable use of resources. As an illustrative ex-

ample, consider Unmanned Aerial Vehicles (UAVs) transmitting partially overlapping observations to a central server. By identifying semantic segments such as vehicles and avoiding repetitive transmissions, the system's cognitive performance is preserved with reduced bandwidth and energy usage (Fig. 2).

While semantic awareness has been explored across various wireless domains, including adaptive video streaming [5] and channel identification [6], its application at the data link layer, particularly within the Medium Access Control (MAC) protocol, still requires further investigation. This direction is promising, as it offers significant potential to improve network efficiency through more intelligent management of data packet transmissions and resource allocation [1], [3], [7], [8]. Early examples, such as the Model Division Multiple Access (MDMA), demonstrate the promise of this approach by allocating resources for shared information transmission among users [9]. Another research direction involves joint semantic encoding for multiple users, known as multi-user semantic communications [10], [11], although these methods often lack adaptability and scalability for 6G systems. Conversely, Multi-Agent Deep Reinforcement Learning (MADRL) is effective in environments with multiple evolving actors [12]. MADRL adapts to non-stationarity by making transmission decisions based on historical data and coordinating through policy sharing and centralized training. Despite its potential [13]–[15], most current MADRL-based multiple access controls do not incorporate semantic context understanding, which is increasingly necessary for the evolving communication landscape.

In response to the existing gap in the current literature, we have recently introduced a novel approach to address the semantic-aware, distributed multiple-access challenge [17]. Referred to as the Semantic-Aware Multi-Agent Double and Dueling Deep Q-Learning (SAMA-D3QL) method and expanded upon the established principles outlined in our prior investigations [18]–[20], our method aims to optimize the trade-off between resource utilization and fairness, taking into account the inherent correlation present in users' data. It operates on the assumption that users engage in the transmission of diverse semantic segments, with certain segments being potentially shared among them. Consequently, when a user initiates transmission, it confers benefits upon all users with shared relevant segments. Drawing upon the insights garnered from SAMA-D3QL, this study offers noteworthy contributions to the field, as outlined below:

- We introduce a comprehensive framework in which users' utilities are not as independent as is commonly assumed

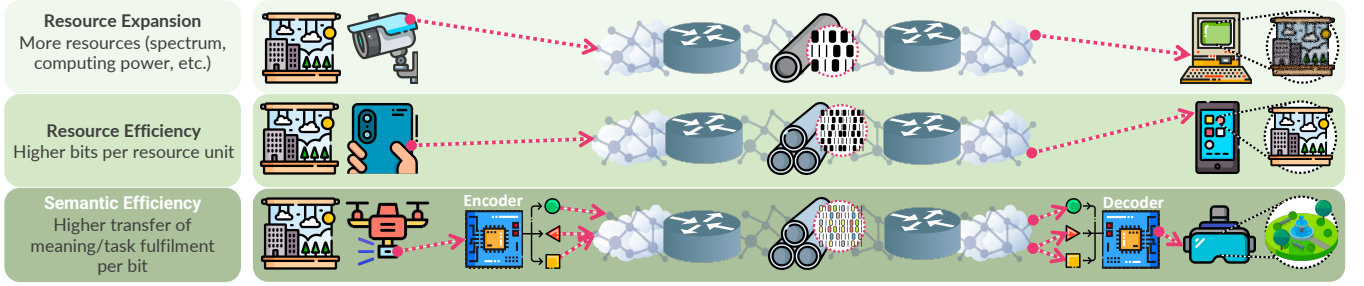


Fig. 1: The emergence of the beyond-Shannon paradigm evolved from unsustainable resource-expansion mindsets (e.g., increasing spectrum bands) and theoretically limited resource-efficiency improvement methods (e.g., novel multiplexing techniques) to enhancing semantic efficiency.



Fig. 2: Two UAVs with overlapping observations of a parking lot, sharing common semantic segments (i.e., vehicles within the red rectangular area), creating an opportunity to avoid excessive resource usage while preserving the system's cognitive performance. Image taken from the DOTA dataset [16].

in traditional multiple access networks. This interdependence arises from the existence of shared segments among certain subsets of users, where the transmissions of one user can impact the task fulfillment of others. Examples of such interdependence are observed in scenarios involving holographic presence, distributed Generative Artificial Intelligence (GenAI), and multi-UAV area coverage [20].

- In pursuit of environmental and societal sustainability within the distributed multiple access capability of 6G networks, we leverage energy efficiency as an objective function, in addition to the fairness-utilization trade-off. Our energy efficiency assessment includes a parameterized consideration of the energy consumption associated with semantic encoding, thereby offering a more precise and realistic understanding of the system and aiming to reduce the environmental footprint of the network.
- We formulate the challenge as a novel optimization problem for each objective function, incorporating the variable packet length of User Equipment (UE) and capturing semantic correlations within the resulting asynchronous setup.
- We perform a comparative analysis of optimal solutions for each problem concerning its objective function: energy efficiency, quantified as the difference between users' throughput and coefficient-weighted consumed energy, and the fairness-utilization trade-off, evaluated using alpha-fairness metrics with varying values of α . While achieving optimal solutions in real-world scenarios is

computationally infeasible due to the complexity of the problems, these solutions serve as target bounds for environmental and societal sustainability in the design of real-time approaches.

- We develop a MADRL-based approach, named PRISM, to solve these problems in a practical manner by incorporating a novel observation space and reward mechanism tailored to semantics and asynchrony. Specifically, we employ only fully-connected layers as the approximation function in the MADRL algorithm, processing short- and long-term histories with an averaged context vector, thereby enhancing the system's sustainability. Moreover, we design objective-specific reward shaping to achieve the right trade-offs for both the α -fairness and energy efficiency objectives.

The subsequent sections of this paper are structured as follows: Section II provides the necessary background and a comprehensive literature review. Section III explains the motivation behind this work, introduces the system model, and presents a thorough problem formulation. Additionally, this section proposes an interpretation of the optimal solutions. Section IV elaborates on PRISM, detailing its learning mechanism and training algorithm. Following this, Section V presents an exposition of the numerical results obtained. Finally, Section VI concludes our work with closing remarks on our findings.

II. BACKGROUND AND RELATED WORK

A. From Bits to Semantics

Historically, the response to emerging applications in wireless communications has predominantly focused on expanding resources (e.g., exploring new spectrum bands) or enhancing resource efficiency (e.g., advancing MIMO technology). Nonetheless, these efforts are insufficient to support innovative communication services, particularly the 6G-based Metaverse, where seamless integration of physical and virtual environments is a primary goal [2]. Fortunately, a semantic revolution is underway, shifting from bit transmission to the transfer of semantics among transceivers, along with semantic-aware resource allocation and orchestration. Semantic Communications (SemCom) aims for the receiver to accurately interpret the transmitted message as intended by the sender [21]. This objective can be realized differently based on system goals

and data modalities. For example, the transmitter may identify and prioritize transmission of key segments in images via segmentation models like the Segment Anything Model (SAM) versions one and two [22]. Moreover, joint source and channel coding can be applied in digital communications to transfer semantics despite channel impairments [23]. Graph-semantic compression is also another option to realize SemCom [24]. Nonetheless, SemCom necessitates mutual understanding, or even a theory of mind [21], among transceivers, which involves an awareness of the entire communicative system regarding semantics, to be discussed in the next subsection.

B. Semantic-Awareness Beyond the Physical Layer

The integration of semantics into decision-making processes at various network layers remains in its early stages of development. Our knowledge management framework for semantic-aware systems, termed *Knowledge Base Management and Orchestration* (KB MANO) [1], alongside the concept of the knowledge space within the Meta-Networking framework [7], the reasoning plane proposed by Lin *et al.* [3], and the 6G-GOALS architecture [8], represent some of the initial efforts to conceptualize network-wide semantic awareness. Additionally, there is an increasing focus on specific semantic-aware functionalities, such as routing, radio resource allocation, and service provisioning, as surveyed by Getu *et al.* [25] and Trevlakakis *et al.* [26]. Among these, the enhancement of multiple access to the frequency spectrum through semantic awareness holds significant potential. As the initial point where user data enters the external environment, the MAC layer can significantly benefit from semantic-aware protocols by eliminating redundant transmissions and optimizing resource use.

MDMA is an early example of exploiting this potential, dedicating a portion of resources to transmit shared information among users [9]. Orthogonal MDMA (O-MDMA) further advances this scheme by introducing the concept of semantic orthogonal signals, i.e., the capability to treat signals generated by different semantic encoders as interference while meeting user performance requirements [27]. O-MDMA also accommodates multiple modalities and variable semantic signal lengths. Similarly, Kassab *et al.* [28] explored scenarios involving a collective of Internet of Things (IoT) devices monitoring correlated on-off events, such as temperature variations. The authors suggest that devices observing the same event could send redundant information, so they propose maximizing the average total event rate of the system instead of the traditional frame throughput. More recently, Qiao *et al.* [29] propose token-domain multiple access (ToDMA), which leverages the semantic orthogonality of tokens. These tokens result from mapping input segments into discrete information units using techniques like vector quantization.

Another research direction, exemplified by DeepSC-IR/MT [10] and DeepMA [11], investigates the joint semantic encoding of data from multiple users, also known as multi-user semantic communications. Although these methods are effective in certain contexts, they do not adequately address the distributed nature of future applications with shared user

information, limiting their scalability and suitability for 6G systems. In this regard, advanced decision-making algorithms, such as RL, could elevate semantic awareness to a higher level.

C. MADRL: An enabler for semantic-aware Multiple Access

In environments characterized by ever-changing, multiple distributed actors, such as unsourced multiple access in 6G [30] or UAV-assisted wireless sensor networks [31], MADRL proves to be an effective approach. Its adaptability to handle non-stationarity, resulting from the co-evolution of agents' policies, underscores its suitability for such decentralized contexts [12]. In MADRL frameworks, agents make transmission decisions based on historical observations, including the success or failure of task completions (e.g., successful transmissions in the multiple access problem), while coordination among agents is facilitated through mechanisms such as policy sharing, message passing, or centralized training.

MADRL-based multiple access control has been discussed in various papers. For instance, Naparstek *et al.* [13] proposed an efficient Nash equilibrium for multi-agent MAC where all agents adopt the same policy. Sohaib *et al.* [14] 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.* [15] introduced the QMIX-advanced LBT (QLBT) MAC protocol, building upon the QMIX algorithm for MADRL problems [32]. QLBT incorporates two reward signals, one for maximizing network utilization and another for maintaining fairness. Miuccio *et al.* [33] designed a MAC algorithm considering communication overhead issues (e.g., scheduling requests and grants) with generalization capability using auto-encoder-based state abstraction. 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.

D. Sustainable Semantic-Awareness: the final puzzle piece

Alongside being distributed, any proposed approach for future systems should consider energy efficiency as a critical objective for optimization. Energy efficiency is regarded as a key factor in environmental sustainability, also impacting business and economic sustainability [34]. The minimalist design approach of semantic communications paves the way to address this challenge towards sustainable networks [4]. The promises of semantic-aware network orchestration could be even more transformative, given its influence on the entire communication system. This approach allows the network to prioritize essential and high-value information, thereby reducing the need to transmit redundant or low-priority data. By minimizing unnecessary transmissions, the overall energy usage of the network is significantly decreased. Additionally, semantic-aware systems can facilitate more efficient resource allocation and management by sharing or deactivating a substantial portion of resources, thereby reducing energy waste in network operations and maintenance. Conversely, the energy

consumption associated with semantic learning and extraction could negatively impact sustainability.

The energy-defined networking concept, which integrates energy measurement components across all system layers [35], offers a viable solution by allowing the consideration of energy consumption in evaluating system efficiency and sustainability. Despite this, only a limited number of studies address the sustainability aspects of semantic communications. Some research targets the computational disparities between the capabilities of different users [36], [37], while others focus on the efficiency of semantic communications training [38]. However, it is necessary to consider more precise energy consumption models and investigate the impact of practical metrics for real-world usage, such as semantic encoding, on system efficiency and effectiveness, drawing inspiration from the work on incorporating decision-making costs into energy efficiency in multiple access [39].

E. Our Approach

Complementary to some existing work that concentrates on the semantic *orthogonality* of users [9], [27], [29], we target *similarities* among users' information. To this end, we propose a comprehensive framework that challenges the conventional assumption of independent user utilities, considering the interdependence arising from shared semantic segments among user subsets. Such interdependence is relevant in scenarios like holographic presence, distributed Generative AI, and multi-UAV area coverage. Exploiting redundancy among users is not new in itself: correlation-based MAC allows a representative node to report on behalf of its correlation neighbors [40], redundant coverage can be provisioned at deployment time in sensor networks [41], and event-driven access rewards the events reported rather than the frames delivered [28]. In these schemes, the redundancy structure is given externally, through deployment geometry, a commonly observed event, or a model shared across encoders, and is therefore known before access decisions are made. Here, in contrast, redundancy is content-defined: which semantic segments are shared, and by which subset of UEs, changes over time and is not known at the UEs. Each UE must instead infer its semantic contribution from access feedback and act on that information in a distributed, asynchronous manner.

Our work incorporates energy efficiency, factoring in the energy costs of semantic encoding as a key objective alongside the fairness-utilization trade-off. The problem is formulated as a novel optimization task that accounts for variable packet lengths and semantic correlations in an asynchronous setting. Here, variable packet length refers to transmitting packets that span more than one frame, highlighting the system's greater freedom in decision-making [18], [42]. In Table I, a comparison of the studied papers about MAC design, including their main ideas and characteristics, is presented.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

An indoor area comprising a Small Base Station (SBS) and a collection of $\mathcal{N}$ intelligent User Equipment (UE) is considered,

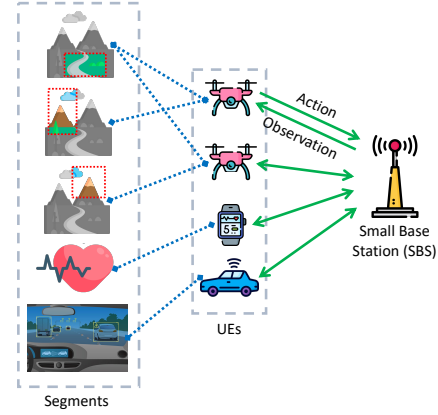


Fig. 3: A sample network consisting of four UEs ($\mathcal{N} = 4$) and five segments ($\mathcal{K} = 5$). As illustrated, one separate segment is associated with each user, in addition to one shared segment among the first and second UEs.

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 uploading their data to the SBS. It is essential to note that simultaneous transmissions over a single channel lead to a collision. Consequently, in the conventional bit-oriented framework, the highest achievable network utilization is capped at $\mathcal{C}$ packets per time slot.

The semantic space is represented by $\mathcal{K}$ distinct entities, termed *segments*, denoted as $k \in \mathbb{K} = \{1, \dots, \mathcal{K}\}$. At each time slot $t \in \{0, \dots, \mathcal{T}\}$, each user u_i is associated/disassociated with segment k , signified by a binary variable $a_{i,k}^t \in \{0, 1\}$, forming a predefined binary association matrix $\mathbb{A} = [a_{i,k}^t]_{\mathcal{N} \times \mathcal{K} \times \mathcal{T}}$. Operationally, $a_{i,k}^t = 1$ means that the semantic descriptor extracted by UE i at slot t is assigned to the global segment k according to a service-specific matching rule, such as label agreement, spatial registration, hash matching, or embedding similarity. 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. 3, the association matrix at time slot t is given as $\mathbb{A}^t = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. The following subsection elaborates further on the details of these semantic correlations and their impact on the system performance.

1) *Semantic Spatial Correlations*: To evaluate the semantic efficiency of the system, we introduce a *segment transmission* indicator, denoted as $y_k^t \in \{0, 1\}$. 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 as¹:

$$y_k^t \triangleq \mathbb{1} \left(\sum_{j \in \mathbb{N}} a_{j,k}^t \bar{m}_j^t > 0 \right), \quad (1)$$

¹A constraint of the form $y \triangleq \mathbb{1}(x > 0)$ (where x is an integer) can be expressed as the following linearized constraints: $yM \geq x$ and $y \leq x$ where M is a large number.

TABLE I: Selected Multiple Access Schemes Comparison.

Reference	Main Idea	Objective Function	Semantic Awareness	Time Awareness
MDMA [9]	Dedicating a portion of resources to transmit shared information among users	Peak signal-to-noise ratio	✓	—
O-MDMA [27]	Semantic orthogonal signals, i.e. the capability to treat signals generated by different semantic encoders as interference while meeting user performance requirements	Multi-scale structural similarity	✓	—
Kassab <i>et al.</i> [28]	Avoid transmitting redundant information by IoT devices observing the same event	Average sum event rate	✓	✓
DeepSC-IR/MT [10]	A Transformer-based transmitter structure to extract semantic information for different tasks, while eliminating interference from other users	- Image: Recall@1 - Text: Bilingual evaluation understudy - Video: Answer accuracy	✓	—
DeepMA [11]	Transmitting multiple data instances simultaneously over a shared communication channel via training the network to recover the target data from a mixed semantic symbol vector	Peak signal-to-noise ratio	✓	✓
Naparstek <i>et al.</i> [13]	Adopt MADRL for distributed dynamic spectrum access + game theory analysis	Average rate	—	✓
Sohaib <i>et al.</i> [14]	Selecting transmission policies over consecutive time slots ensuring short-term fairness under varying user conditions	- Throughput - Fairness	—	✓
QLBT [15]	Leveraging QMIX MADRL algorithm, in addition to novel observations and reward mechanism	- Throughput - Fairness	—	✓
Miuccio <i>et al.</i> [33]	Considering communication overhead issues with generalization capability using auto-encoder-based state abstraction	- Throughput - Generalizability	—	✓
SAMA-D3QL [17] (Our previous work)	Accounting for user data correlation by introducing self- and assisted throughputs in MADRL-based MAC scheme design	α -fairness	✓	✓
PRISM (This work)	Expanding SAMA-D3QL to variable packet length setting, and sustainability awareness	- α -fairness - Energy efficiency	✓	✓

where $\bar{m}_j^t \in \{0, 1\}$ signifies whether UE j has successfully transmitted during time slot t . Based on (1), the normalized semantic throughput of each UE can be determined. Considering that $\mathbb{W}^t = \{t - \mathcal{H}, \dots, t\}$ represents a sliding time window of length $\mathcal{H}$, b_k is the size of segment k in bits, and $\Omega_i^t = 1 / (\sum_{k, t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t b_k)$ acts as the normalization factor, the normalized semantic throughput of user u_i is as follows:

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

By isolating user u_i from other UEs in (2), we can break down x_i^t into two components: *self throughput* ($\dot{x}_i^t$) 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 user u_i to the SBS.

$$\begin{aligned} x_i^t &\stackrel{(1), (2)}{=} \dot{x}_i^t + \ddot{x}_i^t \\ \dot{x}_i^t &= \Omega_i^t \sum_{k, t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t b_k \bar{m}_i^t \\ \ddot{x}_i^t &= \Omega_i^t \sum_{k, t \in \mathbb{K}, \mathbb{W}^t} \left(a_{i,k}^t b_k (1 - \bar{m}_i^t) y_k^{t, -i} \right) \end{aligned} \quad (3)$$

Here, $y_k^{t, -i} = \min \{1, \sum_{j \in \mathbb{N} \setminus \{i\}} a_{j,k}^t \bar{m}_j^t\}$, indicating the transmission of segment k at time t by all users except user i ; if more than one user transmits, only one transmission is counted.

2) *Correlated Utilities*: Following the defined semantic correlations among users, it is evident that users' utilities are not independent, as commonly assumed in traditional networks. Specifically, by allocating more resources to user i at time $t+1$, the utility of another user j may also increase. Mathematically, this can be expressed as $\exists i, j \neq i, t : \mathfrak{R}_i^{t+1} > \mathfrak{R}_i^t \longrightarrow \mathcal{U}_j^{t+1} > \mathcal{U}_j^t$, where $\mathfrak{R}_i^t$ and $\mathcal{U}_i^t$ denote the resources allocated to and the utility of user i at time t , respectively. This utility correlation is observed in various real-world scenarios, including the setup from our previous work on Multi-UAV Area Coverage [20]. In this setup, multiple UAVs collaboratively transmit partially overlapping observations to a central server. A transmission by each UAV may lead to a better reconstructed image and, consequently, more effective tasks like object detection for nearby UAVs.

Although shared semantic information exists among user subsets in all these scenarios, we are not restricted to situations where only a subset of users share data; users may also possess information that plays an equally important role in task completion. For instance, in a collaborative GenAI setup [47], users work together to create synthetic data. Suppose the responsibility of generating a certain subspace of the data is allocated to more than one user. In that case, semantic-aware decision-making will enable the selection of users based on alternative factors such as fairness or users' energy levels.

Table II summarizes additional scenarios and further details. In general, this framework applies to any contention-based many-to-one uplink transmission with local access feedback, where a service-specific rule maps UE descriptors to com-

TABLE II: Potential Use Cases for Our Framework.

Item	Task	Semantic segment	Example construction of $\mathbb{A}$
Multi-UAV Area Coverage [20], [31], [43]	UAVs send overlapping views of the same area.	Objects or Image regions	If two UAVs detect the same object in the overlap, both are linked to the same segment (illustrated in Fig. 2).
Holographic Presence [2]	XR users share parts of their local scenes.	Objects and spatial relations	Matching the same object or relation, e.g., table-next-to-chair, creates a shared segment.
IoT Monitoring System [28]	Sensors monitor the same physical event.	Events, e.g., high temperature	Sensors reporting the same event in the same area and time window share a segment.
Image Segmentation Reconstruction [44]	UEs send parts of an image segmentation map.	Image masks or regions	Masks with the same class and strong overlap are mapped to the same segment.
Cooperative Navigation [45]	Agents share local map observations.	Grid-cell states, e.g., obstacle/free	Reports for the same grid cell and state are assigned to the same segment.
6G Digital Twin [46]	Devices update the same digital replica.	Twin objects or states	Reports about the same component or state are mapped to the same segment.
Collaborative GenAI [47]	Users generate related parts of a shared output.	Generated content blocks	Users working on the same content block (e.g, a room) share a segment.
Satellite Earth Observation [48]	Satellites report changes over Earth regions.	Geo-tagged regions or events	Detections with the same location, event type, and time window share a segment.

mon segment identities, concurrent duplicate deliveries of a segment are counted only once, and semantic overlap is non-negligible and persists during policy learning. In other words, the current framework offers no semantic-redundancy gain for independent payloads or when encryption prevents segment matching, and it is not intended for downlink or centrally scheduled access, hard per-packet guarantees, or services requiring per-UE raw-data provenance.

3) *Fairness-utilization trade-off*: To ensure fairness among UEs while maximizing network utilization, the α -fairness metric can be used as the objective function, as defined by the following formula.

$$\mathcal{F}_\alpha^t = \begin{cases} \sum_{i \in \mathbb{N}} \log(x_i^t) & \text{if } \alpha = 1 \\ (1 - \alpha)^{-1} \sum_{i \in \mathbb{N}} (x_i^t)^{1-\alpha} & \text{if } \alpha \neq 1 \end{cases} \quad (4)$$

When $\alpha \rightarrow 0$, optimizing (4) becomes equivalent to maximizing the sum of throughputs, commonly known as utilitarian fairness or sum-rate maximization. In this scenario, the objective is to maximize the total throughput, with minimal consideration for its distribution among users. Although this method is generally more efficient, it can result in significant inequality. Conversely, when $\alpha \rightarrow \infty$, optimizing (4) shifts to maximizing the minimum user's throughput (i.e., $\max \min_i x_i^t$), known as max-min fairness. Max-min fairness ensures that throughput allocations are adjusted to enhance the allocation received by the user with the least resources, thereby achieving the highest level of equality among users, albeit potentially at the expense of overall efficiency.

One should note that even within a single application, UE contributions are only partly redundant, as each UE also

holds segments that no other UE can deliver, and each access attempt draws transmission and encoder-inference energy from a battery-limited device; α therefore controls whether access prioritizes aggregate semantic throughput ($\alpha \rightarrow 0$) or limits the persistent starvation of individual UEs and the energy concentration it implies. Furthermore, a single application does not imply a single utility holder, as the UEs of one service may belong to distinct stakeholders whose individual utilities remain separately meaningful, such as the portion of a holographic scene that is private to one participant.

4) *Energy Awareness*: To maximize network utilization while minimizing energy consumption, we adopt a linearized energy-efficiency utility and define it for user u_i as the difference between its throughput at time slot t (i.e., x_i^t) and a coefficient ζ of its consumed energy. The consumed energy includes both packet transmissions, which are a function of its own throughput, and semantic encoder inferences, which are a function of the number of decision makings (attempts to transmit on any channel) in $\mathbb{W}^t$, denoted with D_i^t .

$$e_i^t = x_i^t - \zeta(g(\dot{x}_i^t) + h(D_i^t)) \quad (5)$$

Assuming a simple premise, we posit in this paper that $g(\dot{x}_i^t) = \sigma_{tr} \dot{x}_i^t$ and $h(D_i^t) = \sigma_{inf} D_i^t + \sigma_{init}$, where σ_{tr} denotes the energy consumption for transmission during a single time slot, σ_{inf} represents the energy required for the semantic extraction of one packet (i.e., inference on the encoder), and σ_{init} denotes the energy needed to initialize the semantic encoders.

In a bit-oriented network, σ_{inf} and σ_{init} lack relevance. However, in our energy-aware semantic-oriented setting, the relative importance of σ_{inf} and σ_{init} compared to σ_{tr} significantly impacts users' utilities and their behavior. Consequently,

the energy efficiency of the system is determined by the sum of the energy efficiencies of the UEs. In the following formula, $d_1 = \frac{\sigma_{inf}}{\sigma_{tr}}$ and $d_2 = \frac{\sigma_{init}}{\sigma_{tr}}$.

$$\mathcal{E}_{\zeta, d_1, d_2}^t = \sum_{i \in \mathbb{N}} \left(x_i^t - \zeta (\dot{x}_i^t + (d_1 D_i^t) + d_2) \right) \quad (6)$$

As $\zeta \rightarrow 0$, optimizing (6) simplifies to a sum-rate maximization problem (akin to the scenario in (4) as $\alpha \rightarrow 0$). By increasing ζ , the consideration of energy consumption becomes more important in the decision-making, ultimately leading to inactive users avoiding negative objective values (since the maximum of a negative value is zero). Therefore, ζ should be thoughtfully adjusted based on the system characteristics and goals. Another determining parameter is d_1 , and its value influences the average packet size that users transmit (as investigated in section III-C.). It is worth noting that in the special case where d_1 and d_2 equals zero (i.e., $\mathcal{E}_{\zeta, d_1, d_2}^t = \sum_{i \in \mathbb{N}} (x_i^t - \zeta \dot{x}_i^t)$), complexity persists since users' energy consumption is a function of self throughputs.

5) *Practical Challenges of Semantic Orchestration*: Before formulating the problem, it is worth mentioning the practical challenges that arise in defining semantic segments and association matrices. The KB-MANO framework [1] can be utilized as the semantic orchestration entity to address these challenges. Within this framework, the construction of the association matrix begins with a service-specific definition of a semantic segment. The appropriate granularity depends on the downstream task. In collaborative object detection, a segment may correspond either to an object class or to the presence of an object in a region. For scene reconstruction tasks, however, more descriptive semantic segments (e.g., object embeddings within the scene) improve the fidelity of the inferred correlations but also increase the number of candidate segments, the matching complexity, and the signaling overhead. Hence, the segment vocabulary $\mathbb{K}$ and its granularity are service-level design parameters rather than universal quantities.

Another challenge is that the base station, which is responsible for reconstructing users' data or performing downstream tasks on it (e.g., object detection or visual question answering), as well as rewarding users based on their assisting data, does not have access to users' data in advance to construct the association matrix. To address this, the base station can collect semantic metadata from the UEs (e.g., a segment identifier, a quantized embedding, a hash, or a confidence score, rather than the raw observation), assuming that the overhead of transmitting them is negligible. The base station then forwards this metadata to the KB MANO framework, which resolves descriptor identities across users and constructs the estimated association matrix. In slowly varying settings, the matrix can be propagated from previous time slots using object tracking, mobility prediction, or temporal smoothing, with event-triggered recomputation when the observed confidence or application-level QoS/QoE falls below a prescribed threshold. The present optimization uses the estimated association matrix as a deterministic input and does not explicitly account for descriptor-signaling overhead or matrix-estimation errors.

B. Problem Formulation

In this subsection, we present a Mixed-Integer Non-Linear Programming (MINLP) formulation to define two centralized optimization problems. This formulation serves as a centralized collision-free benchmark rather than a distributed MAC protocol. In particular, the optimizer is assumed to have system-wide information over the optimization horizon and to schedule packet starts, packet lengths, and channels. Collisions are therefore excluded by construction. In contrast, collisions may occur in the distributed PRISM protocol introduced in Section IV, where UEs independently select their actions from local observations.

The subsequent subsection will further elaborate on the problems' variables and constraints, as well as the objective function. Notably, the variable packet-length formulation is an enhancement of our previous work [18] to a multi-agent setup.

1) *Decision and Support Variables*: The primary integer decision variable in our problems is $\mathfrak{R} = [r_{i,c}^t]_{\mathcal{N} \times \mathcal{C} \times \mathcal{T}}$, where each element represents the size of the packet that user u_i begins to transmit on channel c at the start of time slot t . Additionally, $\mathbb{Z} = [z_{i,c}^t]_{\mathcal{N} \times \mathcal{C} \times \mathcal{T}}$ denotes an auxiliary integer variable indicating the remaining time slots for the ongoing transmission. The variable $\mathbb{M} = [m_{i,c}^t]_{\mathcal{N} \times \mathcal{C} \times \mathcal{T}}$ serves as another auxiliary variable, set to one whenever $z_{i,c}^t$ is positive. More specifically, $m_{i,c}^t$ is a channel-occupancy indicator for UE i on channel c at time slot t . It is not an independently selected scheduling action; rather, it is determined by the remaining-duration variable $z_{i,c}^t$. If $z_{i,c}^t = 0$, then UE i is not occupying channel c in slot t , and therefore $m_{i,c}^t = 0$. If $z_{i,c}^t > 0$, then UE i is still transmitting a packet on channel c , and therefore $m_{i,c}^t = 1$. For example, if UE i starts a packet of length 3 on channel c at time slot t , then the corresponding occupancy indicator remains 1 for the next 3 occupied slots. This indicator is later used in the channel-occupancy constraints to prevent multiple UEs from occupying the same channel in the centralized benchmark. The following constraints transform the equation $m_{i,c}^t = \mathbb{1}(z_{i,c}^t > 0)$ into a linear form:

$$\begin{aligned} m_{i,c}^t &\leq z_{i,c}^t, \\ m_{i,c}^t R_{max} &\geq z_{i,c}^t, \end{aligned} \quad (7)$$

where R_{max} is the maximum allowed packet size.

2) *Temporal Interdependency Constraints*: Now, we need to establish interdependency over time. First, we must ensure that z decreases by one with each passing time slot for ongoing transmissions. This requirement is enforced by the following constraint.

$$z_{i,c}^t = z_{i,c}^{t-1} - 1 \quad \text{if } ((m_{i,c}^t = 1) \wedge (d_{i,c}^t = 0)) \quad (8)$$

This constraint can be expressed in its linear equivalent forms as follows:

$$\begin{aligned} z_{i,c}^t &\leq (z_{i,c}^{t-1} - 1) + (R_{max}((1 - m_{i,c}^t) + d_{i,c}^t)), \\ (z_{i,c}^{t-1} - 1) &\leq z_{i,c}^t + (R_{max}((1 - m_{i,c}^t) + d_{i,c}^t)), \end{aligned} \quad (9)$$

where $\mathbb{D} = [d_{i,c}^t]_{\mathcal{N} \times \mathcal{C} \times \mathcal{T}}$ is a binary variable used to activate decision making (i.e., starting packet transmission) over specific time slots.

Second, it must be ensured that a new transmission is not initiated until the ongoing one is completed. To achieve this, we establish a set of constraints based on the premise that whenever user u_i initiates the transmission of a packet with length k at time slot t , the decision-making variables for the subsequent $k - 1$ time slots must be zero (i.e., $d_{i,c}^\tau = 0 \forall \tau \in [t + 1, t + k - 1]$):

$$r_{i,c}^t = k \rightarrow \sum_{\tau=t+1}^{t+k-1} d_{i,c}^\tau = 0. \quad \forall k \in [1, R_{max}] \quad (10)$$

This can be expressed in its linear equivalent form as follows (where K_1 and K_2 are auxiliary binary variables):

$$(r_{i,c}^t - k) + 1 \leq R_{max} K_1 \quad (11)$$

$$(k - r_{i,c}^t) + 1 \leq R_{max} K_2 \quad (12)$$

$$\sum_{\tau=t+1}^{t+k-1} d_{i,c}^\tau \leq R_{max}(2 - (K_1 + K_2)) \quad (13)$$

3) *Packet Length Constraints:* Next, we establish the relationship between the packet length variable r and the support variables z and d using (14), where r is equal to z when transmission starts:

$$r_{i,c}^t = \begin{cases} z_{i,c}^t & \text{if } d_{i,c}^t = 1 \\ 0 & \text{if } d_{i,c}^t = 0 \end{cases} \quad (14)$$

We then transform (14) into a linear form, similar to previous constraints:

$$\begin{aligned} r_{i,c}^t &\leq z_{i,c}^t + (R_{max}(1 - d_{i,c}^t)) \\ z_{i,c}^t &\leq r_{i,c}^t + (R_{max}(1 - d_{i,c}^t)) \\ r_{i,c}^t &\leq R_{max}d_{i,c}^t \end{aligned} \quad (15)$$

A simple example, as illustrated in Fig. 4, demonstrates the values of r and the other auxiliary variables. It is important to note that in a single time-slot scenario, the variables r , z , m , and d are equal and binary.

4) *Physical Constraints:* Final constraints, (16) and (17), ensure that each user transmits on at most one channel per time slot, and that no more than one user transmits simultaneously on the same channel that satisfies the required quality condition, as indicated by Q_c^t .

$$\sum_{c \in \mathbb{C}} m_{i,c}^t \leq 1 \quad \forall i \in \mathbb{N}, \forall t \in [0, T] \quad (16)$$

$$\sum_{i \in \mathbb{N}} m_{i,c}^t \leq Q_c^t \quad \forall c \in \mathbb{C}, \forall t \in [0, T] \quad (17)$$

5) *Problems:* Finally, we address two non-linear problems to maximize α -fairness and linearized system energy efficiency utility (presented in P_1 and P_2 , respectively). The non-linearity of both problems arises from the semantic interdependence of packet transmissions. In particular, the definition of $\mathcal{F}_\alpha^t$

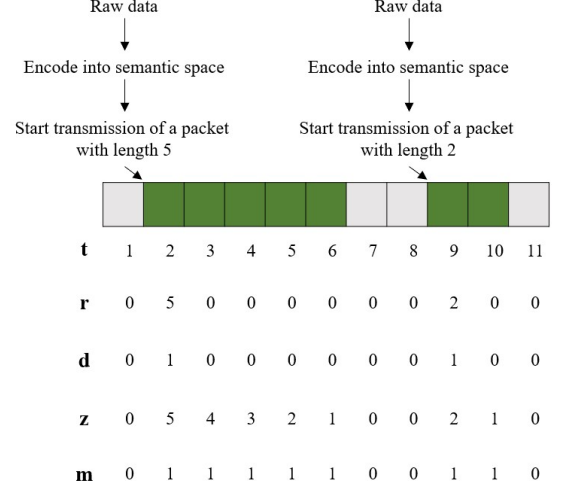


Fig. 4: A scenario in which a UE initiates the transmission of packets with lengths 5 and 2 at time slots $t = 2$ and $t = 9$, respectively. The values of the decision variable r , along with the support variables d , z , and m , can be compared accordingly.

TABLE III: Notations of Symbols Used in the Problem Definitions.

Symbol	Description
u_i	UE with index $i \in \mathbb{N} = \{1, \dots, \mathcal{N}\}$
k	Semantic entity (or segment) $k \in \mathbb{K} = \{1, \dots, \mathcal{K}\}$
y_k^t	Segment transmission indicator for segment k at time slot t
$a_{i,k}^t$	Indicator for association of u_i with segment k at time slot t
$\bar{m}_i^t$	Indicator for successful transmission of u_i during time slot t over all channels (i.e., $\sum_{c \in \mathbb{C}} m_{i,c}^t$)
$\mathbb{W}^t$	Sliding time window of length $\mathcal{H}$, i.e., $\{t - \mathcal{H}, \dots, t\}$
x_i^t	Actual throughput of u_i in $\mathbb{W}^t$
Ω_i^t	Normalization factor (i.e., $1/(\sum_{k,t \in \mathbb{K}, \mathbb{W}^t} a_{i,k}^t)$)
$\dot{x}_i^t$	Self throughput of u_i in $\mathbb{W}^t$
$\ddot{x}_i^t$	Assisted throughput of u_i in $\mathbb{W}^t$
$\mathcal{F}_\alpha^t$	α -fairness metric in $\mathbb{W}^t$, as defined in (4)
e_i^t	Energy efficiency of u_i in $\mathbb{W}^t$
D_i^t	Number of decision makings (attempts to transmit on any channel) in $\mathbb{W}^t$
$\mathcal{E}_{\zeta, d_1, d_2}^t$	Energy efficiency of the system in $\mathbb{W}^t$, as defined in (6)
$m_{i,c}^t$	An indicator for transmission of UE i on channel c in time slot t
$z_{i,c}^t$	Number of time slots remaining until the current transmission of u_i on channel c ends
R_{max}	maximum allowed packet size
$r_{i,c}^t$	Size of the packet u_i starts transmitting on channel c at the beginning of time slot t
$d_{i,c}^t$	An indicator for decision-making of UE i on channel c in time slot t
Q_c^t	Quality indicator for channel c in time slot t

remains non-linear except for $\alpha = 0$, necessitating piecewise

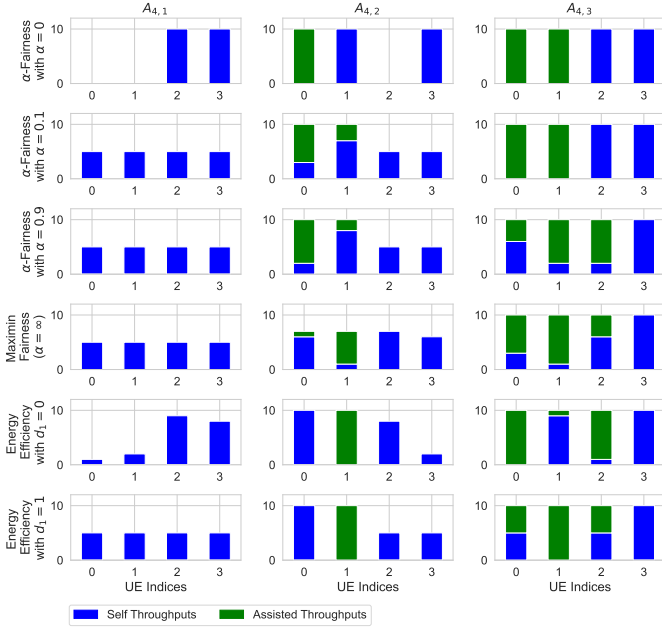


Fig. 5: Optimal self and assisted UE throughputs for different objective functions and association matrices. Note that these optimal points are not necessarily unique.

linear approximation techniques [49].

$$\begin{aligned} \max_R \mathcal{F}_\alpha^T \quad & \text{acc. (4)} \\ \text{s.t. (7), (9), (10), (15), (16), (17)} \end{aligned} \quad (P_1)$$

$$\begin{aligned} \max_R \mathcal{E}_{\zeta, d_1, d_2}^T \quad & \text{acc. (6)} \\ \text{s.t. (7), (9), (10), (15), (16), (17)} \end{aligned} \quad (P_2)$$

C. Optimal Solutions

Despite our efforts to linearize the constraints of problems P_1 and P_2 and employ a linear objective function for energy efficiency, solving them remains intractable due to the presence of multiple integer and binary variables. The solution space for the integer variables $r_{i,c}^t$ and $z_{i,c}^t$ is $(R_{\max} + 1)^{\mathcal{N} \cdot \mathcal{C} \cdot \mathcal{T}}$, and for the binary variables $m_{i,c}^t$ and $d_{i,c}^t$ is $2^{\mathcal{N} \cdot \mathcal{C} \cdot \mathcal{T}}$. Considering all these variables, their solution space sizes must be multiplied together to calculate the solution space of the problems, rendering them impractical to address large numbers of users, channels, maximum packet sizes, or network lifetime. However, to gain insights into problems, we solve them by utilizing the Gurobi optimizer² for a modest setting with $\mathcal{N} = 4$, $\mathcal{C} = 2$, $R_{\max} = 5$, $\mathcal{T} = 10$, and various association matrices $A_{\mathcal{N},q}$, where the first q UEs of $\mathcal{N}$ UEs share their segments ($A_{\mathcal{N},1}$ and $A_{\mathcal{N},\mathcal{N}}$ represent the scenarios of no semantic sharing and full semantic sharing, respectively). The results of these simulations are interpreted and visualized in Fig. 5.

With increasing q in A_q , both objective functions $\mathcal{F}_\alpha^t$ and $\mathcal{E}_{\zeta, d_1, d_2}^t$ improve because users assist each other in their communications. However, this positive impact varies depending

on the objective functions and their parameters. Our observations revealed that metrics focusing less on fairness have a greater potential for improvement since allocating additional resources to users who share segments adds greater value. Precisely, considering the standard deviation among optimal UE throughputs as a measure of *unfairness*³, the objective improvement ratio of $A_{4,3}$ over $A_{4,1}$ has a positive correlation equal to 0.58 with unfairness⁴.

Finally, we provide some intuitions regarding utilization-sustainability trade-offs. In α -fairness maximization, higher α leads to more equal throughputs among users, hence higher societal sustainability. In energy efficiency maximization, when transcoding costs are considered in the objective function ($d_1 = 1$), it results in larger packets. Numerically, the average packet size for energy efficiency with $d_1 = 1$ is 5 while for $d_1 = 0$ equals 2.5.

IV. PRISM

The optimization problems P_1 and P_2 characterize the desired semantic-aware allocation, but they are not suitable for online MAC operation. First, their search spaces grow exponentially with the number of UEs, channels, time slots, and admissible packet lengths, as discussed in Section III-C. Solving them also requires centralized, system-wide information over the optimization horizon. In contrast, a UE must make each access decision promptly from local observations and without knowing the concurrent decisions of the other UEs. Conventional bit-oriented MAC schemes are likewise insufficient because they treat successful packets as independent contributions: they cannot determine whether transmitting a packet delivers new semantic segments or merely repeats segments already supplied by another UE. Moreover, a decision must jointly select a channel and a packet length. Longer packets may improve utilization and amortize semantic-encoding costs, but they occupy a channel for multiple time slots and expose the UE to a longer collision interval. These coupled semantic, temporal, fairness, and energy effects make a fixed or myopic access rule difficult to design.

The resulting decision process is naturally modeled as a Macro-Action Decentralized Partially Observable Markov Decision Process (MacDec-POMDP). A MacDec-POMDP extends a Dec-POMDP by allowing actions with variable durations, such as the multi-slot packet transmissions considered here [50]. Partial observability arises because each UE observes only the channel that it senses or uses, while the success or collision outcome of a multi-slot transmission is revealed only after that transmission terminates. A learning-based solution is therefore needed to infer effective access policies from interaction histories rather than repeatedly solving the full optimization problem. We adopt MADRL because it supports decentralized decisions under local observations while allowing the agents to learn coordinated behavior from a common objective during centralized training. Accordingly,

³For example, in α -fairness with $\alpha \rightarrow \infty$, unfairness is zero since all throughputs are equal.

⁴As a side note, semantic awareness can potentially lead to unfairness as economic advantages tend to favor users with co-related information.

²Gurobi Optimizer, <https://www.gurobi.com>.

we propose PRISM (Protocol for Redundancy Identification in Semantic Multiple-access), in which one DRL agent controls the MAC decisions of each UE. PRISM explicitly exposes self and assisted semantic transmissions to the agents, represents variable-length transmissions as macro-actions, and shapes the common reward according to either the fairness-utilization or energy-efficiency objective. Thus, it preserves the objectives of P_1 and P_2 while providing an online, decentralized execution mechanism.

In this context, agent i 's macro-action a_i is defined as a tuple $\langle I_{a_i}, \pi_{a_i}, \beta_{a_i} \rangle$, representing the initiation set, macro-action low-level policy, and termination set, respectively. In our problems, a new decision on channel sensing or packet transmission (denoted by the variable r in the optimization problems) is made immediately after the preceding macro-action is completed. During the multi-time slot transmission process, the remaining portion of the packet is incrementally reduced until the transmission is complete (i.e., when the variable z reaches zero). For simplicity, we will refer to macro-actions as actions in the remainder of this paper. Furthermore, we define the set of active UEs at time slot t , denoted as $\mathbb{N}_+^t$, as the subset of UEs that have completed their actions and are ready for new transmissions. Formally, this is represented as: $\mathbb{N}_+^t = \{i \in \mathbb{N} \mid \beta_{a_i} = 1\}$. Non-active UEs are excluded from the joint action, state space, and reward calculations, as suggested by Marchesini *et al.* [51].

Traffic is assumed to be saturated, in the sense that semantic content is always available at every UE, but it is generated on demand rather than queued: UE i invokes its semantic encoder at the instant it initiates a transmission action, and the resulting descriptor supersedes any older unsent one, so no backlog is maintained. A sensing action neither invokes the encoder nor alters the buffered descriptor. All segments associated with the current descriptor are carried in a single packet, as assumed in (1), and a collided packet is not retransmitted. Finally, a UE neither observes nor assumes a neighbor's current payload; semantic overlap is resolved at the SBS after reception. In the following sections, we will elaborate on the action and state spaces, reward mechanisms, and the architectural framework of the learning process.

1) *Action Space*: The action space for active UEs at time slot t is specified in (18). In this set, the action of u_i can take one of two forms: $a_i : (0, c_i)$, which indicates sensing channel c_i , or $a_i : (r_i > 0, c_i)$, which denotes the transmission of a packet with length r_i on channel c_i .

$$\mathcal{A}^t = \left\{ \left\{ a_i : (r_i, c_i) \mid \begin{array}{l} r_i \in \{0, \dots, R_{max}\}, \\ c_i \in \{1, \dots, \mathcal{C}\} \end{array} \right\} \mid i \in \mathbb{N}_+^t \right\} \quad (18)$$

Here, $a_i : (0, c_i)$ is an optional one-slot carrier-sensing action and is not mandatory before transmission. It reports only whether c_i is busy or idle and reveals neither the transmitter nor its payload; a busy result therefore does not imply that a semantically related UE is transmitting. Moreover, the channels are semantically agnostic and semantic relationships are represented by $\mathbb{A}$, not by carrier assignment.

2) *State Space*: The state space consists of five distinct components. To provide UEs with real-time fairness information, we incorporate the Delay to Last Successful Transmission

(D2LT) metric [15] 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 next 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 two other state elements called *self transmission* and *assisted transmission* denoted as $\ddot{o}_i \in \mathbb{R}_{\geq 0}$ and $\ddot{o}_i \in \mathbb{R}_{\geq 0}$, respectively. These variables are defined in the next section.

Given that the final element within the state space pertains to UE actions, the state of each UE comprises the most recent $\mathcal{H}_S$ (D2LT, observation, self and assisted transmissions, action) tuples. Additionally, the state includes a vector Ψ_i^t that averages the observation tuples over a longer period, providing a broader context for the agent⁵. Thus, the global system state at time slot t is defined as follows:

$$\begin{aligned} \bar{\mathbf{H}}_{L,i}^t &= \{t - \tau \mid \forall 1 \leq \tau : (i \in \mathbb{N}_+^{t-\tau}) \wedge (|\bar{\mathbf{H}}_{L,i}^t| = \mathcal{H}_L)\} \\ \bar{\mathbf{H}}_{S,i}^t &= \{t - \tau \mid \forall 1 \leq \tau : (i \in \mathbb{N}_+^{t-\tau}) \wedge (|\bar{\mathbf{H}}_{S,i}^t| = \mathcal{H}_S)\} \\ O_i^t &= (\bar{v}_i^t, \dot{o}_i^t, \ddot{o}_i^t, \ddot{o}_i^t, a_i^t) \\ \Psi_i^t &= \text{AVG} \left\{ O_i^q \mid q \in \bar{\mathbf{H}}_{L,i}^t \right\} \\ S^t &= \left\{ s_i^t : \left\{ O_i^h \mid h \in \bar{\mathbf{H}}_{S,i}^t \right\} \cup \Psi_i^t \mid i \in \mathbb{N}_+^t \right\} \end{aligned} \quad (19)$$

In the above equation, AVG denotes the average operator, $\bar{\mathbf{H}}_{L,i}^t$ represents the long-term history, encompassing $\mathcal{H}_L$ time steps prior to t during which UE i decided to transmit or sense, and $\bar{\mathbf{H}}_{S,i}^t$ defines a short-term history with the same definition, encompassing $\mathcal{H}_S$ time steps.

3) *Self and Assisted Transmissions*: At each time slot, the *self transmission* of u_i (denoted with $\ddot{o}_i^t$) determines how much u_i is involved in the transmission of its related segments. Similarly, the assisted transmission of u_i (denoted with $\ddot{o}_i^t$) is the normalized summation of its segments transmitted by other UEs. In practice, calculating these metrics is not straightforward, since UEs with shared segments may not be synchronous. We first assume that the SBS only decodes a segment from the largest packets received containing that segment. The reason is that the information of shorter packets is embedded in the largest packet. Furthermore, by defining D2LT metrics for segments (similar to users' D2LTs with the same interpretation, denoted with δ_k^t), The SBS calculates ν_k^t (Eq. (20)), which represents the portion of recently successfully received packets for segment k that do not intersect with previously received packets. This value indicates the contribution of users who have successfully completed their transmission for this segment. From this, self and assisted transmissions can be determined. In the following equations, $\hat{r}_i^t$ signifies the size of the packet sent by u_i successfully received

⁵In practice, we utilize the one-hot vector representation of observations and actions as it is more compatible with learning algorithms.

by SBS at time t .

$$\nu_k^t = \min \left\{ \delta_k^t, \max_j \{ \hat{r}_j^t \} \right\} \quad (20)$$

$$\ddot{o}_i^t = \min \left\{ \frac{\sum_{k \in \mathbb{K}} (a_{i,k}^t \nu_k^t)}{\sum_{k \in \mathbb{K}} a_{i,k}^t}, \hat{r}_i^t \right\} \quad (21)$$

$$\ddot{o}_i^t = \frac{\sum_{k \in \mathbb{K}} (a_{i,k}^t \nu_k^t)}{\sum_{k \in \mathbb{K}} a_{i,k}^t} - \ddot{o}_i^t \quad (22)$$

As a simple illustrative example, suppose that u_1 and u_2 are associated with a shared segment k and successfully finish transmission of their packets with sizes 4 and 3 at time slot t (clearly on different channels), and the segment D2LT is 2, indicating that 2 time slots have elapsed since the SBS received segment k . In this case, $\nu_k^t = \min \{2, \max\{4, 3\}\} = 2$, indicating that only 2 time slots of the received packets from u_1 and u_2 uniquely contributed by these UEs to deliver segment k to the SBS. Therefore, $\ddot{o}_1^t = \ddot{o}_2^t = 2$ and $\ddot{o}_1^t = \ddot{o}_2^t = 0$. Now assume that u_2 was not successful in its transmission. In this case, $\ddot{o}_1^t = \ddot{o}_1^t = 2$ and $\ddot{o}_2^t = \ddot{o}_2^t = 0$. Finally, assume that the D2LT is 10, meaning that segment k has not been seen by the SBS recently. In this case, $\nu_k^t = \min \{10, \max\{4, 3\}\} = 4$, indicating that all packets can be considered as unique contributions. Therefore, $\ddot{o}_1^t = 4$ and $\ddot{o}_2^t = 3$, resulting in $\ddot{o}_1^t = 0$ and $\ddot{o}_2^t = 1$.

It should be noted that a UE is not given its neighbors' data, actions, or association rows. The sensing outcome $\ddot{o}_i^t \in \{B, I\}$ is obtained locally. After each completed macro-action, the SBS uses $\mathbb{A}$ and δ_k^t to compute $\ddot{o}_i^t$ and $\ddot{o}_i^t$ and returns them with $\bar{v}_i^t$; after a transmission, it also returns $\ddot{o}_i^t \in \{S, C\}$. The policy therefore estimates semantic contribution from the history in (19), rather than receiving it in advance. This feedback is lightweight and downlink-only: one short control message per completed macro-action rather than per time slot, which terminates the macro-action where an acknowledgment would and can thus be piggybacked on it [13], [15], [33], without consuming the $\mathcal{C}$ contended uplink channels. We assume this control feedback is error-free and available before the next decision; therefore delay, loss, and quantization of feedback values are outside the present model.

4) *Reward*: Since we consider a cooperative scenario in which all UEs aim to maximize the global network objective function, the system reward is defined as the average of user rewards, as outlined in (23). In this calculation, the reward of user i at time step t is determined by the selected action and the resulting system state. For a successful action ($\ddot{o}_i^t = S$), the user receives the normalized utility $r_i^t/R_{\max}$, reduced by a strategy-dependent penalty. Under $\mathcal{F}_\alpha^t$, the term $\alpha \max_j \{ \bar{v}_j^t \}$ penalizes the maximum constraint violation. Under $\mathcal{E}_{\zeta, d_1, d_2}^t$, the penalty ζd_1 accounts for the corresponding cost or deviation. A collision or unsuccessful action ($\ddot{o}_i^t = C$) results in a negative normalized reward, whereas no reward is assigned when $\ddot{o}_i^t \in \{B, I\}$. Accordingly, the reward function is defined as:

$$\rho_i^t = \begin{cases} r_i^t/R_{\max} - \alpha \max_j \bar{v}_j^t & \text{if } \ddot{o}_i^t = S \\ & \wedge (\mathcal{U}^t = \mathcal{F}_\alpha^t) \\ r_i^t/R_{\max} - \zeta d_1 & \text{if } \ddot{o}_i^t = S \wedge \\ & (\mathcal{U}^t = \mathcal{E}_{\zeta, d_1, d_2}^t) \\ -r_i^t/R_{\max} & \text{if } \ddot{o}_i^t = C \\ 0 & \text{if } \ddot{o}_i^t \in \{B, I\} \end{cases}$$

$$\rho^t = \frac{1}{N} \sum_{i \in \mathbb{N}} \rho_i^t \quad (23)$$

It should be emphasized that a collided macro-action is re-decided rather than retransmitted. Following the collision penalty $-r_i^t/R_{\max}$ defined above, the next decision is a fresh selection over the entire action space in (18), so the UE may re-attempt immediately, sense first, switch channel, or shorten its packet, and the timing, channel, and length of a retry are learned from (19) instead of being fixed by an ARQ timer and a backoff rule. Complementarily, since a segment is delivered whenever any associated UE succeeds, per (1), spatial semantic redundancy acts as an implicit repetition code across UEs rather than across successive attempts of the same UE, which is precisely what the assisted throughput $\ddot{x}_i^t$ in (3) and the assisted transmission $\ddot{o}_i^t$ quantify. Explicit retransmission with soft combining would add a third, temporal form of redundancy and is deferred to future work.

5) *Training Process*: Our proposed methodology lies between the Value Decomposition Network (VDN) framework [52] and the QPLEX algorithm [53], thus adhering 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, whereas 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.

To implement the VDN, we employ D3QL, as detailed in [17]. In this approach, the expected value of state-action pairs for each UE is represented as the combination of state and advantage values, specifically $Q_i = V_i + \text{adv}_i$, where Q_i is the Q-value, V_i is the state value, and adv_i is the advantage value. This methodology enhances the stability and performance of the learning process by clearly separating the state value from the advantage of specific actions, allowing for more accurate value estimation and decision-making. The overall action-value function is obtained by summing the individual UE action-value functions, enabling the SBS to train all agents using the common system reward defined in (23). This is expressed as:

Algorithm 1: PRISM

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 (4-16)
4   foreach  $i$  in  $\mathbb{N}_+^t$  do
5      $\iota \leftarrow$  generate a random number from  $[0 : 1]$ 
6     if  $\iota > \epsilon$  then
7        $(r_i, c_i) \leftarrow \arg\max_{\rho \in \mathbf{a}_i} Q(s_i^t, \rho, \mathcal{W}_i)$ 
8     else
9       select a random  $(r_i, c_i)$  from  $\mathbf{a}_i$ 
10    if  $r_i > 0$  then
11      start transmitting a packet with length  $r_i$ 
12      on channel  $c_i$ 
13    else
14      sense channel  $c_i$ 
15    obtain  $\hat{o}_i^t \in \{B, I\}$  locally after sensing; receive  $\bar{v}_i^t$ ,
       $\hat{o}_i^t$ ,  $\bar{o}_i^t$  and, after transmission,  $\hat{o}_i^t \in \{S, C\}$  from
      the SBS
16    construct  $s_i^{t+1}$ 
17  * Centralized Training (18-23)
18  calculate  $\rho^t$  according to (23)
19   $memory \leftarrow \{\rho^t\} \cup \{(s_i^t, (r_i, c_i), s_i^{t+1}) \mid i \in \mathbb{N}_+^t\}$ 
20  choose a batch of samples from  $memory$ 
21  train the agent according to (24)
22  if  $\epsilon > \tilde{\epsilon}$  then
23     $\epsilon \leftarrow \epsilon \epsilon'$ 

```

$$Q_{tot}(s^t, \mathcal{A}^t) = \sum_{i \in \mathbb{N}_+^t} Q_i(s_i^t, a_i^t) \quad (24)$$

Algorithm 1 summarizes the interaction between decentralized execution and centralized training. In Lines 4–16, each active UE applies an ϵ -greedy policy to either select the locally optimal action or explore a random sensing or transmission action. After executing the selected action, the UE receives the required feedback from the SBS and constructs its next local state. In Lines 18–23, the SBS computes the common system reward, stores the joint experience in the replay memory, samples a training batch, and updates the UE networks using the centralized value (Q_{tot}). The updated network parameters are then disseminated to the UEs for subsequent decentralized decision-making. The exploration probability is initialized as $\epsilon = 1$ and is updated according to ($\epsilon \leftarrow \epsilon \epsilon'$), where ϵ' is the multiplicative decay factor. This decay continues until ϵ reaches the minimum exploration probability ($\tilde{\epsilon}$).

V. EVALUATION

Two main scenarios were carried out to evaluate the performance of the PRISM scheme, one dedicated to α -fairness maximization (problem P_1), and another to energy efficiency maximization (problem P_2). The experiments are designed to answer three specific questions: (i) what is gained by adding semantic awareness to a modern MADRL-based MAC policy, (ii) how far a decentralized policy is from a centralized oracle under the same system model, and (iii) how the gain changes with the amount of inter-user semantic overlap. The association matrices used for the evaluations are synthetic stress-test instances for these questions; they are not presented

as a community benchmark or as a substitute for application-level datasets.

Our principal learning baseline is PRISM-lite, a literature-aligned semantic-oblivious MADRL policy. Contemporary learning-based MAC schemes commonly combine decentralized decisions from local histories with cooperative multi-agent learning and collision feedback [13], [15], [33]. PRISM-lite follows these general design principles and uses the same CTDE procedure, value-decomposition and double-dueling Q-learning backbone, history representation, channel–packet-length action space, reward, model capacity, and training budget as PRISM. Its only substantive difference is that the self- and assisted-semantic-transmission observations are removed. Thus, PRISM-lite is not intended to imitate a particular published protocol. It is a strong matched baseline that instantiates the established semantic-oblivious MADRL approach in the proposed asynchronous, variable-packet-length environment. Holding the learning and MAC components fixed avoids attributing differences caused by another algorithm, action space, or signaling model to semantic awareness.

CB denotes the model-specific collision-free centralized benchmark used in the evaluation, RND is a sanity check, and ORTH is an idealized semantic-orthogonality capability reference motivated by MDMA [9], O-MDMA [27], and ToDMA [29]. These published schemes show that signals carrying separable semantic content can, under method-specific encoders and decoders, share transmission resources. ORTH abstracts that common capability by allowing collision-free concurrent reception when the transmitted segments do not overlap. This deliberately favorable rule isolates the potential benefit of semantic orthogonality without claiming to reproduce a particular encoder, decoder, modality, or reconstruction metric. Accordingly, ORTH is not a measured implementation or formal upper bound of those methods; it is a best-case, model-level reference against which the distinct benefit of exploiting semantic similarity can be assessed. Optimal throughputs are derived from a centralized scheduler that allocates time slots and channels based on optimal solutions computed by the Gurobi optimizer.

In both scenarios, we set $\mathcal{N} = 8$, $\mathcal{C} = 4$, $R_{max} = 5$, and $\mathcal{T} = 30000$. Notably, across all experiments, a 10% probability of channel failure per time frame is incorporated to preserve the realism of wireless network assumptions. However, the dominant performance limitation in this study is due to collisions among users. Furthermore, similar to investigating optimal solutions, to demonstrate the impact of semantic sharing among UEs, we define and utilize a set of UE-segment association square matrices $A_{\mathcal{N},q}$ in both scenarios. In this section, $A_{\mathcal{N},q}$ translates to a non-diagonal relation that occurs with probability q , while a diagonal relation occurs with probability 1. Table IV contains hyperparameters and configurations of simulations. Notably, our three-layer fully-connected neural network has a Multiply-ACCumulate (MAC) value of $\approx 45k$, indicating its status as a comparatively compact model. On our simulation platform⁶, each training run

⁶An NVIDIA RTX 4000 Ada Generation (20 GB) and an Intel(R) Core(TM) Ultra 9 285K (3.70 GHz) CPU

TABLE IV: Training Configuration.

Parameter	Value
Short-term history size ($\mathcal{H}_S$)	5 experiences
Long-term history size ($\mathcal{H}_L$)	20 experiences
Capacity of experience memory	1000 experiences
Batch size	64
Discount factor (γ)	0.9
Learning rate	0.001
Exploration parameters $\tilde{\epsilon}, \epsilon'$	0.005, 0.999
Approximator model	A fully-connected network with 256, 128 and 64 units
Training frequency	Every 1 step
Target network update frequency	Every 50 steps

consisting of 30,000 time steps requires approximately 70–90 minutes to complete. This duration includes the forward pass, backward pass, and memory replay. Finally, each association matrix is fixed within a run and varied across runs; hence, the evaluation assumes stationary association statistics during learning. Instantaneous semantic contributions still vary with δ_k^t and access outcomes. Adaptation to within-run drift is left for future work.

A. α -fairness Maximization

In this scenario, we compare algorithms based on their ability to maximize the fairness-utilization trade-off, quantified using the α -fairness objective function defined in (4). Experiments are conducted for α values of 0 (i.e., sum-rate maximization), 0.1, 0.5, 1 (Proportional Fairness), 2, and $\alpha \rightarrow \infty$ (max-min fairness). In Fig. 6-(a), the evolution of α -fairness during the training procedure of 8 UEs over time is depicted for $\alpha = 0.1$ on the left. On the right, the subfigure shows the all-time average self and assisted throughputs for each UE. The results compare $A_{8,0.3}$ (little shared semantics among users) and $A_{8,0.9}$ (high rate of semantic sharing among users). When there is little shared semantic information to exploit (e.g., users' non-overlapping observations), there is less benefit in incurring the additional complexity of semantic orchestration. On the bright side, when users' streams are highly interdependent, semantic-aware multiple access can approach optimal performance in both scenarios.

Aggregated results for all values of α and association matrices are presented in Fig. 7-(a), illustrating that PRISM outperforms both random and PRISM-lite approaches, approaching optimal performance in many scenarios. It also outperforms the upper bound of orthogonal approaches in environments with a high shared semantics rate. For $\alpha = 0$ (sum-rate maximization), the focus is on maximizing total throughput without considering fairness, potentially allowing some UEs to monopolize the channel. As α increases to 0.1 and beyond, the trade-off shifts towards more equitable throughput distribution among UEs. The comparison between $A_{8,0.1}$ and $A_{8,0.9}$ demonstrates that higher rates of semantic sharing lead to better performance in terms of both self and assisted throughputs.

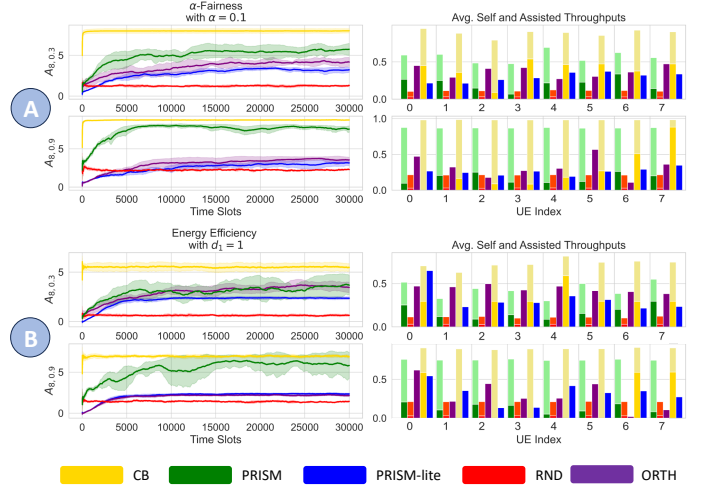


Fig. 6: Objective functions over time and all-time average UE throughputs, encompassing selected configurations of experiments A and B, comparing CB, PRISM, PRISM-lite, RND, and ORTH. Each line represents the average of five executions with random channel impairments, with shaded areas indicating the standard deviation. It is noteworthy that, when present, the lighter segment of each UE throughput signifies the assisted throughput for that particular UE.

The results suggest that PRISM successfully balances fairness and utilization, with greater capability observed for lower α values. This is likely due to the inherently cooperative nature of fairness, making it challenging to achieve in a distributed multiple-access setup. The incorporation of semantic sharing among UEs significantly enhances the system's performance, especially in terms of fairness and utilization trade-offs. This highlights PRISM's capability to manage resources efficiently in a distributed multiple-access environment, demonstrating superior performance and approaching optimal outcomes.

The comparison with ORTH further illustrates the difference between semantic orthogonality and semantic similarity. ORTH can be competitive when q is small because non-overlapping segments are common, making its idealized collision-free concurrent reception assumption favorable. As q grows, fewer transmissions satisfy the non-overlap condition, while PRISM benefits directly from the increased semantic commonality.

B. Energy Efficiency Maximization

In this scenario, we evaluate the algorithms based on their ability to maximize overall energy efficiency, as defined in (6). The parameters ζ and d_2 are set to 0.1 and 0.01, respectively. However, d_1 is assigned two values: 0 to exclude semantic transcoding costs, and 1 to include them. More detailed results are illustrated in Fig. 6-(b) for $d_1 = 1$. The subfigure on the left shows the evolution of energy efficiency during the training procedure of 8 UEs over time, while the right side displays the all-time average self and assisted throughputs for each UE. The results compare $A_{8,0.3}$ and $A_{8,0.9}$, similar to the previous scenario. Aggregated results for all values of d_1 and association matrices are depicted in Fig. 7-(b).

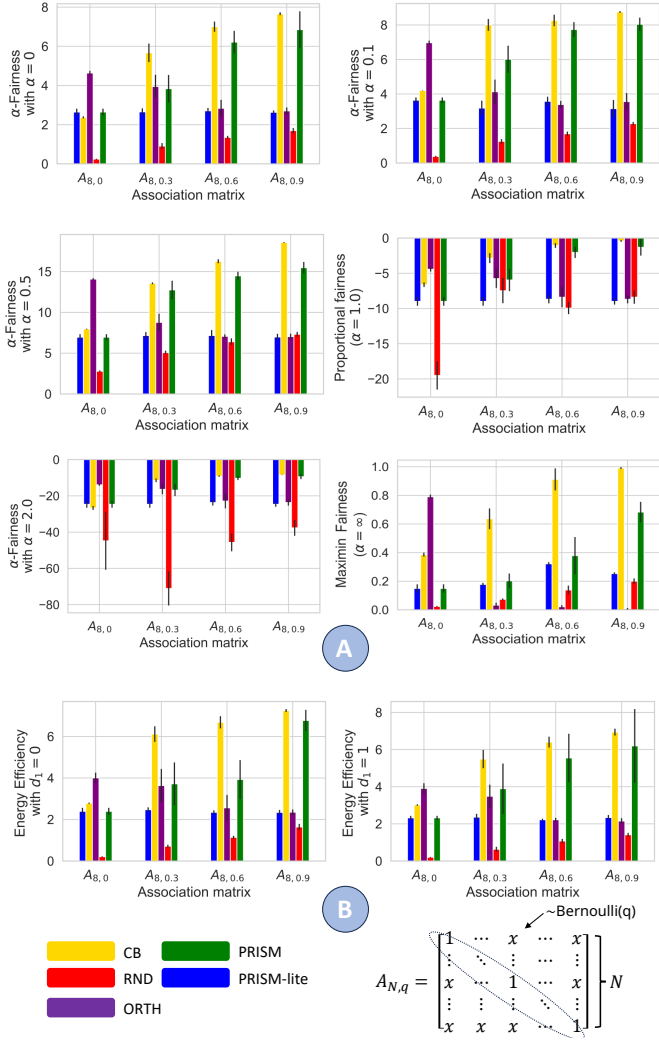


Fig. 7: Average of the last 1000 time slots of objective functions, encompassing experiments A and B, comparing CB, PRISM, PRISM-lite, RND, and ORTH. The reported results are based on averaging the outcomes of five rounds of simulations with random channel impairments, with error bars indicating the standard deviation.

The figures illustrate that PRISM outperforms both random and PRISM-lite approaches, approaching optimal performance in terms of energy efficiency. For both $d_1 = 0$ and $d_1 = 1$, the energy efficiency shows significant improvements, indicating the effectiveness of PRISM in managing energy consumption, even when semantic transcoding costs are included. The comparison between different association matrices demonstrates that higher semantic sharing rates lead to improved performance. These results highlight PRISM's capability to manage resources efficiently in a distributed multiple-access environment, particularly in terms of energy efficiency. Moreover, compared to orthogonal approaches, PRISM performs better in scenarios with higher semantic sharing rates, promising that these approaches can be used complementarily.

It is worth noting that the two scenarios optimize and report $\mathcal{F}_\alpha^t$ and $\mathcal{E}_{\zeta, d_1, d_2}^t$ separately; they do not cross-evaluate

each trained policy. Optimizing one objective therefore does not guarantee the other. Nevertheless, the two objectives are aligned in the direction that matters here: both reward the same utilization term and differ only in the penalized quantity, i.e., throughput dispersion under $\mathcal{F}_\alpha^t$ versus transmission and inference energy under $\mathcal{E}_{\zeta, d_1, d_2}^t$, while assisted throughput relieves both penalties at once, as it raises a UE's throughput with no additional energy cost and lifts precisely the UEs that did not transmit successfully. Consistently, the ordering of the schemes and the growth of the semantic gain with q reported above were observed under both objectives. Cross-evaluation and a scalarized fairness–energy Pareto analysis are left for future work.

C. Sensitivity to the Number of UEs and Channels

Fig. 8 evaluates whether the observed behavior persists when the network size changes. The top row varies the number of UEs, and the bottom row varies the number of channels. The left column reports energy efficiency with $d_1 = 1$, while the right column reports α -fairness with $\alpha = 0.5$.

As the number of UEs increases, PRISM benefits from the larger pool of potentially related semantic segments. The widening performance gap between PRISM and PRISM-lite across both objectives demonstrates that semantic awareness becomes increasingly valuable under heavier contention, rather than constituting an advantage limited to small-scale networks. Moreover, the evaluation across varying numbers of users reveals that ORTH improves at a slower rate than PRISM, since its gains depend on semantic non-overlap, whereas PRISM can exploit the increasing number of overlapping segments. Increasing the number of channels reduces the occurrence of collisions, thereby improving the performance of PRISM-lite and RND, particularly in terms of α -fairness. Nevertheless, PRISM remains the strongest decentralized method, as it jointly benefits from the collision reduction enabled by additional channels and the assistance provided by semantic information. Regarding energy efficiency, the evaluation across varying numbers of channels does not exhibit strictly monotonic behavior at every point. In particular, when $d_1 = 1$, additional transmission opportunities may also incur further transcoding costs. This observation highlights that PRISM addresses a joint communication-computation trade-off, rather than merely maximizing the number of successful transmissions.

VI. CONCLUSION

In this paper, we presented a novel framework for distributed multiple access that capitalizes on the revolutionary potential of semantic communication. By explicitly modeling the correlation among user source data—a critical feature of emerging 6G applications like the Metaverse and holographic telepresence—we established a foundation for optimizing spectral usage. We formulated two distinct optimization problems targeting the fairness-utilization trade-off and energy efficiency, deriving the PRISM (Protocol for Redundancy Identification in Semantic Multiple-access) scheme to solve them. Our evaluation demonstrates that integrating semantic intelligence

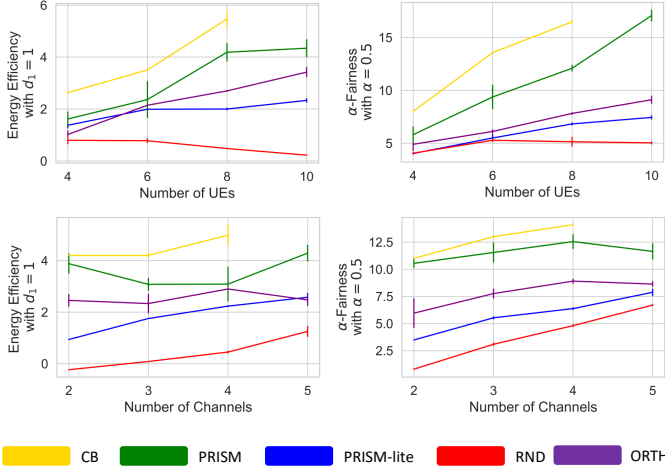


Fig. 8: Sensitivity of energy efficiency and α -fairness to the number of UEs and channels, comparing CB, PRISM, PRISM-lite, RND, and ORTH. The top row varies the number of UEs, and the bottom row varies the number of channels. The reported results are averaged over five executions with random channel impairments, with error bars indicating the standard deviation. For some cases, the optimal values could not be obtained within a reasonable computational time.

into distributed MAC decisions significantly enhances network sustainability compared to semantic-oblivious baselines.

Looking ahead, we aim to extend this framework to scenarios involving dynamic or imperfectly characterized association matrices, reflecting the fluid nature of real-time user data. This includes UEs joining or leaving the deployment, or changing their transmission and semantic-generation patterns during operation, which breaks the stationarity assumed in the Evaluation section and calls for online adaptation of the learned policies. We propose leveraging Continual Learning (CL) algorithms to handle these transitions, ensuring forward and backward transfer as user-segment associations evolve. Furthermore, we intend to investigate the granularity trade-off inherent in semantic segmentation: fine-grained segmentation offers precision at the cost of complexity, while coarse granulation risks association errors. We also intend to incorporate more realistic wireless channel models, including time-correlated fading and dynamic channel quality estimation, to better capture practical communication environments. In the same direction, we plan to relax the error-free control-feedback assumption of Section IV and to incorporate explicit retransmission mechanisms, such as HARQ with soft combining, into the macro-action structure via a redundancy-version selector and an SINR-based capture model, so that a partially decodable transmission is no longer treated as a total loss. Finally, we plan to validate the PRISM framework by deploying it within a real-world testbed for one of the identified 6G use cases.

ACKNOWLEDGMENT

The research work is supported in part by the Federal Ministry of Research, Technology, and Space (BMFTR), Germany, through the Project 6GEM+ under Grant 16KIS2411; and in part by the European Union's Horizon Europe research and innovation programme under the 6G-Path project (Grant No. 101139172).

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*, 2024.
- [2] H. Yu, M. Shokrnezhad *et al.*, "Towards 6G-Based Metaverse: Supporting Highly-Dynamic Deterministic Multi-User Extended Reality Services," *IEEE Network*, vol. 37, no. 4, pp. 30–38, 2023.
- [3] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Communications Surveys & Tutorials*, 2024.
- [4] P. Agheli, N. Pappas, and M. Kountouris, "Semantic filtering and source coding in distributed wireless monitoring systems," *IEEE Transactions on Communications*, 2024.
- [5] K. Wang, Z. Bao, K. Liu, H. Liang, C. Dong, X. Xu, and L. Li, "Adaptive Bitrate Video Semantic Increment Transmission System Based on Buffer and Semantic Importance," *IEEE Transactions on Network and Service Management*, vol. 22, no. 4, pp. 3586–3600, 2025.
- [6] A. R. Khan, H. U. Manzoor, R. N. Bin Rais, S. Hussain, L. Mohjazi, M. A. Imran, and A. Zoha, "Semantic-Aware Federated Blockage Prediction (SFBP) in Vision-Aided Next-Generation Wireless Network," *IEEE Transactions on Network and Service Management*, vol. 22, no. 2, pp. 1531–1543, 2025.
- [7] Y. Lin, C. Wu, J. Wu, L. Zhong, X. Chen, and Y. Ji, "Meta-networking: Beyond the Shannon limit with multi-faceted information," *IEEE Network*, vol. 37, no. 4, pp. 256–264, 2023.
- [8] E. C. Strinati, P. Di Lorenzo, V. Sciancalepore, A. Aijaz, M. Kountouris, D. Gündüz, P. Popovski, M. Sana, P. A. Stavrou, B. Soret *et al.*, "Goal-oriented and semantic communication in 6G AI-native networks: The 6G-GOALS approach," in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2024, pp. 1–6.
- [9] P. Zhang, X. Xu, C. Dong, K. Niu, H. Liang, Z. Liang, X. Qin, M. Sun, H. Chen, N. Ma *et al.*, "Model division multiple access for semantic communications," *Frontiers of Information Technology & Electronic Engineering*, pp. 1–12, 2023.
- [10] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, "Task-oriented multi-user semantic communications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 9, pp. 2584–2597, 2022.
- [11] W. Zhang, K. Bai, S. Zeadally, H. Zhang, H. Shao, H. Ma, and V. C. Leung, "DeepMA: End-to-end deep multiple access for wireless image transmission in semantic communication," *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 2, pp. 387–402, 2023.
- [12] S. Gronauer and K. Diepold, "Multi-Agent Deep Reinforcement Learning: A Survey," *Artificial Intelligence Review*, pp. 1–49, 2022.
- [13] 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.
- [14] 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.
- [15] Z. Guo, Z. Chen, P. Liu, J. Luo, X. Yang, and X. Sun, "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.
- [16] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3974–3983.
- [17] H. Mazandarani, M. Shokrnezhad, and T. Taleb, "A Semantic-Aware Multiple Access Scheme for Distributed, Dynamic 6G-Based Applications," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024.

- [18] —, “A Novel Multiple Access Scheme for Heterogeneous Wireless Communications Using Symmetry-Aware Continual Deep Reinforcement Learning,” *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 3, pp. 353–368, 2025.
- [19] H. Mazandarani, M. Shokrnezhad, T. Taleb, and R. Li, “Self-sustaining multiple access with continual deep reinforcement learning for dynamic metaverse applications,” in *2023 IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom)*. IEEE, 2023, pp. 65–70.
- [20] H. Mazandarani, M. Shokrnezhad, and T. Taleb, “Semantic-Aware Dynamic and Distributed Power Allocation: a Multi-UAV Area Coverage Use Case,” in *IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, 2025.
- [21] L. X. Nguyen, A. D. Raha, P. S. Aung, D. Niyato, Z. Han, and C. S. Hong, “A contemporary survey on semantic communications: Theory of mind, generative ai, and deep joint source-channel coding,” *arXiv preprint arXiv:2502.16468*, 2025.
- [22] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson *et al.*, “SAM 2: Segment Anything In Images and Videos,” *arXiv preprint arXiv:2408.00714*, 2024.
- [23] J. Park, Y. Oh, S. Kim, and Y.-S. Jeon, “Joint source-channel coding for channel-adaptive digital semantic communications,” *IEEE Transactions on Cognitive Communications and Networking*, 2024.
- [24] A. K. Sangaiah, J. Anandakrishnan, A. R. Devarapelly, M. L. A. B. Mohamad, G.-B. Bian, M. J. Alenazi, and S. A. AlQahtani, “R-UAV-Net: Enhanced YOLOv4 With Graph-Semantic Compression for Transformative UAV Sensing in Paddy Agronomy,” *IEEE Transactions on Cognitive Communications and Networking*, 2024.
- [25] T. M. Getu, G. Kaddoum, and M. Bennis, “A survey on goal-oriented semantic communication: Techniques, challenges, and future directions,” *IEEE Access*, 2024.
- [26] S. E. Trevlakakis, N. Pappas, and A.-A. A. Boulogeorgos, “Towards Natively Intelligent Semantic Communications and Networking,” *IEEE Open Journal of the Communications Society*, 2024.
- [27] H. Liang, K. Liu, X. Liu, H. Jiang, C. Dong, X. Xu, K. Niu, and P. Zhang, “Orthogonal Model Division Multiple Access,” *IEEE Transactions on Wireless Communications*, 2024.
- [28] R. Kassab, A. Destounis, D. Tsilimantous, and M. Debbah, “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.
- [29] L. Qiao, M. B. Mashhadi, Z. Gao, and D. Gündüz, “Token-Domain Multiple Access: Exploiting Semantic Orthogonality for Collision Mitigation,” *arXiv preprint arXiv:2502.06118*, 2025.
- [30] P. Agostini, J.-F. Chamberland, F. Clazzer, J. Dommel, G. Liva, A. Munari, K. Narayanan, Y. Polyanskiy, S. Stanczak, and Z. Utkovski, “Evolution of the 5G New Radio Two-Step Random Access towards 6G Unsourced MAC,” *arXiv preprint arXiv:2405.03348*, 2024.
- [31] M. L. Betalo, S. Leng, H. N. Abishu, A. M. Seid, M. Fakirah, A. Erbad, and M. Guizani, “Multi-Agent DRL-Based Energy Harvesting for Freshness of Data in UAV-Assisted Wireless Sensor Networks,” *IEEE Transactions on Network and Service Management*, vol. 21, no. 6, pp. 6527–6541, 2024.
- [32] T. Rashid, M. Samvelyan, C. S. De Witt, G. Farquhar, J. Foerster, and S. Whiteson, “Monotonic value function factorisation for deep multi-agent reinforcement learning,” *The Journal of Machine Learning Research*, vol. 21, no. 1, pp. 7234–7284, 2020.
- [33] L. Miuccio, S. Riolo, S. Samarakoon, M. Bennis, and D. Panno, “On Learning Generalized Wireless MAC Communication Protocols via a Feasible Multi-Agent Reinforcement Learning Framework,” *IEEE Transactions on Machine Learning in Communications and Networking*, 2024.
- [34] H. Ahmadi, M. Rahmani, S. B. Chetty, E. E. Tsiropoulou, H. Arslan, M. Debbah, and T. Quek, “Towards Sustainability in 6G and Beyond: Challenges and Opportunities of Open RAN,” *arXiv preprint arXiv:2503.08353*, 2025.
- [35] A. Galis, “Future Sustainable Internet Energy-Defined Networking,” *Future Internet*, vol. 16, no. 1, p. 23, 2024.
- [36] L. X. Nguyen, Y. L. Tun, Y. K. Tun, M. N. Nguyen, C. Zhang, Z. Han, and C. S. Hong, “Swin transformer-based dynamic semantic communication for multi-user with different computing capacity,” *IEEE Transactions on Vehicular Technology*, 2024.
- [37] A. Albaseer and M. Abdallah, “Tailoring Semantic Communication at Network Edge: A Novel Approach Using Dynamic Knowledge Distillation,” *arXiv preprint arXiv:2401.10214*, 2024.
- [38] L. X. Nguyen, H. Q. Le, Y. L. Tun, P. S. Aung, Y. K. Tun, Z. Han, and C. S. Hong, “An Efficient Federated Learning Framework for Training Semantic Communication Systems,” *IEEE Transactions on Vehicular Technology*, 2024.
- [39] H. R. Mazandarani and S. Khorsandi, “Energy-aware Multiple Access Using Deep Reinforcement Learning,” in *2021 29th Iranian Conference on Electrical Engineering (ICEE)*. IEEE, 2021, pp. 521–525.
- [40] M. Vuran and I. Akyildiz, “Spatial correlation-based collaborative medium access control in wireless sensor networks,” *IEEE/ACM Transactions on Networking*, vol. 14, no. 2, pp. 316–329, 2006.
- [41] L. Chen, Y. Jiang, Z. Qiu, D. Zhu, Z. Liu, and Z. Tang, “Toward Energy-Saving Deployment in Large-Scale Heterogeneous Wireless Sensor Networks for Q-Coverage and C-Connectivity: An Efficient Parallel Framework,” *IEEE Transactions on Network and Service Management*, vol. 23, pp. 1242–1259, 2026.
- [42] Y. Yu, S. C. Liew, and T. Wang, “Non-uniform time-step deep q-network for carrier-sense multiple access in heterogeneous wireless networks,” *IEEE Transactions on Mobile Computing*, vol. 20, no. 9, pp. 2848–2861, 2021.
- [43] Yang, Erhe and Yu, Zhiwen and Zhang, Yao and Cui, Helei and Huang, Zhaoxiang and Wang, Hui and Ren, Jiaju and Guo, Bin, “Joint Semantic Extraction and Resource Optimization in Communication-Efficient UAV Crowd Sensing,” *IEEE Transactions on Network and Service Management*, vol. 22, no. 6, pp. 5900–5914, 2025.
- [44] Q. Pan, H. Tong, J. Lv, T. Luo, Z. Zhang, C. Yin, and J. Li, “Image segmentation semantic communication over internet of vehicles,” in *2023 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2023, pp. 1–6.
- [45] M. K. Abdel-Aziz, M. S. Elbamby, S. Samarakoon, and M. Bennis, “Cooperative Multi-Agent Learning for Navigation via Structured State Abstraction,” *IEEE Transactions on Communications*, 2024.
- [46] X. Lin, L. Kundu, C. Dick, E. Obiodu, T. Mostak, and M. Flaxman, “6G digital twin networks: From theory to practice,” *IEEE Communications Magazine*, 2023.
- [47] H. Du, R. Zhang, D. Niyato, J. Kang, Z. Xiong, D. I. Kim, X. S. Shen, and H. V. Poor, “Exploring collaborative distributed diffusion-based AI-generated content (AIGC) in wireless networks,” *IEEE Network*, 2023.
- [48] V.-P. Bui, T. Q. Dinh, I. Leyva-Mayorga, S. R. Pandey, E. Lagunas, and P. Popovski, “Semantic Image Encoding and Communication for Earth Observation with LEO Satellites,” *IEEE Transactions on Cognitive Communications and Networking*, 2024.
- [49] K. Eriksson, D. Estep, C. Johnson, K. Eriksson, D. Estep, and C. Johnson, “Piecewise linear approximation,” *Applied Mathematics: Body and Soul: Volume 2: Integrals and Geometry in IR n*, pp. 741–753, 2004.
- [50] Y. Xiao, J. Hoffman, and C. Amato, “Macro-action-based deep multi-agent reinforcement learning,” in *Conference on Robot Learning*. PMLR, 2020, pp. 1146–1161.
- [51] E. Marchesini, Y. Xiao, and C. Amato, “Value Factorization for Asynchronous Multi-Agent Reinforcement Learning,” 2023.
- [52] P. Sunehag, G. Lever, A. Gruslys, W. M. Czarnecki, V. Zambaldi, M. Jaderberg, M. Lanctot, N. Sonnerat, J. Z. Leibo, K. Tuyls *et al.*, “Value-decomposition networks for cooperative multi-agent learning,” *arXiv preprint arXiv:1706.05296*, 2017.
- [53] J. Wang, Z. Ren, T. Liu, Y. Yu, and C. Zhang, “Qplex: Duplex dueling multi-agent q-learning,” *arXiv preprint arXiv:2008.01062*, 2020.