Semantic Aerial Open Radio Access Networks

Andrew Ashdown, Francesca Meneghello, and Francesco Restuccia

Institute for Intelligent Networked Systems, Northeastern University, United States

***Abstract*— The disaggregation and interoperability enabled by Open Radio Access Networks (RANs) is redefining the possible for NextG cellular standards like 5G and 6G. While this revolution started with traditional ground-based RANs, there are now efforts to deploy fully functional Open RANs on Unmanned Aerial Vehicle (UAV) swarms to provide aerial network coverage in remote locations where installing terrestrial RANs is not an option or is overly costly. Despite the promise of such endeavors, a key deployment challenge that remains to be addressed is the optimization of the wireless fronthaul connection between UAV-based Open Radio Units (O-RUs) and UAV-based Open Distributed Units (O-DUs). Specifically, meeting the Hybrid Automatic Repeat Request (HARQ) timing constraints may become infeasible when the number of tasks requested by users increases. In this paper, we tackle this issue from a semantic perspective, accounting for the specific users' task requirements. HARQ-imposed timing constraints are semantically alleviated by modifying both MAC-Layer HARQ and Radio Link Control (RLC)-mode retransmission functionalities. Our solution is a dynamic task-aware UAV-swarm-based Open RAN architecture entailing four O-DU robustness routes with differing RLC-mode and HARQ retransmission capabilities. Tasks are routed through one of these four O-DU routes based on semantic task requirements, UAV compute power, and wireless channel conditions. Our experimental results show up to a 51.0% reduction in HARQ failure rate with respect to a UAV-swarm-based Open RAN baseline with traditional HARQ and RLC retransmission functionalities.**

***Index Terms*—Open Radio Access Networks (O-RANs), Unmanned Aerial Vehicle (UAV), Hybrid Automatic Repeat Request (HARQ), Wireless Fronthaul.**

I. Introduction

The United States Presidential Memorandum *Winning the 6G Race* from December 2025 states that "The next generation of mobile communications networks (6G) will be foundational to the national security, foreign policy, and economic prosperity of the United States" [1]. Providing service continuity and reliability is a key aspect in this regard. As such, mobilizing the Radio Access Network (RAN) via Unmanned Aerial Vehicle (UAV) swarm support will allow for wireless coverage in tactical or civilian scenarios where terrestrial networks are compromised or nonexistent and where deploying new ground-based infrastructure is too costly or time-consuming [2]. This effort is supported by continuous investments in improving drone technology. For example, the White House Executive Order *Unleashing American Drone Dominance* from June 2025 states that "It is the policy of the United States to ensure continued American leadership in the development, commercialization, and export of [Unmanned Aircraft Systems]" [3].

From a communication technology point of view, a framework that can help in the deployment of aerial RANs is represented by Open RANs, which introduces unprecedented flexibility and customization opportunities into the cellular domain supported by the O-RAN Alliance. In particular, Open RAN's disaggregation and standardized interoperability are paving the way for components such as the Open Distributed Unit (O-DU) and Open Radio Unit (O-RU) to be deployed in different UAVs and connected wirelessly through a common set of interfaces.

Despite ongoing efforts in this direction [4]–[9], a critical roadblock to realizing functional UAV-swarm-based RANs is linked with the wireless fronthaul connection between the O-RU and O-DU components that are deployed on separate drones. Specifically, the wireless fronthaul timing constraints imposed by the Hybrid Automatic Repeat Request (HARQ) protocol require O-DU drones to have significant computing resources and high-performance radios for inter-swarm communications (i.e., communications between O-RUs and O-DUs). This becomes especially problematic as the number of networking tasks increases and/or when drone compute power is limited.

In this paper, we propose approaching the problem from a semantic communications perspective where User Equipment (UE) task requirements are used to map tasks to drone routes with the appropriate level of semantic robustness (in terms of retransmission and latency capabilities). Tasks with softer requirements are routed through O-DU drones that do not use HARQ retransmissions (removing HARQ alleviates compute utilization for those HARQ-less O-DU drones) which frees up O-DU drones with HARQ retransmissions enabled to serve tasks with stricter constraints.

Summary of Novel Contributions

- We develop a semantic aerial Open RAN base station architecture entailing four distinct O-DU routes with unique HARQ and Radio Link Control (RLC) retransmission characteristics.
- We develop O-DU Selection Algorithm (DUSA), a semantic task routing algorithm that is executed at a remote, ground-based Open Centralized Unit (O-CU). DUSA maps each UE task to an appropriate O-DU route based on O-DU route capabilities, UE task semantic requirements, drone compute resources, and wireless channel characteristics.
- We validate DUSA through a realistic python-based simulation that integrates experimental wireless channel data. The results show up to 51.0% reduction in HARQ failure rate with respect to a traditional Open RAN implementation.

The remainder of the paper is structured as follows: In Section II we discuss the related work; in Section III we outline the system design for our semantic aerial base station architecture; in Sections IV-V we showcase our simulation environment and results; Section VI concludes the work. Additionally, the simulation code, generated data, and plots

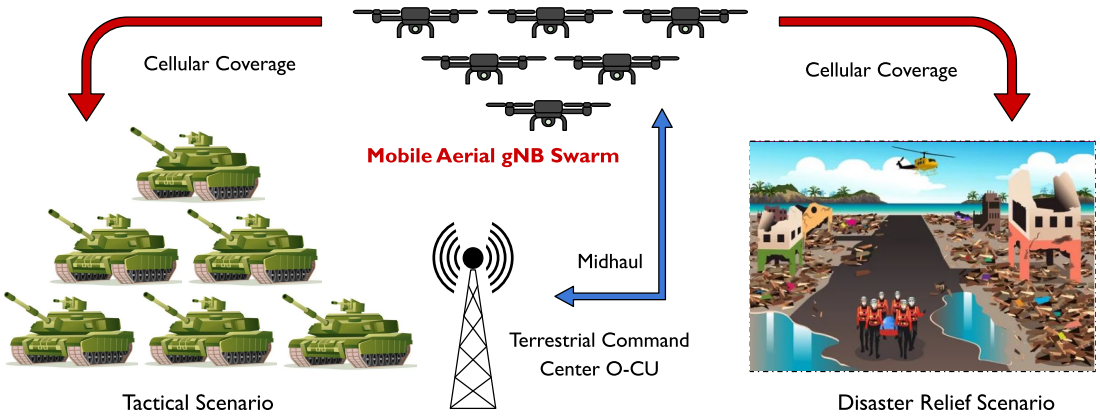

Fig. 1. Semantic Aerial Open RAN. In NextG networks drone swarms will function as aerial base stations to provide cellular connectivity in locations where physical infrastructure is unavailable (e.g., tactical scenarios or disaster relief situations).

for this paper can be found at https://github.com/Restuccia-Group/Drones-MILCOM.

II. Related Work

Recent research efforts are integrating and leveraging semantic concepts in Open RAN for different purposes such as for semantic message compression (see, e.g., [10]–[13]). However, to our knowledge no work focuses specifically on leveraging semantic concepts in the context of the Open RAN wireless fronthaul link, which poses serious latency challenges for an aerial Next Generation Node B (gNB) where O-RUs are separated spatially from their controlling O-DUs.

Several techniques have been proposed to overcome the HARQ-timing-related wireless fronthaul challenges in UAV-enabled base stations (e.g., [4]–[9]). However, we are not aware of any work that relaxes wireless fronthaul latency constraints by modifying the HARQ functionalities and changing the RLC mode. This represents a significant oversight because in many instances it is not feasible to meet the standardized HARQ timing constraints in a wireless fronthaul link. Hence, these constraints should be relaxed or removed if possible. Fronthaul compression strategies that seek to decrease the amount of data being sent over the wireless fronthaul link have been proposed in the literature [4] [14, Annex A] [5], [15], [16]. However, these research efforts present Low-Physical (PHY) layer compression techniques, and none of them explore HARQ or RLC mode adjustments as a means of wireless fronthaul compression. Thus, they are more concerned with reducing fronthaul throughput utilization and they fail to address the strict HARQ-related *latency* constraints.

Lastly, there are approaches in the literature that use semantic concepts to optimize the HARQ functionality (e.g., see [17]), but they do not do so to relax wireless fronthaul latency constraints; rather, they are concerned with determining semantically effective ways to do end-to-end HARQ (i.e., essentially end-to-end semantic retransmissions) and HARQ-imposed roundtrip latency constraints remain in place (as opposed to our approach where HARQ is entirely removed in some semantically strategic cases which eliminates the HARQ timing difficulties in those instances). Furthermore, the HARQ discussed in such papers is not really the 5G Medium Access Control (MAC) Layer HARQ we are referencing. They are using HARQ in a more abstract way for "*retransmit when needed*" at the application layer or some layer above the MAC. Thus, they do not relax the HARQ timing constraints in any way. As a result, they are of no use in addressing the HARQ timing challenges in wireless fronthaul links.

III. Semantic Aerial Open RAN Architecture

We consider an aerial gNB consisting of a tightly grouped swarm of UAVs $\mathcal{S}$ where each drone $\delta^i \in \mathcal{S}$ is equipped with a lightweight mmWave transceiver for inter-swarm communications. A single remote, ground-based O-CU implements the Packet Data Convergence Protocol (PDCP), Radio Resource Control (RRC), and Service Data Adaptation Protocol (SDAP) gNB protocol stack layers. The O-CU is connected via wire to the core network and wirelessly connected to the drone swarm $\mathcal{S}$. The drones serve as O-DUs and O-RUs as detailed next.

O-DUs are implemented in a subset of drones $\mathcal{S}_{DU} \subset \mathcal{S}$ that provide the High-PHY, MAC, and RLC gNB protocol stack layer functionalities (specifically there are four different O-DU drones in our framework). Each drone $\delta^i_{DU/\mu} \in \mathcal{S}_{DU}$ implements a different MAC/RLC combination defined by the index $\mu \in [0,\ 1,\ 2,\ 3]$. These four semantic O-DU routes are summarized in Figure 2 (top) and detailed in the following. We use *DU/0* and *DU/1* to indicate UAVs that use MAC with traditional HARQ and implement RLC Acknowledgment Mode (RLC-AM) and RLC Unacknowledgment Mode (RLC-UM), respectively; these are similar to traditional cellular networks in that they both enable HARQ retransmissions which results in strict timing constraints between the O-RU and O-DU components. In turn, these O-DU drones require significant computing resources to meet the tight HARQ-related latency requirements, especially as the ratio between drone compute power and the number of tasks routed through these O-DUs decreases. For UAVs identified as *DU/2* and *DU/3*, the MAC layer does not use HARQ (i.e., no HARQ processes are retransmitted at the MAC layer) and the RLC

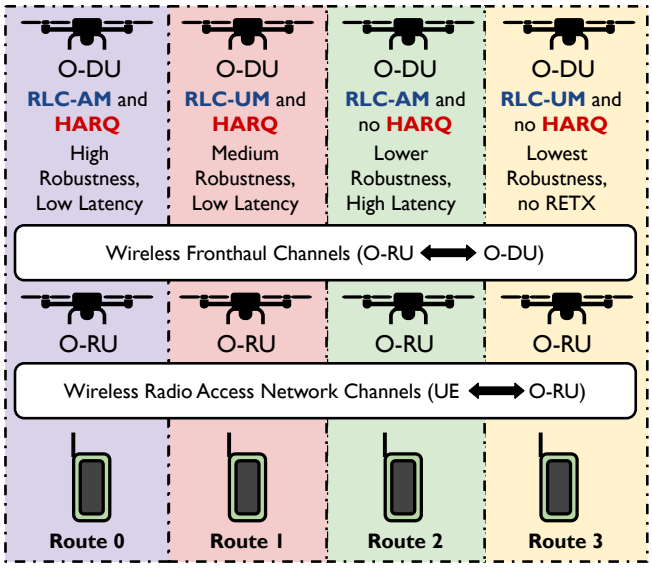

Fig. 2. Semantic O-DU Routes: *Route 0* implements both HARQ and RLC retransmissions which enables high-robustness, low latency retransmissions. *Route 1* implements only HARQ retransmissions which enables medium-robustness, low latency retransmissions. *Route 2* implements only RLC retransmissions which enables lower-robustness, high latency retransmissions. *Route 3* does not implement any retransmissions.

mode implemented is RLC-AM and RLC-UM, respectively. Thus, these O-DUs have lower robustness than the first two but they do not suffer as severely from compute max-out which is crucial for larger network sizes with many UE tasks.

Another set of drones, $\delta^i_{RU/\mu} \in \mathcal{S}_{RU} \subset \mathcal{S}$ function as O-RUs and they implement the Low-PHY and Radio Frequency (RF) gNB protocol stack layers; this represents the O-RAN Alliance's standardized 7.2x split configuration where the Low-PHY and High-PHY are split between O-RU and O-DU, respectively. Each RU drone is equipped with a sub-6 GHz transceiver for communication with ground-based UEs. The operation mode index μ for O-RUs has the same meaning as the one defined for the O-DUs (see Figure 2). The operation mode must be equal for every $\delta^i_{DU/\mu}$ / $\delta^i_{RU/\mu}$ pair. We refer to the $\delta^i_{DU/\mu}$ drone and the corresponding $\delta^i_{RU/\mu}$ drones as part of *Route* μ throughout this paper.

We consider a set $\mathcal{U}$ of ground-based UEs where each UE $u^j \in \mathcal{U}$ is equipped with a sub-6 GHz transceiver. UEs are grouped into two categories. UEs in the first category $u^j_{CLASS} \in \mathcal{U}_{CLASS} \subset \mathcal{U}$ offload a classification task to the gNB. Specifically, UEs in this category transmit images to be classified by the aerial gNB. The classification must meet a *minimum accuracy confidence level* $A^j_{min} \in \{0,\ 1\}$ and be completed within a *maximum latency* L^j_{max}. As classifying certain objects is inherently more challenging than classifying others, we use D^j to indicate the *semantic classification difficulty* involved in completing UE u^j_{CLASS}'s task. UEs in the second category $u^j_{COM} \in \mathcal{U}_{COM} \subset \mathcal{U}$ perform non-classification, bit-wise, communication tasks with *maximum Bit Error Rate (BER)* BER^j_{max} and *maximum latency* L^j_{max}. Each UE $u^j \in \mathcal{U}$ has the capability to switch its HARQ functionalities on or off depending on which O-DU drone is serving the user. A summary of system design symbols can be found in Table I.

TABLE I
SUMMARY OF THE SYMBOLS USED IN THE PAPER.

Symbol	Description
$\mathcal{S}$	UAV Swarm
δ^i	Drone i in swarm $\mathcal{S}$
$\mathcal{S}_{DU}$	The O-DU subset of UAV swarm $\mathcal{S}$
$\delta^i_{DU/\mu}$	O-DU drone i from O-DU route μ
$\mathcal{S}_{RU}$	The O-RU subset of UAV swarm $\mathcal{S}$
$\delta^i_{RU/\mu}$	O-RU drone i from O-DU route μ
$\mathcal{U}$	Set of ground-based UEs
u^j	UE j from set $\mathcal{U}$
$\mathcal{U}_{CLASS}$	Subset of UEs that execute classification tasks
u^j_{CLASS}	UE j executing classification task
$\mathcal{U}_{COM}$	Subset of UEs that execute communication tasks
u^j_{COM}	UE j executing communication task
t	Timestep
BER_{max}	Max Bit Error Rate
L_{max}	Max latency
A_{min}	Min classification accuracy confidence level
D	Semantic classification difficulty

A. Semantic 5QI

Traditional 5G networks use what is called a *5G Quality of Service Identifier (5QI)* to map UE tasks to Quality of Service (QoS) traffic flows. Each 5QI value defines a "one-to-one mapping to a standardized combination of 5G QoS characteristics" [18]. For our semantic-aware architecture, we define a semantic version of the 5QI, which we call *Semantic 5QI (Sem5QI)*. Each Sem5QI value in our system maps to a set of semantic task requirements which include *maximum BER* ($BER_{\max}$) and *maximum latency* ($L_{\max}$) for bit-wise communication tasks, and *minimum classification accuracy confidence level* (A_{min}), *inherent semantic classification difficulty* (D), and *maximum latency* ($L_{\max}$) for classification tasks. Each UE task is associated with a Sem5QI value which is used by the aerial gNB for semantic task routing.

TABLE II
EXAMPLES OF SEM5QI VALUES AND CORRESPONDING REQUIREMENTS.

Sem5QI Value	$BER_{\max}$	$A_{\min}$	D	$L_{\max}$
1	16%	N/A	N/A	5 ms
2	N/A	60%	70%	5 ms
3	N/A	68%	80%	5 ms
4	N/A	77%	90%	5 ms
5	N/A	85%	100%	5 ms

The first five entries of our Sem5QI table are reported in Table II. Sem5QI values with a valid value for *maximum BER* correspond to bit-wise communication tasks, while Sem5QI values with a valid value for the *minimum classification accuracy confidence level* correspond to classification tasks. Each consecutive group of five Sem5QI values form a Sem5QI group (the first group is shown in Table II and the second group would represent Sem5QI values 6-10, etc.). A Sem5QI group consists of one bit-wise Sem5QI value

and four corresponding classification Sem5QI values. The *minimum classification accuracy confidence level* for each classification Sem5QI value is obtained as

$$A_{\min} = D \cdot e^{-BER_{\max}}, \quad (1)$$

where D represents the *semantic classification difficulty* associated with that Sem5QI value, and $BER_{\max}$ is the *maximum BER* from the Sem5QI group's bit-wise Sem5QI value. For instance, $A_{\min}$ for Sem5QI value 3 in Table II is calculated as $A_{\min} = 0.8 \cdot e^{-0.16} = 68\%$. In the full table, $BER_{\max}$ / $A_{\min}$ / $L_{\max}$ range between 12% - 22% / 57% - 89% / 0.5ms - 34ms.

B. DUSA Semantic Task Routing Algorithms

Algorithms 1 and 2 outline our semantic task routing procedure, DUSA. Algorithm 1 is executed on the ground-based O-CU and determines the O-DU route through which each UE task is routed and when a UE should be handed over to a more semantically robust route. Algorithm 2 is executed within Algorithm 1 when a UE's task is repeatedly failing due to task constraint violations in its current O-DU route (see Algorithm 1, line 15) and it sets forth the procedure for handing a UE over to a more semantically robust route. Both algorithms are described in the following.

1) Algorithm 1: At start time, a Sem5QI value for each UE task is received by the O-CU, each task is assigned to an initial O-DU route using a predefined mapping (set by the O-CU) between the Sem5QI value and O-DU route, and all UEs are then served by their respective O-DU routes (lines 1-4 in Algorithm 1). Upon each task completion, a task success indicator vector (i.e., *TSI*) records which, if any, of the task requirements were violated. *TSI[0]* is for $BER_{\max}/A_{\min}$ for $COM/CLASS$ tasks and *TSI[1]* is for $L_{\max}$. If the task success indicator corresponding to *TSI[x]* failed, this is recorded in a *fails* vector in vector position x (lines 7-13). If any task fails more than $N_{\text{fail}}^{\max}$ times due to a task constraint violation, Algorithm 2 is executed (line 15) to handover that UE to a more semantically robust O-DU route; otherwise, the UE remains on its current route (lines 14-24).

2) Algorithm 2: This is executed every time any element of the *fails* vector exceeds $N_{\text{fail}}^{\max}$, which indicates the need for O-DU handover. The new route to which a task is handed over depends on which task constraint exceeded the max-fail threshold. For example, if the task failed due to a $BER_{\max}$ or $A_{\min}$ violation (i.e., $fails[0] \geq N_{\text{fail}}^{\max}$), the task will be handed over to an O-DU route with greater retransmission robustness to reduce BER or boost classification accuracy confidence level (lines 6-17 in Algorithm 2). The same reasoning holds for the $L_{\max}$ constraint; tasks with deadline violations will be handed over to routes with lower latency/delay (lines 6-12 and 18-26).

Algorithm 1 Semantic O-DU Selection Algorithm

1: At task timestep $t = 0$:
2: CU receives a Sem5QI value for each UE/task
3: Map Sem5QIs to DU routes
4: Initialize vector $fails \leftarrow [0, 0]$
5: **while** $t < \infty$ **do**
6: **for** every task j **do**
7: Receive TSI vector for previous j attempt
8: **if** $TSI[0] =$ FAIL **then**
9: $fails[0] \leftarrow fails[0] + 1$
10: **end if**
11: **if** $TSI[1] =$ FAIL **then**
12: $fails[1] \leftarrow fails[1] + 1$
13: **end if**
14: **if** $fails[0] \geq N_{\text{fail}}^{\max}$ **or** $fails[1] \geq N_{\text{fail}}^{\max}$ **then**
15: Execute Semantic O-DU Handover (see Alg. 2)
16: **if** $fails[0] \geq N_{\text{fail}}^{\max}$ **then**
17: $fails[0] \leftarrow 0$
18: **end if**
19: **if** $fails[1] \geq N_{\text{fail}}^{\max}$ **then**
20: $fails[1] \leftarrow 0$
21: **end if**
22: **end if**
23: **end for**
24: **end while**

Algorithm 2 Semantic O-DU Handover Algorithm

1: Route 0: RLC-AM and HARQ
2: Route 1: RLC-UM and HARQ
3: Route 2: RLC-AM and no HARQ
4: Route 3: RLC-UM and no HARQ
5: **for** every task j with a task constraint violation **do**
6: **if** $fails[0] > N_{fail}^{max}$ **and** $fails[1] > N_{fail}^{max}$ **then**
7: **if** $current\ route \geq 2$ **then**
8: Transfer task to $route\ 1$
9: **else**
10: Transfer task to $route\ 0$
11: **end if**
12: **end if**
13: **if** $fails[0] > N_{fail}^{max}$ **and** $fails[1] \leq N_{fail}^{max}$ **then**
14: **if** $current\ route - 1 \geq 0$ **then**
15: Transfer task to $current\ route - 1$
16: **end if**
17: **end if**
18: **if** $fails[0] \leq N_{fail}^{max}$ **and** $fails[1] > N_{fail}^{max}$ **then**
19: **if** $current\ route + 1 = route\ 2$ **then**
20: **if** $current\ route + 2 \leq 3$ **then**
21: Transfer task to $current\ route + 2$
22: **end if**
23: **else if** $current\ route + 1 \leq 3$ **then**
24: Transfer task to $current\ route + 1$
25: **end if**
26: **end if**
27: **end for**

IV. Experimental Evaluation

We created a realistic python-based UE-to-O-DU simulation using Sionna [19] to validate DUSA. The simulation consists of four logical O-DU routes corresponding to those in Figure 2, where each of the four drone-hosted O-DUs controls three drone-hosted O-RUs. We collected mmWave wireless channel samples using a 60 GHz Software Defined Radio (SDR)-based testbed [20] and use them for the O-DUs–O-RUs connections. We also experimentally collected sub6 GHz wireless channel samples using Wi-Fi devices implementing the IEEE 802.11ax standard [21] and use them to emulate the O-RUs–UEs connections. Random mmWave / sub6 GHz channels are pulled from a pool of 64 mmWave / 8,392 sub6

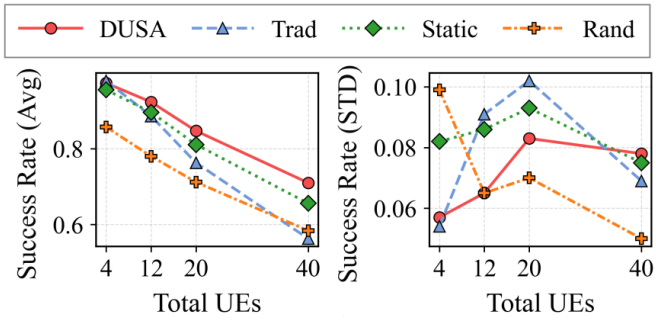

Fig. 3. Average overall UE task success rate vs. total UEs.

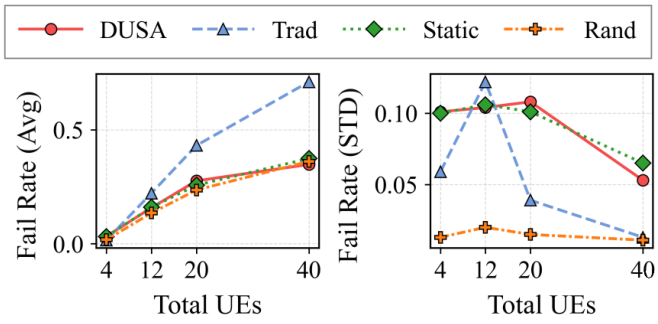

Fig. 4. Average overall fraction of transmissions where at least one HARQ round failure would have occurred vs. total UEs.

GHz channels at each new iteration (70 task transmissions for each UE make up one iteration and we conducted 70 iterations total). Every drone has a maximum computing budget; if it is exceeded, tasks are queued and can be delayed as a result. At the start of each iteration, every UE selects a Sem5QI value that corresponds to its semantic task requirements for that iteration. Note that our simulation does not deal with HARQ failures in any unique way; rather for each task transmission, we retrospectively record if a HARQ timing violation *would have* resulted given the latency. For the sake of space, we show only uplink experiments, but downlink performance would show the same trends; the only difference is the order in which the tasks traverse the wireless channels (uplink is Sub6 GHz $\rightarrow$ mmWave, downlink is mmWave $\rightarrow$ Sub6 GHz).

A. Evaluation Baselines

We consider three baselines to assess the performance of DUSA. The first baseline, named *Trad*, represents a traditional cellular network where all four O-DU routes use HARQ. Two of the routes in *Trad* use RLC-AM and two use RLC-UM. The second baseline, named *Static*, exhibits the same O-DU route structure as our DUSA approach, but once tasks are initially mapped to an O-DU route using the Sem5QI values, they are never handed over to another O-DU route. Thus, this baseline exhibits some of the semantic qualities of DUSA but is not dynamic in the same way. As a third baseline, named *Rand*, we consider the same O-DU route structure as our DUSA approach, but UE tasks are randomly assigned O-DU routes.

V. Experimental Results

As a first evaluation, we show that as the total number of UEs in the network increases, the baselines (particularly

TABLE III
Total Number of Tasks-per-Route for 40 UE Case

Technique	*Route 0*	*Route 1*	*Route 2*	*Route 3*
DUSA	41018	52993	46065	55924
Trad	50750	47950	50050	47250
Static	46760	51940	49700	47600
Rand	49079	49092	48789	49040

and most importantly *Tard*) struggle more to effectively meet semantic task requirements and/or they generate many more HARQ timing violations than DUSA. The results are reported in Figures 3-4. As shown, DUSA exhibits a 26.1% / 8.2% / 21.6% higher average task success rate (task success means all semantic task requirements are met) than *Trad* / *Static* / *Rand* (Figure 3) and a 51.0% / 7.0% / 3.6% lower HARQ failure rate (Figure 4) for the highest UE count case (i.e., 40 UEs). For *Trad*, more than half of task transmissions would have exhibited at least one HARQ timing violation (which can generate major network inefficiency and likely lowers task success rate even further) for the 40 UE case because drones are computationally overwhelmed and introduce computational delay that renders HARQ infeasible (we set the O-RU-to-O-DU round trip HARQ deadline to 0.5 ms in our simulations). Whereas HARQ is more feasible for DUSA, *Static*, and *Rand* because two of the routes in each of these do not have HARQ enabled. In our experiments, DUSA outperforms all baselines considering the task success rate because it better maps tasks to O-DU routes (Table III shows the number of tasks per route for each technique for the 40 UE case). Note that DUSA is not necessarily guaranteed to achieve higher task success rate than the baselines in all cases because this depends partially on the strictness of the specific UE tasks, wireless channel conditions, and drone compute (for instance see the 4 UE case where DUSA, *Trad*, and *Static* have similar task success). However, it should achieve at least comparable success rate in most instances. Further, even though DUSA is agnostic to HARQ failure, the task-to-O-DU mapping it chose for the 40 UE case routes more tasks away from HARQ-enabled *Route 0* (as shown in Table III) which results in a superior HARQ failure rate as well with respect to *Static* and *Rand*. As DUSA, *Static*, and *Rand* exhibit the same number of non-HARQ O-DU routes, DUSA is not necessarily guaranteed to achieve lower HARQ failure rate than those schemes in the same way it should typically beat *Trad* for compute-constrained cases (recall that HARQ is enabled for all *Trad* O-DUs).

As a second set of evaluations, we show the same type of performance as the max compute budget is decreased for the drones in the network. The results are presented in Figures 5-6. For each *Compute Config* in Figures 5-6, the swarm drones have a different max compute budget (compute budgets are internal simulation constants). Across our experiments, *Compute Config 1* corresponds to the least computational power per drone and *Compute Config 4* corresponds to the most computational power per drone. *DUSA* exhibits a 21.7% / 12.5% / 24.1% higher average task success rate than *Trad*

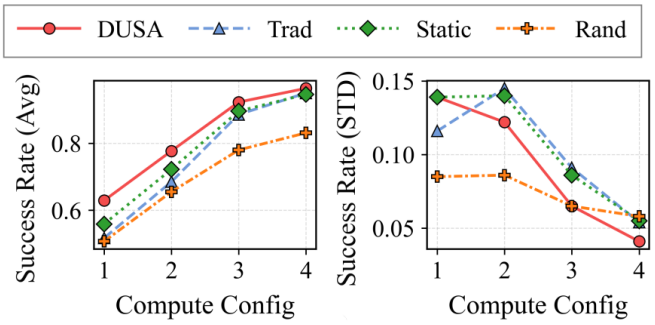

Fig. 5. Average UE task success rate vs. swarm computational budget configuration (x-axis).

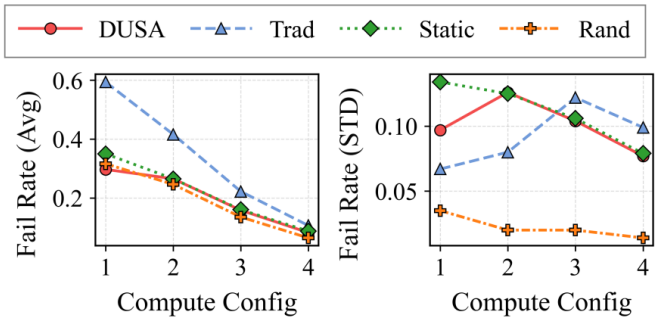

Fig. 6. Average fraction of transmissions where at least one HARQ failure would have occurred vs. swarm compute budget (x-axis).

/ *Static* / *Rand*, respectively, and a 50.0% / 15.1% / 6.0% lower HARQ failure rate (Figure 6) in the most compute-constrained configuration. Importantly, DUSA exhibits up to 50.0% less HARQ failure rate than traditional open RAN and simultaneously achieves up to 21.7% higher task success rate.

Key Experimental Takeaway *- Traditional HARQ is infeasible when drones are computationally overwhelmed as the drone-induced computational delay causes HARQ timing violations. Our experiments demonstrate that a semantic, task-oriented approach without mandatory HARQ can significantly decrease HARQ infeasibility (with up to 51.0% reduction) without impacting UE task success rate.*

VI. Conclusion

In this paper, we studied aerial Open RAN and proposed a strategy to address the wireless fronthaul challenge linked with the HARQ timing constraints by accounting for UE task semantics. We have shown that routing each task through a semantically robust O-DU route (i.e., a route where the HARQ and RLC retransmission functions, drone compute, and wireless channels match semantic task requirements) maintains or increases average overall UE task performance and decreases HARQ failure rate with respect to a UAV-swarm-based Open RAN baseline with traditional HARQ and RLC retransmission. We believe that our framework will pave the way toward efficient aerial Open RAN.

VII. Acknowledgments

This work was supported in part by NSF under Grant OAC-2530896, ARO under Grant W911NF261A165, ONR under Grants N00014-23-1-2221 and N00014-26-1-2376, and AFOSR under Grant FA9550-23-1-0261. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-23-F-0014 in the amount of $129,600.

References

[1] Trump, Donald J., "Winning the 6G Race." Presidential Memorandum, The White House, Dec. 2025.

[2] X. Gu and G. Zhang, "A survey on UAV-assisted wireless communications: Recent advances and future trends," *Computer Communications*, vol. 208, pp. 44–78, 2023.

[3] Trump, Donald J., "Unleashing American Drone Dominance." Executive Order No. 14307, The White House, June 2025.

[4] M. Jiang, J. Cezanne, A. Sampath, O. Shental, Q. Wu, O. Koymen, A. Bedewy, and J. Li, "Wireless fronthaul for 5G and future radio access networks: Challenges and enabling technologies," *IEEE Wireless Communications*, vol. 29, no. 2, pp. 108–114, 2022.

[5] D. Townend, R. Husbands, S. D. Walker, and A. Sutton, "Challenges and Opportunities in Wireless Fronthaul," *IEEE Access*, vol. 11, pp. 106607–106619, 2023.

[6] C. Zhang, E. Medeiros, I. Almeida, P.-E. Eriksson, S. Höst, and P. Ödling, "Robust Fronthaul-Aware Link Adaptation for UAV Communication With Lossy Wireless Fronthaul," *IEEE Communications Letters*, 2025.

[7] G. Fontanesi, A. Zhu, and H. Ahmadi, "Outage analysis for millimeter-wave fronthaul link of UAV-aided wireless networks," *IEEE Access*, vol. 8, pp. 111693–111706, 2020.

[8] M. T. Dabiri, M. Hasna, and W. Saad, "Downlink interference analysis of UAV-based mmWave fronthaul for small cell networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 5, pp. 5560–5575, 2022.

[9] C. Diaz-Vilor, A. Lozano, and H. Jafarkhani, "Cell-free UAV networks with wireless fronthaul: Analysis and optimization," *IEEE Transactions on Wireless Communications*, vol. 23, no. 3, pp. 2054–2069, 2023.

[10] Y. Sun, L. Zhang, L. Guo, J. Li, D. Niyato, and Y. Fang, "S-RAN: Semantic-Aware Radio Access Networks," 2024.

[11] P. Li and A. Aijaz, "Open RAN meets Semantic Communications: A Synergistic Match for Open, Intelligent, and Knowledge-Driven 6G," in *Proc. of IEEE CSCN*, pp. 87–93, 2023.

[12] E. Zeydan, L. Blanco, C. J. Vaca-Rubio, M. Caus, and K. Dev, "Semantic Non-Terrestrial Communications With Open RAN-Enabled 6G," *IEEE Communications Standards Magazine*, vol. 9, no. 4, pp. 72–78, 2025.

[13] C. Puligheddu, J. Ashdown, C. F. Chiasserini, and F. Restuccia, "SEM-O-RAN: Semantic and flexible O-RAN slicing for NextG edge-assisted mobile systems," in *Proc. of IEEE INFOCOM*, 2023.

[14] O-RAN Alliance, "O-RAN Fronthaul Control, User and Synchronization Plane," tech. rep., O-RAN Alliance, July 2019. O-RAN.WG4.CUS.0-v02.00.

[15] S. Lagén, X. Gelabert, A. Hansson, M. Requena, and L. Giupponi, "Fronthaul Compression Control for shared Fronthaul Access Networks," 2022.

[16] J. Cezanne, M. Jiang, O. Shental, A. M. Bedewy, A. Sampath, O. Koymen, and J. Li, "Design of Wireless Fronthaul With mmWave LOS-MIMO and Sample-Level Coding for O-RAN and Beyond 5G Systems," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 1893–1912, 2023.

[17] Y. Sheng, L. Liang, H. Ye, S. Jin, and G. Y. Li, "Semantic Communication for Cooperative Perception Using HARQ," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2025.

[18] 3GPP, "System Architecture for the 5G System (5GS)," Technical Specification TS 23.501, 3rd Generation Partnership Project (3GPP), July 2022. Release 17.

[19] J. Hoydis, S. Cammerer, F. Ait Aoudia, M. Nimier-David, L. Maggi, G. Marcus, A. Vem, and A. Keller, "Sionna," 2022. https://nvlabs.github.io/sionna/.

[20] K. F. Haque, F. Meneghello, K. Rumman, and F. Restuccia, "m3MIMO: An 8× 8 mmWave Multi-User MIMO Testbed for Wireless Research," in *Proc. of ACM WiNTECH*, 2024.

[21] F. Gringoli, M. Cominelli, A. Blanco, and J. Widmer, "AX-CSI: Enabling CSI extraction on commercial 802.11 ax Wi-Fi platforms," in *Proc. of ACM WiNTECH*, 2022.