Toward Secure and Efficient O-RAN Deployments: Secure Slicing xApp Use Case

Joshua Moore, Nisha Adhikari, Aly S. Abdalla, and Vuk Marojevic Dept. of Electrical and Computer Engineering, Mississippi State University, USA Emails: {jjm702; na731; asa298; vuk.marojevic}@msstate.edu

Abstract-The open radio access network (O-RAN) is recognized for its modularity and adaptability, facilitating swift responses to emerging applications and technological advancements. However, this architecture's disaggregated nature, coupled with support from various vendors, introduces new security challenges. This paper proposes an innovative approach to bolster the security of future O-RAN deployments by leveraging RAN slicing principles. Central to this security enhancement is the concept of secure slicing. We introduce SliceX, an xApp designed to safeguard RAN resources while ensuring strict throughput and latency requirements are met for legitimate users. Leveraging the open artificial intelligence cellular research (OAIC) platform, we observed that the network latency averages around ten microseconds in a default configuration without SliceX. The latency escalates to over seven seconds in the presence of a malicious user equipment (UE) flooding the network with requests. SliceX intervenes, restoring network latency to normal levels, with a maximum latency of approximately 2.3 s. These and other numerical findings presented in this paper affirm the tangible advantages of SliceX in mitigating security threats and ensuring that O-RAN deployments meet stringent performance requirements. Our research demonstrates the realworld effectiveness of secure slicing, making SliceX a valuable tool for military, government, and critical infrastructure operators reliant on public wireless communication networks to fulfill their security, resiliency, and performance objectives.

Index Terms—O-RAN, security, slicing, throughput, latency, OAIC, xApp.

I. Introduction

In today's ever-evolving digital landscape, the conventional configurations, setups, and deployments of cellular networks are in contrast with the demand for increased interconnectivity and technological convergence. The Open Radio Access Network (O-RAN) is a technology innovation poised to exert a transformative impact on the telecommunications domain by endowing it with unparalleled attributes of adaptability, interoperability, and cost-effectiveness [1].

Unlike the closed and proprietary network solutions, O-RAN shifts to a new paradigm characterized by a disaggregated architecture that decouples the hardware and software components [2]. O-RAN champions the principle of vendor neutrality, breaking free from the traditional constraints that tied network operators to specific vendors. Instead, it promotes open interfaces, virtualization, programmability, cloud integration, intelligent network management, and standardized protocols to enable smooth cooperation between hardware and software from various providers. This approach fosters healthy competition, catalyzes innovation within the industry, and facilitates network improvements, expansions, and the introduction of novel services, thereby enhancing

network flexibility, performance, and scalability [3]. The inherent disaggregated architecture of O-RAN provides network operators with the unique advantage of harnessing commercial off-the-shelf (COTS) hardware and virtualization technology. This strategic shift leads to a reduction in both capital and operational expenditures, contributing to cost-efficiency and sustainability in network management.

The adaptable and versatile architecture grants network operators the agility to promptly respond to emerging application or service needs [4]. This adaptability streamlines the seamless assimilation of novel features, functionalities, and network enhancements, thereby fortifying the network's resilience and ensuring its readiness to cater to future demands. It establishes a framework for the evolution of modern cellular network architectures toward next generation networks.

In an era characterized by interconnected networks, the demand for low-latency, secure, and high-throughput services has become paramount. This paradigm reflects the pervasive integration of digital technologies into nearly every facet of our lives, from communication and entertainment to business operations and critical infrastructure. Low-latency services refer to the ability of a network or system to minimize delays in transmitting data or processing requests. It ensures that actions, such as online gaming, video conferencing, vehicleto-vehicle communication, and financial transactions, occur with minimal delay [5]. As networks facilitate communication, collaboration, and data sharing on a global scale, the need for robust security measures is imperative [6]. Cyber threats and data breaches are ever-present risks. High-profile security incidents underscore the importance of safeguarding sensitive information, including personal data, financial transactions, and critical infrastructure. Secure services employ encryption, authentication, access controls, and monitoring to protect against unauthorized access, data breaches, and cyberattacks [7]. High-throughput services relate to the capacity of a network to efficiently handle a substantial volume of data or requests [8]. High-throughput networks can accommodate this data traffic without congestion or performance degradation. This is critical for bandwidth-intensive activities such as streaming 4K/8K video, cloud-based applications, and largescale data transfers [9].

This research explores the unfolding possibilities of O-RAN with a primary emphasis on RAN slicing. The central aim is to elucidate the practical applications of this emerging technology in enabling secure, high-throughput, and low-latency communications. In this paper, we introduce *SliceX*, a RAN slicing microservice hosted by the near-real-

1

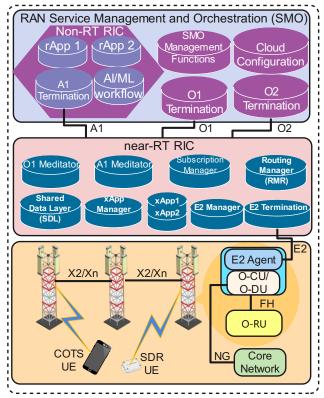


Fig. 1: The O-RAN architecture.

time RAN Intelligent Controller (RIC). *SliceX* is designed to effectively secure the communications between O-RAN and the served users while satisfying their low latency and high throughput requirements. Leveraging the Open Artificial Intelligence Cellular (OAIC) research platform [10], we not only implement *SliceX* but also demonstrate its practicality and effectiveness for RAN optimization, and management.

The rest of this paper is organized as follows: Section II provides background related to O-RAN, RAN slicing, and the integration of security measures, throughput optimization, and latency reduction. In Section III, we introduce the research methods and tools, including the OAIC research platform. Section IV outlines the experimental setup, analyzes the performance and security enhancements achieved with *SliceX*, and discusses the key findings and outcomes from our experiments. Section V presents the concluding remarks.

II. BACKGROUND AND PRIOR WORK

A. O-RAN Architecture

The architectural framework of O-RAN stands as a complex network structure, comprising numerous disaggregated components interconnected through open interfaces. Figure 1 illustrates the O-RAN architecture, highlighting its central network elements, network management components, and key interfaces.

O-RAN introduces the Non-Real-Time (non-RT) RIC and the Near-Real-Time (near-RT) RIC. These controllers play critical roles in managing and optimizing the RAN [11]. The non-RT RIC operates on longer timescales, overseeing tasks such as machine learning model management and resource policy guidance. The non-RT RIC is situated as an integral part of the Service Management and Orchestration (SMO) framework and host microservices called rApps. These rApps

actively engage with both the RAN infrastructure and the near-RT RIC, operating within time frames that span from seconds to minutes. The near-RT RIC operates at a timescale between 10 ms and 1 s, enabling applications such as RAN slicing and load balancing managed by xApps [12].

The near-RT RIC forms connections with the Central and Distributed Units (CUs/DUs) within the disaggregated RAN infrastructure, facilitated through the E2 interface. By defining precise service models, xApps acquire the capacity to collect a wide range of essential performance metrics and other data from the RAN. Furthermore, these xApps have the capacity to send control signals to the RAN infrastructure, enhancing vital cellular network functions such as handover management and RAN slicing, among other capabilities.

B. RAN Slicing

RAN slicing involves partitioning the RAN to create isolated and dedicated network segments where each of these segments, or slices, operates independently and is tailored to meet the unique requirements of specific use cases, applications, or services [13]. RAN slicing was originally introduced to enable the sharing of physical infrastructure among various virtual network operators. RAN slicing can also be employed to optimize the allocation of radio resources to meet the diverse needs of heterogeneous devices and applications. Each RAN slice can have its own set of quality-of-service (QoS) parameters, latency characteristics, bandwidth needs, and security policies, making it suitable for scenarios with varying performance demands. RAN slicing is especially relevant in the context of 5G and next-generation wireless networks, where the diverse range of applications, from massive Internet of Things (IoT) to ultra-reliable lowlatency communications (URLLC), requires a flexible and efficient RAN infrastructure [14]. It enables network operators to efficiently and dynamically allocate a pool of shared resources enabling the coexistence of diverse services on a single physical infrastructure.

While RAN slicing holds significant promise for enhancing network performance, several pressing challenges persist. Notable among these challenges is the efficient sharing of resources, especially in the face of dynamic network conditions, the constraints imposed by static slice management practices on network scalability and adaptability, and the absence of adequate mechanisms for real-time slice reconfiguration in response to changing network dynamics. Moreover, such an operating environment where the physical infrastructure is virtualized to support multiple logical networks presents a potential vulnerability of granting illegitimate users access to network resources and services [15]. The integration of O-RAN emerges as a strategic solution to address these challenges. Within O-RAN deployments, RAN slicing primarily involves leveraging the capabilities of the non-RT and near-RT RICs. These controllers facilitate the fine-tuning of resources allocated to slices. By harnessing the interactions between the non-RT and near-RT RICs, the resource allocation can be dynamically optimized in response to changing network conditions.

Recently, there has been a surge of interest and focus, both within academic research and industry circles, on the practical application of RAN slicing as a means to enhance and optimize the performance of RAN and O-RAN deployments. Johnson et al. [16] implement RAN slicing through a customized xApp. This xApp collects essential metrics from the RAN and, guided by a predefined policy, exercises control over slices using a tailored, slice-aware scheduler. While it demonstrate RAN slicing within the O-RAN context, the given resources to each slice are reconfigured manually controlling scheduling decisions that are static and inflexible to changing network demands. Furthermore, control mechanisms are exposed and are, thus, insecure allowing attackers to gain access to resources easily without any security strategy in place to counteract them.

In support of the implementation of services relying on URLLC, [17] introduces a two-tier RAN slicing framework within the O-RAN architecture, where primary objective of the framework is optimizing the allocation of both communications and computation resources. The study formulates the resource allocation problem at the RAN slicing level as a single-agent Markov decision process. A deep reinforcement learning solution is proposed, enabling dynamic adaptation of resource allocation decisions. It is noted that while the flexibility offered by the proposed architecture presents advantages, it may also introduce occasional deviations from the specified slice characteristics or parameters. This observation underscores the need for further exploration and refinement of the proposed approach to strike an optimal balance between flexibility and adherence to predefined slice requirements.

Shi et al. [18] highlight a significant vulnerability within 5G RAN slicing, specifically the susceptibility to flooding attacks. These attacks can severely impair the performance of legitimate service requests by overwhelming resources with fraudulent users. Although the primary emphasis of the study is on the exploitation of a specific network slicing scheme, the paper establishes that flooding attacks can prove effective against alternative network slicing approaches. Furthermore, the research delves into an examination of various system parameters, encompassing factors such as the weight and rate at which fake requests are generated. This multifaceted investigation sheds light on a vulnerability of network slicing and its potential ramifications, but it does not inform about mitigation strategies that could be applied in real-world use cases.

III. SliceX METHODOLOGY

The O-RAN architecture confronts security challenges arising from the RAN disaggregation and the involvement of multiple vendors [19]. Protecting data during its transmission between users, RAN components, and other network entities is vital to prevent data and information exposure or other types of attacks. The security challenges encompass authentication, access control, confidentiality, and integrity. It is critical to recognize that security breaches, whether deliberate or inadvertent, can have a cascading impact on the overall network performance, specifically in terms of throughput and latency. Malicious attacks or security vulnerabilities can disrupt normal network operations, leading to congestion and increased latency, which in turn can negatively affect the overall reliability, performance, and trustworthiness.

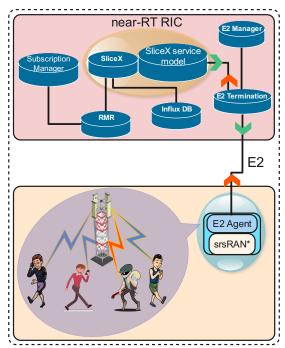


Fig. 2: The SliceX xApp life cycle.

The central objective of our endeavor is the development of an innovative O-RAN secure slicing xApp. *SliceX* is strategically designed with a primary focus on providing high throughput and low latency communications. Its core mission is to go beyond traditional network operations by proactively addressing potential security threats and diligently ensuring the fulfillment of service level agreements within RAN slices. In essence, *SliceX* embodies a cutting-edge approach to network slicing, which not only optimizes resource allocation but also protects users and resources. It aims at ensuring that communications services remain uninterrupted and that service level agreements can be consistently met even in harsh operating environments.

Figure 2 illustrates the life cycle of SliceX. It starts when the xApp is introduced, or onboarded, into the near-RT RIC. The xApp then connects to the RIC Routing Manager (RMR), which manages communications among O-RAN components and enables advanced functions such as connecting with the Subscription Manager. The Subscription Manager provides subscription services to E2 Nodes, and it is used to initiate subscription requests to E2 Nodes that can use the xApp's features. These nodes are identified by their unique IDs. The subscription request is then routed to the E2 Termination, which forwards it to the E2 Agent at an E2 Node. The E2 Agent deciphers the RAN Function ID and service from the message, matches them with the corresponding KPM RAN function, and sends the subscription message to the E2SM-KPM agent for further processing. The E2SM-KPM agent examines the subscription message to determine how often reports should be generated, setting a timer accordingly. The E2-SM KPM agent collects performance metrics from the O-DU, O-CU-CP, or O-CU-UP. These metrics are grouped into containers, each assigned a unique ID. Additionally, each E2 message includes a header with a global RAN ID (PLMN ID) to specify the particular RAN node. These metrics are then combined to create an indication message (KPM Report) using predefined structures in the Subscription

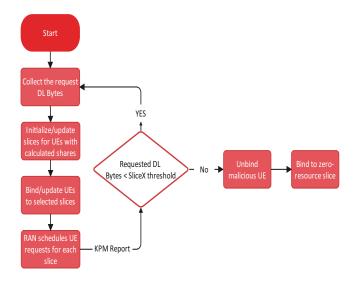


Fig. 3: The SliceX XApp flow chart.

Manager and encoded using ASN.1 encoding. Periodically, these indication messages are sent to *SliceX*, which uses the same Subscription Manager and encoding definitions to decode these messages and extract the metrics [20].

Within this operational context, the Subscription Manager assumes a pivotal and multifaceted role, encompassing both REPORT and CONTROL services. The REPORT service operates on a timed mechanism, aggregating KPM Reports at predetermined intervals and subsequently disseminating them periodically to the xApp. Conversely, the CONTROL service operates asynchronously, initiated by the xApp. This entails the handling of incoming requests transmitted via an interface, their systematic mapping onto E2 convocations, and their ultimate cascading down to the connected nodes.

SliceX serves as a versatile tool, offering comprehensive control and real-time monitoring through an established interface. It supports essential functions such as the creation, update, and deletion of slices. Slices can be linked to the resource block scheduler to adhere to the slice's specified policy when managing the associated UEs. The slice scheduler implements a sophisticated subframe-based proportional slicing strategy that optimizes data transmission on the physical downlink shared channel (PDSCH). This approach ensures efficient resource allocation in each subframe, with a single slice taking precedence for resource allocation, except for the periodic special subframe. In this environment, when a slice is dedicated to a single UE, that UE adheres to the bandwidth configuration defined within its designated slice, ensuring network integrity. The SliceX service model facilitates fundamental operations such as slice creation, modification, or removal by translating them into specific E2 CONTROL messages which implements actions to bind a slice or update its share, among others.

Figure 3 shows the flow of *SliceX* from its initial deployment to the decision-making processes that secures communications. First, the application handles the collection of requests from the operator with respect to downlink (DL) bytes required for a slice. Then the slices are created or updated in accordance with these requests. In an attempt to drive slices to the same overall throughput, it systematically examines distribution disparities among slices and calculates

new allocation shares if necessary, particularly when at least 30% of the available physical resource blocks (PRBs) of the reporting E2 Node are in active use. This balanced slice throughput policy ensures that slices set with low throughput still can operate effectively and are not unfairly starved. These slices are then bound to UEs or existing shares are updated for previously bound slices which will be scheduled resources in accordance to their calculated shares. The metrics collected from the RAN are used to generate a KPM report, which is then assessed during each iteration. This evaluation checks if the requested DL bytes per UE in a slice fall below the threshold set by SliceX for the current number of connected UEs. If this condition is not met, the system takes subsequent action to unbind the potentially malicious UE from its current operating slice by constructing a CONTROL message and transmitting it over the E2 interface to the E2 Node handling the malicious UE. Another CONTROL message is sent that rebinds it to a slice operating with no resource share which will effectively cause the UE to disconnect from the network.

IV. EXPERIMENTAL DEPLOYMENT AND RESULTS

In this section, we present a real-world implementation of the *SliceX* xApp and contrast it with a network not employing *SliceX* using the OAIC platform [20]. We analyze the effectiveness of *SliceX* in securing high throughput and low latency communications by mitigating intrusion attacks. The OAIC platform empowered by a software radio testbed is an open-source O-RAN research platform that enables demonstrating and validating xApps. *SliceX* is designed, developed, and deployed to address key security challenges of O-RAN deployments. It continuously monitors RAN slices for anomalies and deploys security mechanisms in response to emerging attack scenarios.

The experimental setup encompasses two distinct scenarios with a simulated flooding attack where a malicious UE attempts to overwhelm the network with requests for resources. We evaluate the case of a default network configuration without any deployed xApp and compare it against the case which features SliceX actively operating within the near-RT RIC. Figure 4 illustrates the results obtained from default case. In subplot 4(a), we chart the network latency, while subplot 4(b) provides insights into the alterations in throughput when a single UE adopts a malicious stance. In this default configuration, where no proactive monitoring of network slices is in place and no remedial actions are taken in response to the presence of a malicious UE, the consequences are a notable degradation in both network latency and throughput performance. Normal network latency on average is around ten microseconds which increases to over seven seconds when a malicious UE is present and is flooding the network with requests. As depicted in Fig. 4(a), as requests accumulate, the network experiences a degradation in performance, which becomes noticeable at approximately 180 s and reaches its zenith at around 250 s, resulting in a peak sustained latency of 7 s. Without loss of generality, each of the three UEs is configured here to obtain a 10 Mbps throughput share. As the malicious UE achieves a higher resource share, the other UEs experience a drop in resources available to them.

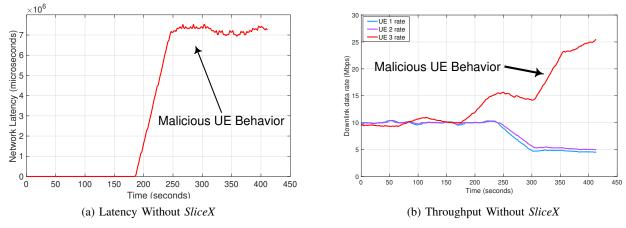


Fig. 4: Network latency and throughput without the SliceX xApp.

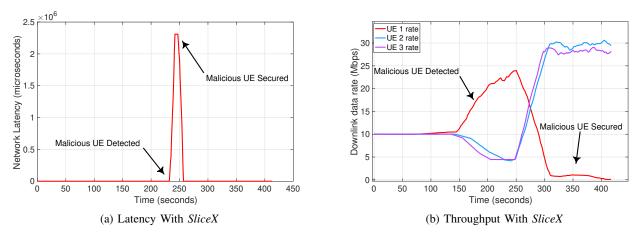


Fig. 5: Network latency and throughput with the SliceX xApp.

The second scenario replicates the previous situation, albeit with an intervention orchestrated by SliceX to isolate the malicious user within a slice devoid of resources. This intervention yields a conspicuous advantage, primarily in terms of network latency and throughput optimization as shown in Fig. 5. Comparing the two cases with respect to latency shows not only a return to normal network latency after the intervention from SliceX, but it also shows that the impact of a malicious user can be confined to, in this case, a temporary network latency of roughly 2.3 s as compared to the default case having a sustained network latency of over 7 seconds. The outcome shows that under the same attack, the SliceX enabled network has a higher than normal network latency for 6.25% of the time it was evaluated as compared to the network without SliceX exhibiting higher than normal network latency for over 50% of the time. Similarly, in terms of throughput, the malicious UE is able to capture more than its prescribed share of resources for over 50% of the run time resulting in legitimate UEs receiving 50% fewer resources resulting in a 5 Mbps DL data rate as opposed to 10 Mbps. *SliceX* is able to reduce the harm to only 25% of the total evaluation period where the legitimate UEs experience a similar throughput loss but only temporarily. While legitimate users without SliceX have no hope to regain consistency with their predefined slice's allotment, SliceX enables users to regain the resources, plus consume the resources that the

malicious UE previously leveraged.

This experiment demonstrated how *SliceX* can play a role in ensuring compliance with service level agreements for each network slice. The proactive secure slicing approach not only improves the overall network performance but also reinforces the network's ability to uphold strict agreements, ensuring a consistently high level of service quality. It highlights the essential role of proactive network management solutions in safeguarding network integrity and elevating the operational efficiency of contemporary wireless networks.

V. CONCLUSIONS

This paper has introduced a proactive RAN slicing xApp which is designed to protect RAN resources and users against malicious UEs while concurrently maximizing throughput and ensuring secure low-latency communications. Leveraging the OAIC research platform we have implemented *SliceX* and demonstrated its capability in achieving the objectives related to RAN optimization and management in a harsh operating environment. The results show the effectiveness of actively monitoring the attached UEs and promptly administering security measures upon detecting malicious behaviors. The findings of this research provide valuable insights to network operators and security practitioners, equipping them with the means to effectively mitigate security concerns through the deployment of *SliceX* in the near-RT RIC. The objective

of this research is to make substantial contributions to the advancement of secure wireless communications while maximizing user throughput, minimizing latency, and providing minimal network overhead.

ACKNOWLEDGEMENT

This work was supported in part by NSF award 2120442 as well as NSF and Office of the Under Secretary of Defense (OUSD) – Research and Engineering, under Grant ITE2326898, as part of the NSF Convergence Accelerator Track G: Securely Operating Through 5G Infrastructure Program.

REFERENCES

- [1] A. S. Abdalla, P. S. Upadhyaya, V. K. Shah, and V. Marojevic, "Toward Next Generation Open Radio Access Networks—What O-RAN Can and Cannot Do!" *IEEE Network*, pp. 1–8, 2022.
- [2] M. Polese, L. Bonati, S. D'oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, 2023.
- [3] M. Dryjański, Ł. Kułacz, and A. Kliks, "Toward modular and flexible open RAN implementations in 6G networks: Traffic steering use case and O-RAN xApps," MDPI Sensors, vol. 21, no. 24, p. 8173, 2021.
- [4] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Open, programmable, and virtualized 5g networks: State-of-the-art and the road ahead," *Computer Networks*, vol. 182, p. 107516, 2020.
- [5] P. Popovski, J. J. Nielsen, C. Stefanovic, E. De Carvalho, E. Strom, K. F. Trillingsgaard, A.-S. Bana, D. M. Kim, R. Kotaba, J. Park et al., "Wireless access for ultra-reliable low-latency communication: Principles and building blocks," *Ieee Network*, vol. 32, no. 2, pp. 16– 23, 2018.
- [6] M. Shin, J. Ma, A. Mishra, and W. A. Arbaugh, "Wireless network security and interworking," *Proceedings of the IEEE*, vol. 94, no. 2, pp. 455–466, 2006.
- [7] A. S. Abdalla, "Physical layer security with unmanned aerial vehicles for advanced wireless networks," Ph.D. dissertation, Bagley College of Engineering, Mississippi State University, 2023.

- [8] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.
- [9] C. De Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 836–886, 2021.
- [10] "Open AI Cellular (OAIC) Website," 2021. [Online]. Available: https://www.openaicellular.org/
- [11] L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and learning in o-ran for data-driven nextg cellular networks," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 21–27, 2021.
- [12] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Colo-RAN: Developing Machine Learning-Based xApps for Open RAN Closed-Loop Control on Programmable Experimental Platforms," *IEEE Transactions on Mobile Computing*, vol. 22, no. 10, pp. 5787–5800, 2023.
- [13] S. E. Elayoubi, S. B. Jemaa, Z. Altman, and A. Galindo-Serrano, "5g ran slicing for verticals: Enablers and challenges," *IEEE Communications Magazine*, vol. 57, no. 1, pp. 28–34, 2019.
- [14] P. L. Vo, M. N. Nguyen, T. A. Le, and N. H. Tran, "Slicing the edge: Resource allocation for ran network slicing," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 970–973, 2018.
- [15] A. S. Abdalla and V. Marojevic, "Multi-agent learning for secure wireless access from uavs with limited energy resources," *IEEE Internet of Things Journal*, pp. 1–15, 2023.
- [16] D. Johnson, D. Maas, and J. K. V. der Merwe, "NexRAN: closed-loop RAN slicing in POWDER - a top-to-bottom open-source open-RAN use case," Tech. Rep., 2021.
- [17] A. Filali, B. Nour, S. Cherkaoui, and A. Kobbane, "Communication and Computation O-RAN Resource Slicing for URLLC Services Using Deep Reinforcement Learning," *IEEE Communications Standards Magazine*, vol. 7, no. 1, pp. 66–73, 2023.
 [18] Y. Shi and Y. E. Sagduyu, "Adversarial Machine Learning for Flooding
- [18] Y. Šhi and Y. E. Sagduyu, "Adversarial Machine Learning for Flooding Attacks on 5G Radio Access Network Slicing," in 2021 IEEE International Conference on Communications Workshops (ICC Workshops), 2021, pp. 1–6.
- [19] A. S. Abdalla and V. Marojevic, "End-to-End O-RAN Security Architecture, Threat Surface, Coverage, and the Case of the Open Fronthaul," arXiv preprint arXiv:2304.05513, 2023.
- [20] P. S. Upadhyaya, A. S. Abdalla, V. Marojevic, J. H. Reed, and V. K. Shah, "Prototyping next-generation O-RAN research testbeds with SDRs," arXiv preprint arXiv:2205.13178, 2022.