IET Research Journals
Research Article
IET Journals The Institution of Engineering and Technology

Leveraging Synergy of SDWN and Multi-Layer Resource Management for 5G Networks

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Mahsa Derakhshani [1] *Saeedeh Parsaeefard* [2] *Tho Le-Ngoc* [3] *Alberto Leon-Garcia* [4]

[1] *Wolfson School of Mechanical, Electrical & Manufacturing Engineering, Loughborough University, UK*
[2] *Communication Technologies Department, ITRC, Tehran, Iran*
[3] *Department of Electrical & Computer Engineering, McGill University, Canada*
[4] *Department of Electrical & Computer Engineering, University of Toronto,Canada*
* *E-mail: m.derakhshani@lboro.ac.uk*

Abstract: Fifth-generation (5G) networks are envisioned to predispose service-oriented and flexible edge-to-core infrastructure to offer diverse applications. Convergence of software-defined networking (SDN), software-defined radio (SDR), and virtualization on the concept of software-defined wireless networking (SDWN) is a promising approach to support such dynamic networks. The principal technique behind the 5G-SDWN framework is the separation of control and data planes, from deep core entities to edge wireless access points. This separation allows the abstraction of resources as transmission parameters of users. In such user-centric and service-oriented environment, resource management plays a critical role to achieve efficiency and reliability. In this paper, we introduce a converged multi-layer resource management (CML-RM) framework for SDWN-enabled 5G networks, that involves a functional model and an optimization framework. In such framework, the key questions are if 5G-SDWN can be leveraged to enable CML-RM over the portfolio of resources, and reciprocally, if CML-RM can effectively provide performance enhancement and reliability for 5G-SDWN. In this paper, we tackle these questions by proposing a flexible protocol structure for 5G-SDWN, which can handle all the required functionalities in a more cross-layer manner. Based on this, we demonstrate how the proposed general framework of CML-RM can control the end-user quality of experience. Moreover, for two scenarios of 5G-SDWN, we investigate the effects of joint user-association and resource allocation via CML-RM to improve performance in virtualized networks.

1 Introduction

1.1 Motivation

Cellular wireless networks have been waiting to evolve by 2020 in the context of 5G wireless networks [1, 2]. 5G needs to offer a more service-oriented, flexible, and spectrum/energy-efficient structure to improve quality-of-service (QoS) for end-users by enabling heterogeneity over the utilization of different technologies. Besides, aiming to provide seamless connection, all the legacy wireless networks will be integrated under the 5G structure from edge to the core [3–5]. Among all the new contexts, three emerging trends in computer and communication networks—SDN, SDR, and virtualization—are expected to converge under the umbrella of 5G-SDWN to cater increasing demands for diverse services. Offered by these three pillars, this new paradigm, called 5G-SDWN, provides numerous advantages, ranging from higher spectrum/energy efficiency and lower end-to-end transmission delay, to lower costs and time required for launching new applications and services [2, 6, 7].

Such advantages of immigrating to 5G-SDWN stem from the fact that wireless procedures and functional units of infrastructure entities can be moved into software with the aid of SDR and SDN, for both edge access points (APs) and core nodes in 5G. Furthermore, the separation of control and date planes, enabled by SDN and SDR, provides the ground to offer network and wireless virtualization [6, 8–11]. Allowing abstraction of resources, virtualization is a technique to share network infrastructure among different service providers (SPs) and to bound resources for a specific set of users over the concept of slicing [12–15].

5G-SDWN stands at the shoulders of these three networking layers, where it inherits all their flexibilities by transition from hardware to software-based implementation. SDR acts as a physical layer of 5G-SDWN where APs are reprogrammable and adjustable. SDN takes care of management and controlling messages among nodes and cellular network functionalities. More specifically, it is a translator of all protocols, standards and vendors together in such a way that *communications*, *transactions* and *transmissions* among different entities are technology and vendor-agnostic [16, 17]. Over these two layers, virtualization is surfing as an application for SPs and slices aiming to improve infrastructure utilization of 5G-SDWN.

This simple high-level structure of 5G-SDWN opens the door of cross-layer, dynamic and efficient implementation of functionalities and procedures related to each network entity from the core to the edge, while introducing the software intelligence over the entire 5G-SDWN. From this programmable structure of 5G-SDWN, the centralized and comprehensive view of all network entities is feasible, where all network infrastructure can be re-arranged and adjusted for each requested service to complete the transmission path. In other words, the flexible structure and cross-domain integrity of 5G-SDWN provide the capability to abstract resources from the infrastructure level and to deploy comprehensive resource management. In this paper, we introduce a converged multi-layer resource management (CML-RM) framework for SDWN-enabled 5G networks, in which all transmission parameters and connections can actively be manipulated based on requested services, user conditions, and 5G-SDWN available resources. Such CML-RM can handle, harmonize, and distribute user traffics among APs and core entities in such an efficient manner to cope with the underutilization of resources, QoS provisioning, and user-association from the core to the edge [18–20].

To fully realize features and potentials of CML-RM in 5G-SDWN, we first present and elaborate a general architecture and its cross-layer stack protocol for 5G-SDWN, which can handle all the required functionalities in a more cross-layer manner. Then, we propose a modular functional model for CML-RM, which explains what the required functional components are and how they interact with each other. We also provide a general optimization framework for CML-RM, which describes possible utility functions and defines different constraints. We believe that the proposed CML-RM enabled by 5G-SDWN can considerably increase the network performance. To manifest the importance of CML-RM, we present two case studies for association control and resource allocation in homogeneous virtualized macro-cells and small-cells (i.e., 802.11 wireless local area networks (WLANs)) networks. It is shown that association control leveraging 5G-SDWN principles can be of benefit to improve overall throughput, isolation among SPs, and coverage in both types of wireless networks.

1.2 Related Works

To acquire the full benefits of softwarization over 5G, two categories of works have been developed in this context. The first class of works has been focused on designing the system level of resource management framework and concepts in SDWN-enabled 5G networks, while in the second category, the effects of this type of softwarization on the network performance have been investigated by applying efficient resource allocation techniques. Besides new surveys and tutorials in this context, in [21], the evolutionary and revolutionary views of resource management toward 5G are investigated.

As part of the first class, in [22], the effects of network slicing on resource management are investigated. In [23], the effect of cloud radio access networks structure (C-RAN) on 5G resource management is studied considering different functional splitting in this context. Moreover, [24] proposes an SDWN-enabled spectrum management architecture.

In the second class of works, mainly, resource management algorithms are developed based on the softwarization techniques and the performance has been studied. For instance, the user association based on the new structure of 5G is investigated in [25]. In [26], a resource management structure for multi-tenant heterogeneous networks towards 5G is proposed and they show how this approach can achieve objectives of both network's and users' perspectives, while it is a scalable solution. Along this direction of works, there are plenty of works investigating the performance gain of proposed resource management structure for 5G SDWN, e.g., [27].

Following the above two directions of research, in this work, we propose a converged multi-layer resource management framework for 5G SDWN and investigate the effects of the proposed CML-RM framework on improving the 5G network performance via two examples.

1.3 Organization of Paper

The remainder of this paper is organized as follows. In Section 2, a brief review of main technologies, SDN, SDR, and virtualization is presented as well as the overall architecture of 5G-SDWN. In Section 3, a functional model and an optimization framework for converged multi-layer resource management in 5G-SDWN are proposed. Section 4 presents two illustrative examples to show how this new paradigm of resource management in 5G can improve the network performance, followed by the conclusion and highlight on the future works in Section 5.

2 5G-SDWN Generic Architecture

We believe that flexibility and software-based features of 5G-SDWN lead to more intertwined and cross-layer stack protocol design. Clearly, by providing *software intelligence* to the physical entities of wireless networks, tasks interrelated to different Open Systems Interconnection (OSI) layers can be handled by only one layer over the stack protocol of 5G-SDWN. This will lead to lower delay for requested services by increasing the processing speed and decreasing the latency over 5G-SDWN-based network. In this section, first, we briefly review the concepts of SDN, SDR, and virtualization. Next, as a combination of these three concepts, we present and elaborate a general architecture for 5G-SDWN and its cross-layer stack protocol.

2.1 Brief Overview of SDN, SDR and Virtualization

Traditional data networks follow a tightly coupled structure of data and control planes, embedded in each network element. While supporting network's resilience, this decentralized structure causes a bottleneck for extending and updating the data networks due to its complex and relatively static architecture. The software-defined networking (SDN) paradigm cuts this coupling design of control and data planes by (i) removing control functionality from network nodes and turning them into simple data/packet forwarding nodes, (ii) changing data decision/routing forwarding from destination-based to flow-based, and (iii) moving control logic to an external entity, called the SDN controller [6].

These SDN controllers have an overall view of the network nodes. Thus, they are able to control all the nodes and their interfaces in using high-level languages. The SDN controller translates these programs into actions for each network element and hides different interface commands for nodes. In this structure, SDN controllers have two main interfaces, called northbound and southbound. The southbound takes care of the interactions between the controller, network elements and programmable interfaces at the edge elements. The northbound interface is the communication bridge between the controller and control applications. SDN and its decoupling feature can bring more flexibility and efficiency into networks. Furthermore, indirectly, SDN can reduce capital and operational expenses (CAPEX and OPEX) of communication networks, promoting wireless operators to provide new services with lower prices.

Delivering similar benefits as in SDN, the decoupling of the control and transmission layers can be applied in edge wireless transceivers with the help of *software-defined radio* (SDR). In particular, SDR is the edge transceiver with two basic units, i.e., radio and processing units. First, the radio unit is responsible for transmitting and receiving radio frequency (RF) signals. Ideally, this unit should work at different frequencies and standards. Second, the processing unit is responsible for all radio operating functions including modulation/demodulation, coding/decoding, encryption/decryption, and medium access control (MAC) procedures. All these functionalities are implemented over programmable processing technologies such as field-programmable gate array (FPGA) and a generic central processing unit (CPU) [3]. Therefore, SDR allows traditionally hardware-integrated wireless functionalities to be controllable through software-based controllers.

From the 5G-SDWN prescriptive, the SDR transceiver is a flexible and smart entity, which can enable self-organizing networking (SON) solutions and provide the portfolio of wireless resources for optimizing the network performance considering the wireless channel conditions, interference level, and QoS requirements of each user. Most importantly, it can adjust the multi-user access techniques and MAC protocols based on all mentioned parameters. Flexible reconfiguration of MAC protocol and transmission mode selection is a key enabler to reach higher spectrum and power efficiency based on the comprehensive resource management for 5G-SDWN.

Virtualization allows the flexible reuse and sharing of the existing infrastructure among different SPs (also called tenants), which is another enabler for reaching higher spectrum and energy efficiency. This concept was initiated in computing and backbone networking domains, and now by cooperation of SDN and SDR, it has been extended to the wireless edge of networks. Therefore, there exists a broad range of sharing or virtualization from high-level network management, service allocation or application sharing to low-level hardware or physical resource sharing [12, 13].

In this paper, virtualization is defined as the abstraction and bundling all kinds of resources and equipment by tenants, which is considered as a main application, while networking shares context over the entire infrastructure of 5G-SDWN. In this context, the virtual instance of a set of bundled resources for one tenant is called a slice. Depending on how and in which layer resources are sliced for one tenant, the depth of slice can be determined over the stack protocol of 5G-SDWN. The important implementation issue of slices is to provide isolation among slices. It means that any change in one slice because of new users' arrival, mobility, and channel fluctuations, should not affect services offered to other slices.

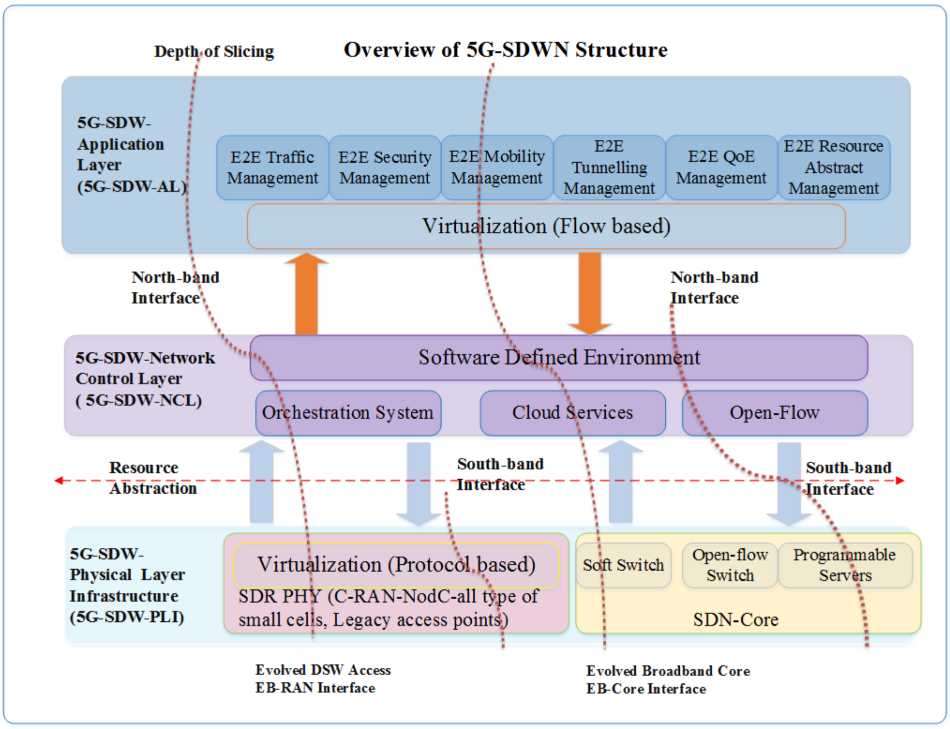


Fig. 1: Illustration of 5G-SDWN generic architecture

2.2 *5G-SDWN Architecture*

All of the aforementioned technical advancements can be gathered under the umbrella of the 5G-SDWN, which could be a major revolution of the wireless generations. A basic architecture and outline of generic 5G-SDWN structure have been put forward as shown in Fig. 1 and Fig. 2, respectively. This architecture has three main layers: 5G-SDW application layer (5G-SDW-AL), 5G-SDW network control layer (5G-SDW-NCL), and 5G-SDW physical layer infrastructure (5G-SDW-PLI).

2.2.1 5G-SDW-AL: This layer contains all applications related to the wireless networks such as mobility management, connection control, and security for both access and core. In the following, the detailed functionalities, procedures, and processes related to each and every module in 5G-SDW-AL are listed.

- *Security management*: User authentication, encryption, and key management
- *Mobility management*: Handover over the network from access into gateways, interactions with other networks, roaming, tracking, and location update of users in different states
- *Traffic management*: Connection management, load balancing, traffic sharing, and role control of all entities in both access and core
- *Data tunneling management*: Routing and tunneling of data, session, or flow
- *Quality-of-Experience (QoE) management*: QoS assignment, flow assignment, and service admission policy
- *Resource abstract management*: MAC allocation, software selection, standard selection, and admission control.

All applications of 5G-SDW-AL are related to the end-to-end transmission , considering both core and access. They are all defined based on the requirements of each user considering its requested services over specific virtualization scenario. The flow-based virtualization can be initiated from this layer by 5G-SDWN as demonstrated in Fig. 2.

2.2.2 5G-SDW-NCL: This layer contains all components and functionalities similar to the SDN controller as mentioned in Section 2.1. In this layer, an overall view of the network is achieved. Through this layer, every action is first transformed into a vendor-independent programming language and then transferred to 5G-SDW-AL (via northband interface) or to 5G-SDW-PLI (via southband interface). The abstraction of different resources in 5G-SDWN is provided here from 5G-SDW-PLI to 5G-SDW-AL. This intelligence and decoupling come from SDN (OpenFlow) and orchestration system (i.e., a broker between the applications and the network elements), and recent cloud-based protocols. All the functionalities related to the control and management of 5G-SDWN belong to this layer, including SDN server, SDN controller, routers, and switches as demonstrated in Fig. 2.

2.2.3 5G-SDW-PLI: This layer encompasses all the physical entities required to handle all functionalities of 5G-SDWN including two major parts:

- *5G-SDW-P Access*, which contains access points from different technologies and generations [3], e.g., legacy small cells (femto, pico, Home eNodeB (HeNB)), 5G small cells (including millimeter wave (mm-Wave) small cells), macro cells of legacy networks (e.g., base transceiver station (BTS) of 2G, NodeB (NB) of 3G, evolved NodeB (eNB) of LTE), macro cells for 5G including cloud radio access network (C-RAN), NodeC, massive multiple input multiple output (massive MIMO) access points, relays, and back-haul and front-haul links [28]. All these elements are demonstrated in the related layer in Fig. 2.
- *5G-SDW-P Core*, which contains all front-haul, soft switches, open-flow switches, gateways, and any programmable servers for deployment applications including home-location registration, cloud servers, and data centers, as depicted in Fig. 2.

To reach the seamless connection from the end-users point of view, 5G-SDWN will deal with different wireless generations, technologies, and standards specifically in 5G-SDW-P Access due to the broad range of radio access technologies. Since 5G-SDW-P Access plays a crucial role for the end-users and management of the other functionalities of 5G-SDWNs, we propose the following classification for the 5G-SDWN.

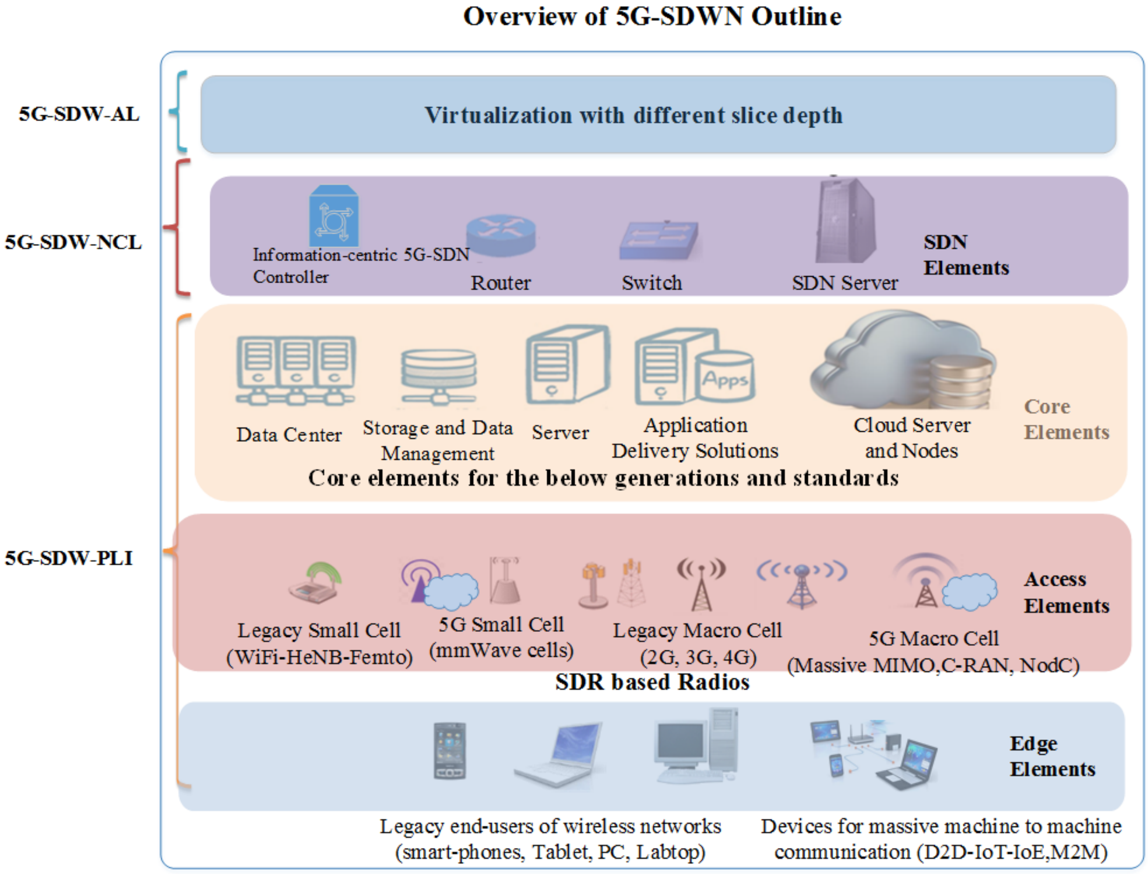


Fig. 2: 5G-SDWN generic outline based on proposed architecture

- *Coordinated multi-AP 5G-SDWN (CM-5G-SDWN)*: In CM-5G-SDWN, SPs within one specific region are served by the transceivers belonged to only one generation of wireless technologies, e.g., (C-RAN, massive-MIMO-based AP, 2G, 3G, LTE, or WiFi). In this setup, both coverage and capacity will be provided by the same technology for all users of networks.
- *Heterogeneous multi-tier 5G-SDWN (Het-5G-SDWN)*: In Het-5G-SDWN, different generations and technologies support the coverage and capacity of wireless access. For example, in a wireless access network, one of the legacy generations (3G or 4G) and/or C-RAN, NodeC or massive-MIMO-based APs can be deployed to provide the best coverage. This layer can be considered as a *coverage/overlay layer*. Simultaneously, the high traffic hot-spots can be served by 5G small cells with LTE, WiFi and/or up-coming 5G dense deployment of small cells, as a *capacity/underlay layer*.

Obviously, implementation, planning, and optimization of CM-5G-SDWN are easier than those for Het-5G-SDWN. Nevertheless, the latter delivers enormous capacity and coverage improvement.

In Fig. 2, we propose two groups of end-users for 5G-SDWN. First group includes legacy users of wireless networks, including all existing data-hungry users such as smart phones, tablets, and laptops. This type of users, which are capable to run a wide range of data applications, should connect directly to the APs of 5G-SDW-PLI. Second group represents devices for machine-type communications, including all device-to-device (D2D), machine-to-machine (M2M), Internet of things (IoT)-capable devices, which can connect to each other in addition to the 5G-SDW-PLI APs. Via CML-RM over 5G-SDWN, transmission modes and types of the connection among the latter devices can be controlled aiming to increase the spectral and energy efficiency.

According to the software-based and programmable structure of 5G-SDWN, a portfolio of network resources is available, which is abstracted from 5G-SDW-AL to the higher layers. Therefore, 5G-SDWN can leverage a CML-RM to optimize the network performance. Such integrated management of converged resources can provide energy-aware and efficient resource allocation. For each user, such allocation will be applied over the management parameters of *transmission and control planes* based on its service requirements, depth of its corresponding slice, and virtualization type. We call this procedure as "*CML-RM over 5G-SDWN*".

3 Toward a converged multi-layer resource management over 5G-SDWN

The flexible structure and cross-domain integrity of 5G-SDWN provide the capability to abstract resources from the infrastructure level and to deploy the converged multi-layer resource management over the network. In the 5G-SDWN structure, all network resources can be divided into three categories: 1) wireless resources: spectrum, transmit power, antenna, beam, time, and code; 2) computing resources including all the storage, computing units of clouds, and base-band units of C-RAN, 3) infrastructure resources, all APs, switches, links, front-haul, and back-hauls links. From the 5G-SDWN point of view, all these resources are abstracted to diffident grids and tables, which can be assigned by the CML-RM and divided to the different layers of networks.

3.1 Hierarchical Functional Model for CML-RM

Here, we present a hierarchical functional model for CML-RM in the 5G SDWN, which is illustrated in Fig. 3. This model consists of different elements including Software-defined Virtual Resource Management (SD-VRM), Software-defined Common Resource Management (SD-CRM), and Software-defined Local Resource Management (SD-LRM) entities. With this layering structure, CML-RM

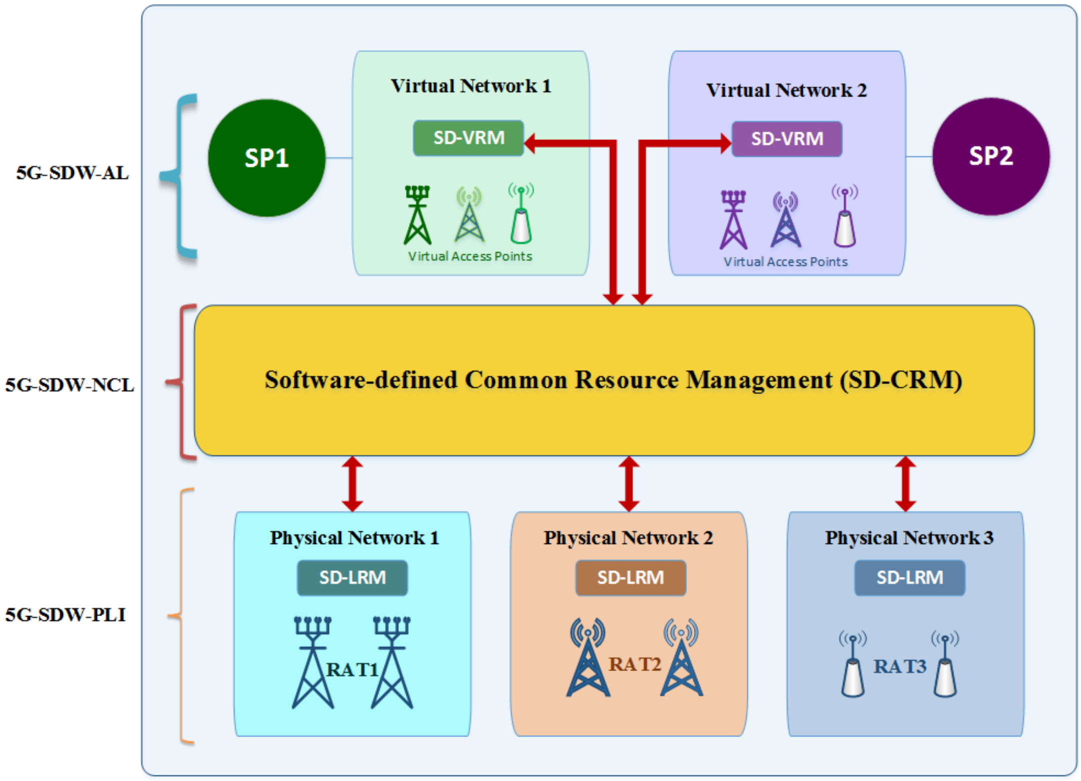


Fig. 3: A functional model for CML-RM in 5G-SDWN

can consider each device as a packet forwarding node, where all of its transmission parameters can be adjusted based on its grid of resources. Consequently, the final performance of node can be translated based on the requirement of each layer as well. Via this approach, CML-RM is attainable over 5G-SDWN.

In the following, we elaborate on the functionalities of each and every component of our proposed functional model.

3.1.1 The SD-VRM entity: This entity is placed on top of this hierarchy and is run for each separate virtual network, which belongs to an individual SP/slice. SD-VRM needs to translate QoS requirements and service level agreements (SLAs) for the lower levels. Depending on the level of isolation a SP is needed, scheduling end-users within the virtual network can be of its responsibility. But, such scheduling at SD-VRM is not dealing with physical resources and only virtual resources are created and allocated to provide the capacity required by the SP. The physical resource scheduling would be delegated to the lower levels.

3.1.2 The SD-CRM entity: This entity is responsible to manage a pool of different resources, including wireless and computing resources as well as infrastructure resources of different radio access networks (RANs), in a coordinated manner. In this layer, the physical resources of different RATs and core network are abstracted as resource blocks and SD-CRM is in charge of scheduling and resource block allocation for SPs and/or end-users. With such centralized management, it can be ensured that the resource block scheduling takes into account the QoS requirements of all slices/SPs as well as the resource availability in all SD-LRM entities. The concept of Common *Radio* Resource Management (CRRM) has been already introduced by 3GPP (Third Generation Partnership Project) for heterogeneous multi-tier wireless networks [29], although it is not applicable for 5G-SDWN due to the requirement for slice isolation and resource provisioning in virtualized networks.

3.1.3 The SD-LRM entity: This entity performs the management of the resources of a specific RAN. For each physical network, SD-LRM is responsible to map the scheduling of resource blocks onto physical ones. The decision on resource block scheduling is made by SD-CRM and reported back to SD-LRM. Moreover, SD-LRM is in charge of informing SD-CRM about the available capacity and measurements (e.g., channel state information) taken at the access points.

In summary, each SD-VRM acts as a local resource manager and scheduler for a virtual network of a SP. First, SD-VRM needs to communicate the QoS requirements of its corresponding SP in terms of the total number of resources (resource-based provisioning) and/or the total rate requirement (rate-based provisioning) based on its service level agreement. Then, in a centralized manner, SD-CRM will allocate the requested resources to each virtual network considering network (including access, core, fronthaul and backhaul) constraints and other slice SP requirements. Subsequently, each SD-VRM allocates resources assigned by SD-CRM to different users to ensure that each slice can employ custom user scheduling policies. It should be noted that depending on the level of isolation required, SD-CRM might take the responsibility for user-level resource management as well.

In a heterogeneous environment where several RATs coexist and multiple SPs share the infrastructure, the modularity in this design helps to handle network-wide resource management and isolation among different SPs. In comparison to the models proposed in [30, 31], this model introduces a centralized coordinated management in the SD-CRM along with several resource manager/schedulers (i.e., SD-VRM), each exclusively for a virtual network, ensuring that SP could manage scheduling in its own virtual network.

3.2 General Optimization Framework for CML-RM

Here, we introduce a general optimization framework for CML-RM as

$$\max f_0(\mathbf{x}), \tag{1a}$$

subject to:

$$\mathrm{C}_i^m: \; f_i^m(\mathbf{x}) \leq 0, \quad m \in \mathcal{M}, \tag{1b}$$

$$\mathrm{C}_j^g: \; f_j^g(\mathbf{x}) \leq 0, \quad g \in \mathcal{G}, \tag{1c}$$

where $\mathbf{x}$ is a vector representing the optimization variables including all resources (i.e., wireless, computing, and infrastructure resources) to be allocated to reach the goal of maximizing $f_0(\mathbf{x})$. In this optimization problem, $f_0(\mathbf{x}) = (f_0^1(\mathbf{x}), \ldots, f_0^{N_0}(\mathbf{x}))$ represents the objective function which in general might involve multiple objectives related to different network layers. Each $f_0^n(\mathbf{x})$ can describe different network utilities such as maximizing total throughput, maximizing number of admitted flows, maximizing the revenue, minimizing operational costs, minimizing buffer usage, minimizing delay, and minimizing power consumption.

The constraints of this optimization problem reflect different limitations in network as well as isolation requirements of different slices/SPs sharing the network. The CML-RM ensures isolation of network resources across SPs in relation to the key performance indicators, and QoS and QoE parameters of each slice. More specifically, we introduce two categories of constraints. In the first group (1b), $f_i^m(\mathbf{x})$ represents network constraints such as communication link constraints, computing and storage limitations, infrastructure restrictions. In the proposed CML-RM framework, isolation is interpreted as slice-level constraints and reflected in (1c), where $f_j^g(\mathbf{x})$ describes the QoS and isolation requirements of each slice $g \in \mathcal{G}$. Such constraints provide provisioning for different slices, either as resource-based provisioning or QoS-based provisioning [15]. In the resource-based provisioning, the slice requirement of each SP is defined as a preserved share of network and wireless resources, e.g., the number of sub-carriers using orthogonal frequency division multiple access (OFDMA). Instead of preserving a fixed number of resources, QoS-based provisioning ensures a certain quality level of service for each slice in terms of total rate requirement, and/or delay requirements.

This CML-RM framework can handle, harmonize, and distribute the user traffic between APs and core entities. However, such CML-RM is not trivial and straightforward over SDWN. There are technical challenges as

- Diversity of control parameters,
- Complexity of framework,
- Feasibility of defined optimization problem,
- Scalability and performance trade-off.

More specifically, since the parameters for CML-RM are diverse, the main challenge is how to translate different layer parameters and regularize them in similar dimensions. After this step, compatible mathematical formulations need to be defined in a form of optimization problems. Such problems are generally nonconvex, combinatorial, and thus computationally complex. Due to various QoS requirements over different slices, with high probability these problems generally suffer from infeasibility issues. The other issue is how much integration over CML-RM is right and enough. Unfortunately, since number of network parameters is large, considering all of them in the problem is not reasonable and comes at the cost of high computational complexity and low scalability.

One example of such CML-RM is network-wide association control in wireless networks. In a network with densely deployed APs, before a user can access the network, it needs to make a decision about which AP to associate with. In most current vendor implementations, a connection/handover is initiated by users. In particular, users choose the AP with the highest received signal-to-noise ratio (SNR) to connect with. However, the AP with the maximum SNR may not always have enough capacity or resources to occupy an additional user. Furthermore, since the user density is often uneven in the network, the Max-SNR approach can lead to an unbalanced distribution of users among APs, causing unfairness.

Thus, delegating management rights to the network operator to decide how to associate users with APs can be useful to guarantee connectivity, manage QoS, and balance the traffic load. By remote assistance of the controller, SDWN could enable network-originated association control. In the next section, we present two case studies for association control and resource allocation enabled by SDWN in homogeneous virtualized 802.11 and cellular networks. These studies manifest the challenges and importance of network-wide resource management and association control—that can be achieved in an SDWN architecture—for service customization and QoS provisioning.

4 User Association over SDWN

Multi-tier multi-technology "de-cell-ization" is an inherent structure for 5G radio access, which is highly promised by the cloud-based RAN [2, 32]. Consequently, the user association to the appropriate access points belonging to different tiers and technologies is of high importance. Accordingly, the user association over CML-RM is not only essential but also feasible for 5G due to the end-to-end software defined based structure. In this section, we will present our problem formulations for user association for the scenarios of CM-5G-SDWN and present how we can overcome the computational complexity. The extension of user association over Het-5G-SDWN will remain for our future works. Notably, the user association problem over 5G and traditional wireless networks has been drawn a lot of attention recently. For instance, the user association problems for the multi-cell wireless networks equipped with massive-MIMO are studied in [33–36]. In the following works, we will present how the user association factor can be defined in the wireless networks, which can combine and interrelate different implementation limitations in this context with considering new sets of constraints.

4.1 Association and Airtime Control in Virtualized 802.11 Networks

In virtualized 802.11 WLANs, transmissions of different virtual WLANs (V-WLANs) are closely coupled, although administrative virtualization (i.e., one physical AP advertises multiple service set identifiers (SSIDs)) can already differentiate groups of flows. With a contention-based MAC based on carrier sense multiple access with collision avoidance (CSMA /CA), unavoidable collisions act to couple the transmissions of different V-WLANs. Moreover, since the network capacity is shared yet constrained, the increase of traffic in one V-WLAN may reduce the available network capacity to another. Thus, an efficient resource allocation among V-WLANs is essential to manage the MAC-layer couplings. To overcome such MAC-layer couplings and balance the load, we propose a user-level management approach in virtualized 802.11 networks.

More specifically, we consider an IEEE 802.11-based WLAN that consists of a number of APs. APs operate on non-overlapping frequency channels. Let $\mathcal{A}$ be the set of APs and $N_a = |\mathcal{A}|$ be the total number of APs. Each AP has a limited coverage area and all users are randomly distributed in the field. The network carries traffic belonging to a number of different SPs (also referred to as V-WLANs). Let $\mathcal{G} = \{1, \ldots, G\}$ be the set of SPs using the network. Furthermore, let $\mathcal{N}_g$ be the set of users of SP $g \in \mathcal{G}$. Furthermore, let $\mathcal{N}$ be the set of all users and $N = \sum_{g \in \mathcal{G}} N_g$ be the total number of users in the network. The network is administratively virtualized, i.e., each AP will broadcast multiple different SSIDs, one for each SP. Fig. 4a illustrates an example of the network architecture with four physical APs and two SPs.

In a WLAN with APs densely deployed, users need to determine which APs to connect with. We aim to generalize the association control problem by adjusting the transmission probability of each user at any AP, rather than selecting one AP to associate with. Thus, we define $\tau_{n_g}^a$ $(0 \leq \tau_{n_g}^a \leq 1)$ as the probability that user n_g attempts to transmit at AP a in a general time-slot.

Let $T_{n_g}^a$ be the throughput of user n_g at AP a and T_{air,n_g}^a be the total access airtime for user n_g at AP a. In [37], based on the CSMA/CA operation, it is shown that $T_{n_g}^a$ and T_{air,n_g}^a can respectively be calculated as

$$T_{n_g}^a = \frac{x_{n_g}^a R_{n_g}^a t}{\prod_{\forall g \in \mathcal{G}} \prod_{\forall n'_g \in \mathcal{N}_g} (1 + x_{n'_g}^a) - t'}, \tag{2}$$

and

$$T_{\mathrm{air},n_g}^a = \frac{x_{n_g}^a \prod_{\forall g \in \mathcal{G}} \prod_{n'_g \in \mathcal{N}_g, n'_g \neq n} (1 + x_{n'_g}^a)}{\prod_{\forall g \in \mathcal{G}} \prod_{\forall n'_g \in \mathcal{N}_g} (1 + x_{n'_g}^a) - t'}, \tag{3}$$

where $x^a_{n_g} = \frac{\tau^a_{n_g}}{1-\tau^a_{n_g}}$, $t = \frac{T_{\text{TXOP}}}{T}$, $t' = \frac{T-\delta}{T}$, δ is the duration of an idle time-slot, T_{TXOP} is the duration of a data frame, T is the duration of a successful transmission, and $R^a_{n_g}$ represents the transmission data rate of the link between user n_g and AP a.

Apparently, based on (2) and (3), the achievable throughput and access airtime of each user is not only a function of its transmission probability, but also significantly depends on the transmission probabilities of other users associated with the AP. Such coupling between transmission strategies of different users necessitates a careful and comprehensive optimization.

Taking into account user transmission rates and SP airtime reservations, in this approach, we aim to jointly optimize the transmission probability of all users at different APs to maximize the overall network throughput of all users at all APs. Furthermore, for each SP (e.g., SP g), we set a constraint to keep the total airtime of all users belonging to SP g larger than a minimum requirement. This set of constraints enables controlling SPs' share of access airtime regardless of their number of users. Therefore, the user-association optimization problem can be formulated as

$$\max_{\boldsymbol{X}} \sum_{\forall g \in \mathcal{G}} \sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} T^a_{n_g} \tag{4a}$$

$$\text{subject to :}$$

$$\text{C1}: \sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} T^a_{\text{air},n_g} \geq \eta_g, \; \forall g \in \mathcal{G} \tag{4b}$$

where $\boldsymbol{X} = [x^a_{n_g}]$ $(x^a_{n_g} \geq 0)$ and η_g denotes the target share of the airtime for SP g.

Substituting $T^a_{n_g}$ and T^a_{air,n_g} based on (2) and (3), more specifically, (4) can be written as

$$\max_{\boldsymbol{X}} \sum_{\forall g \in \mathcal{G}} \sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} \frac{x^a_{n_g} R^a_{n_g} t}{\prod_{\forall g \in \mathcal{G}} \prod_{\forall n'_g \in \mathcal{N}_g} (1 + x^a_{n'_g}) - t'}, \tag{5a}$$

$$\text{subject to :}$$

$$\text{C1}: \sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} \frac{x^a_{n_g} \prod_{n'_g \in \mathcal{N}_g, n'_g \neq n_g} (1 + x^a_{n'_g})}{\prod_{\forall g \in \mathcal{G}} \prod_{\forall n'_g \in \mathcal{N}_g} (1 + x^a_{n'_g}) - t'} \geq \eta_g, \forall g \in \mathcal{G} \tag{5b}$$

Each constraint in (5b) describes the isolation requirement of each SP in terms of its minimum airtime reservation. Controlling airtime usage of the users provides the opportunity to optimize the SP performance (e.g., improving throughput by exploiting multi-user diversity) as well as another degree of freedom to guarantee fairness among the SPs.

In order to solve the optimization problem (5), in [37], we have proposed an iterative algorithm based on complementary geometric programming (CGP). Details of the developed iterative algorithm to obtain jointly optimal transmission probabilities can be found in our study [37]. This SDWN-enabled algorithm searches for the jointly optimal transmission probabilities of all users to maximize the total throughput, while guaranteeing the minimum requirement of each SP.

Here, the performance of the proposed SDWN-enabled and conventional Max-SNR association approaches are compared in two examples under different user density and SP load. The SDWN-enabled algorithm referred to the algorithm that solves the optimization problem in (5). Max-SNR association is the traditional and heuristic technique, implemented by most vendors, in which a user will be connected to the access point from which receives the highest received signal-to-noise ratio (SNR).

For the numerical results in Fig. 4b and Fig. 4c, the simulation assumes a network with 4 APs placed at the centers of four $5 \times 5\ m^2$ grids to provide seamless coverage. To eliminate interference between APs, four non-overlapping 20 MHz channels are assigned to four APs. The users are distributed in the entire area according to the two-dimensional Poisson point process (PPP). Let define ρ_1 (SP 1 load) as the ratio of number of users served by SP 1 to the total number of users in the network.

The wireless channel model includes path loss and small-scale fading. Generally, the channel power gain can be expressed as $h = \chi d^{-\alpha}$, where d is the distance between a user and an AP, $\alpha \geq 2$ is the path loss exponent and χ represents the small-scale Rayleigh-fading component. In the numerical results, we set $\alpha = 3$ and h' is an exponential random variable with mean 1. The received SNR at user n is equal to $\frac{P h^a_{n_g}}{\sigma^2}$ where P is the transmission power, $h^a_{n_g}$ is the channel power gain from user n_g to AP a, and σ^2 is the noise power. In the numerical results, P/σ^2 is assumed to be 10dB.

The MAC layer parameters used in our simulations are set as follows: idle time slot $\delta = 9\mu$s, fixed transmission duration $T_{\text{TXOP}} = 1$ ms, and the duration of a successful transmission $T = 1.1$ ms. Moreover, the target airtime share for each SP g (i.e., η_g) is set equal to the number of APs divided by the number of SPs. In other words, we assume that the SPs have the same minimum airtime reservation and share the total airtime in a fair manner.

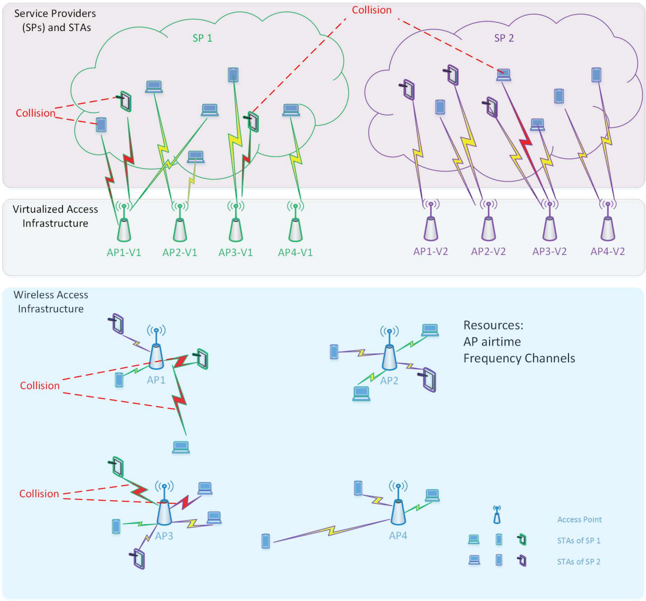
The first example in Fig. 4b shows the total throughput (i.e., $\sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} T^a_{n_g}$ based on (5a)) achieved by the two association algorithms versus λ_{mean} (which represents the average number of users per AP) for a homogeneous user distribution. For any fixed ρ_1, it is shown that SDWN-enabled association significantly improves the total throughput as compared with the Max-SNR. This is mainly because the SDWN-enabled algorithm jointly chooses the transmission probabilities of all users to maximize the total throughput of all users, while max-SNR algorithm selects the AP with the largest $R^a_{n_g}$ for each user without considering the impact on other users. According to (2), by increasing the transmission probability of user n_g at AP a, there is a tradeoff between increasing the throughput of user n_g at AP a and decreasing the throughput of other users associated to this AP, SDWN-enabled approach can improve the total throughput, taking care of the tradeoff between different users depending on the network conditions such as transmission data rate of the links and user distributions.

Moreover, for a fixed ρ_1, the total throughput by both algorithms increases with the user density. But, the throughput increase rate is decreasing with λ_{mean}. This is because the wireless channel is underutilized when the user density is low. Thus, the increase in the user density will improve the total throughput. But, when the user density is large, increasing the user density further will result in a higher collision probability, and hence, slow down the total throughput improvement.

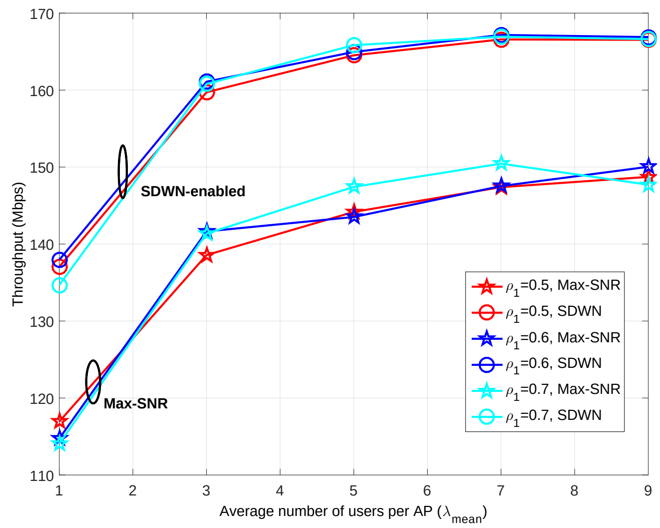
The second example in Fig. 4c measures the fairness by employing the Jain's fairness index over $T_g = \sum_{\forall n_g \in \mathcal{N}_g} \sum_{\forall a \in \mathcal{A}} T^a_{n_g}$, which is the achieved throughput for all the users of SP g. It should be noted that $\mathcal{N}_g$ denotes the set of users belonging to SP g and $\mathcal{A}$ represents the set of all APs in the networks. From Fig. 4c, it is clear that the proposed SDWN-enabled association approach can always guarantee perfect fairness between the SPs regardless of the user density or ρ_1. This can be explained by the fact that the SDWN-enabled algorithm provisions the minimum airtime reservations for each SP based on the constraint (5b). Since in this example the target airtime share for each SP is the same, the SDWN-enabled algorithm guarantees that SPs share the total airtime in a fair manner. However, the achieved fairness level by Max-SNR association is always worse than SDWN-enabled, especially when the user load is highly unbalanced between SPs (i.e., ρ_1 is not close to 0.5). This is because the max-SNR algorithm is oblivious to the user density of each slice, user distribution among APs, and minimum reservations of each SP. Therefore, in unbalanced situations, the max-SNR user association results in unfair throughput achievements by different SPs.

4.2 User Association and Resource Allocation in Virtualized Cellular Networks

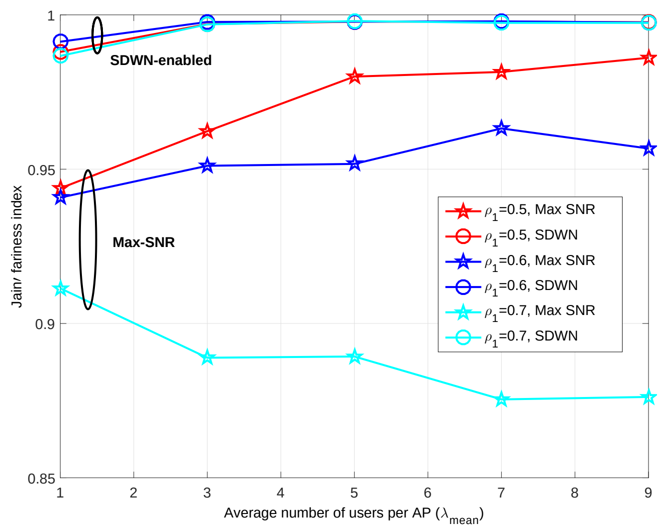
Here, we study one of the applications of SDWN in a cellular radio access network for traffic shaping at the user level. Using the SDWN and the centralized view it provides, data plane cooperation across base stations (BSs) can be realized for performance optimization in a multi-cell scenario. One of such optimization is coordinated



(a) Illustration of virtualized 802.11 network



(b) Total throughput vs. user density for different ρ_1



(c) Fairness vs. user density for different ρ_1

Fig. 4: System model and performance evaluation for homogeneous virtualized 802.11 networks

association control in virtualized cellular networks, aiming to balance the load, manage the intra-cell interference, and guarantee QoS requirements customized for each SP.

By allowing data to be controlled centrally, we present a cohesive approach on wireless resource management (power and sub-carrier allocation) and user association. Specifically, in a multi-cell OFDMA-based network, we optimally solve the problem of joint user association and power/sub-carrier assignment.

We consider the downlink transmission of a virtualized wireless network, where the coverage of a specific area is provided by a set of BSs, i.e., $\mathcal{M} = \{1, \ldots, M\}$. The total bandwidth of B Hz is divided into a set of sub-carriers, $\mathcal{K} = \{1, \ldots, K\}$ and shared by all BSs through orthogonal frequency-division multiple-access (OFDMA). The bandwidth of each sub-carrier, i.e., $B_c = \frac{B}{K}$, is assumed to be much smaller than the coherent bandwidth of the wireless channel, so that the channel response in each sub-carrier is flat. This set of BSs serves a set of slices, $\mathcal{G} = \{1, \ldots, G\}$, where the slice g has a set of users $\mathcal{N}_g = \{1, \ldots, N_g\}$ and requests for a minimum reserved rate of R_g^{rsv} and $N = \sum_{g \in \mathcal{G}} N_g$ is the total number of users.

Let h_{m,k,n_g} and P_{m,k,n_g} be the channel power gain (also representing the channel state information (CSI)), and the allocated power, respectively, of the link from BS $m \in \mathcal{M}$ to user n_g of slice g on sub-carrier k. Due to the OFDMA limitation, each user is assigned to one BS, and to avoid intra-cell interference, orthogonal sub-carrier assignment is assumed among users in a cell. The binary-valued user association factor (UAF) $\beta_{m,k,n_g} \in \{0, 1\}$ represents both sub-carrier allocation and BS assignment indicator for user n_g of slice g on sub-carrier k of BS m. The UAF is set to one, i.e., $\beta_{m,k,n_g} = 1$, when BS m allocates sub-carrier k to user n_g, and $\beta_{m,k,n_g} = 0$, otherwise. Consider $\mathbf{P} = [P_{m,k,n_g}]_{\forall m,g,n_g,k}$ and $\boldsymbol{\beta} = [\beta_{m,k,n_g}]_{\forall m,g,n_g,k}$ as the vectors of all transmit powers and UAFs of users, respectively.

Consequently, the transmission rate of user n_g at sub-carrier k of BS m can be expressed as

$$R_{m,k,n_g}(\mathbf{P}) = \log_2 \left[1 + \frac{P_{m,k,n_g} h_{m,k,n_g}}{\sigma^2 + I_{m,k,n_g}} \right], \tag{6}$$

where

$$I_{m,k,n_g} = \sum_{\forall m' \in \mathcal{M}, m' \neq m} \sum_{\forall g \in \mathcal{G}} \sum_{\forall n'_g \in \mathcal{N}_g, n'_g \neq n_g} P_{m',k,n'_g} h_{m,k,n'_g}$$

is the interference incured to user n_g in cell m and sub-carrier k, and σ^2 is the noise power. Without loss of generality, noise power is assumed to be equal for all users in all sub-carriers and BSs.

Subsequently, aiming to maximize the total transmission rate, we formulate the joint power, sub-carrier and BS assignment as

$$\max_{\boldsymbol{\beta}, \mathbf{P}} \sum_{m \in \mathcal{M}} \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} \beta_{m,k,n_g} R_{m,k,n_g}(\mathbf{P}), \tag{7}$$

subject to:

$$\tilde{\text{C}}1: \sum_{m \in \mathcal{M}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in \mathcal{K}} \beta_{m,k,n_g} R_{m,k,n_g}(\mathbf{P}) \geq R_g^{\text{rsv}}, \forall g \in \mathcal{G}$$

$$\tilde{\text{C}}2: \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \sum_{k \in K} P_{m,k,n_g} \leq P_m^{\max}, \quad \forall m \in \mathcal{M},$$

$$\tilde{\text{C}}3: \sum_{g \in \mathcal{G}} \sum_{n_g \in \mathcal{N}_g} \beta_{m,k,n_g} \leq 1, \qquad \forall m \in \mathcal{M}, \quad \forall k \in \mathcal{K}.$$

$$\tilde{\text{C}}4: [\sum_{k \in \mathcal{K}} \beta_{m,k,n_g}][\sum_{\forall m' \neq m} \sum_{k \in \mathcal{K}} \beta_{m',k,n_g}] = 0,$$

$$\forall n_g \in \mathcal{N}_g, \ \forall g \in \mathcal{G}, \ \forall m \in \mathcal{M}.$$

where $\tilde{\text{C}}1$ represents the required minimum rate of slice $g \in \mathcal{G}$, $\tilde{\text{C}}2$ expresses the maximum transmit power limitation of each BS where $P_m^{\max}$ is the maximum transmit power of BS m, $\tilde{\text{C}}3$ guarantees the OFDMA exclusive sub-carrier allocation within each cell m, and

$\tilde{C}4$ implies that each user can be associated to only one BS. More specifically, $\tilde{C}4$ ensures when any sub-carrier k is assigned to user n_g by BS m, that user would not be assigned any sub-carriers by other BSs m'.

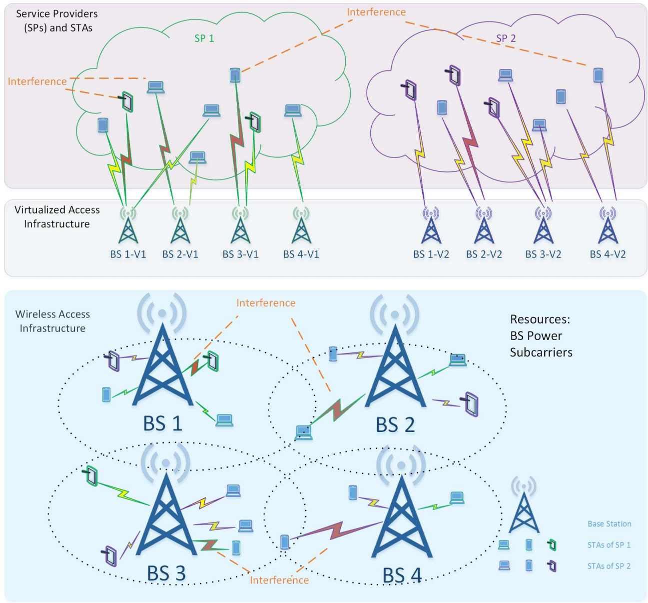
The optimization problem (7) has a non-convex objective function due to inter-cell interference and involves non-linear constraints with combination of continuous and binary variables, i.e., $\mathbf{P}$ and $\boldsymbol{\beta}$. As the proposed optimization problem is inherently non-convex and NP-hard, by applying the successive convex approximation (SCA) and complementary geometric programming (CGP), we develop an efficient iterative approach with low computational complexity to solve the proposed problem. This algorithm SDNW-enabled algorithm searches for jointly optimal user-association, power, and sub-carrier allocations. Details of the developed iterative algorithm to obtain optimal resource allocation can be found in our study [38].

Here, the performance of proposed SDWN-enabled and conventional Max-SNR association approaches are compared in two examples. In these two examples, i.e., Fig. 5b and Fig. 5c, the simulation setting assumes a virtualized cellular network with $M = 4$ BSs with $K = 4$ sub-carriers serving $G = 2$ slices (service providers) in a 2×2 square area (See Fig. 5a). The 4 BSs are located at coordinates: (0.5, 0.5), (0.5, 1.5), (1.5, 0.5) and (1.5, 1.5). The channel power gains are based on the path loss and Rayleigh fading model, i.e., $h_{m,k,n_g} = \chi_{m,k,n_g} d_{m,n_g}^{-\alpha}$ where $\alpha = 3$ is the path loss exponent, $d_{m,n_g} > 0$ is the distance between the BS m and user n_g and χ_{m,k,n_g} is the exponential random variable with mean of 1. We use the noise power in a sub-carrier bandwidth as reference (i.e., normalized to 1 or 0 dB) and hence express transmit power or interference power in dB relative to noise power. The simulation results are taken over the average of 100 different channel realizations. For all the following simulations, we set $R^{\text{rsv}} = R_g^{\text{rsv}}$ for all $g \in \mathcal{G}$ and $P^{\max} = P_m^{\max}$ for all $m \in \mathcal{M}$.

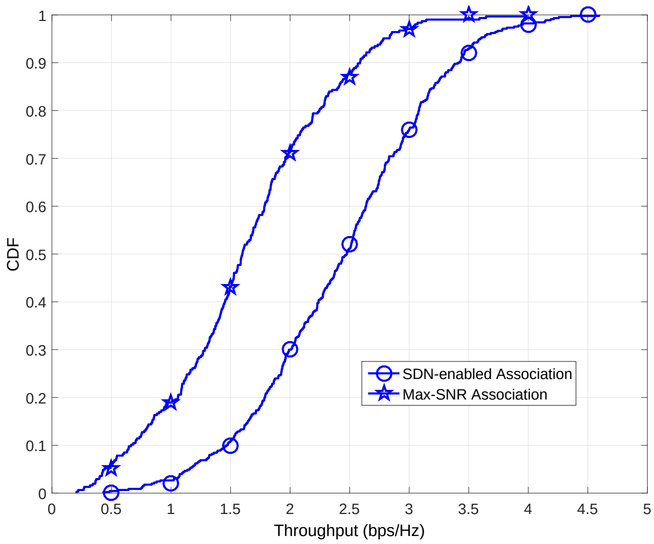
In any cellular network, the coverage is one of the most important planning parameters which can be measured by the signal to interference plus noise ratio (SINR) or achieved total throughput of users at the cell boundaries. To study the performance of SDWN-enabled association control to increase the coverage of our scenario, we consider the simulation setup in which majority of users' are located in the cell-edge, consequently, these users experience high interference from other BSs.

The first example in Fig. 5b demonstrates the cumulative distribution function (CDF) of the total throughput of cell-edge users for both the SDWN-enabled and Max-SNR algorithms. It can be seen that the SDWN-enabled association outperforms the Max-SNR for the cell-edge users where 50% of users in the cell-edge achieve a throughput of 2.5 bps/Hz in the case of SDWN-enabled association, while their throughput is around 1.5 bps/Hz in the case of Max-SNR. However, the performance of both algorithms is similar for the cell-center users. It is because via user-association in SDWN-enabled association, the interference among different cells will be controlled through joint allocation optimization, while Max-SNR cannot control the interference since the association of each user is predetermined based the received SNR levels of reference signals regardless of other users' connectivity. In other words, the max-SNR makes the association decision separatley for each individual user and does not take into account the user distribution and/or the slice requirements. Therefore, depending on the user distribution, the system could end up with uneven load balancing between BSs and severe interference issues at boundaries of different cells. Consequently, SDWN-enabled association can provide better coverage even for cell-edge users for virtualized multi-cellular networks, which is very desirable from implementation perspective.

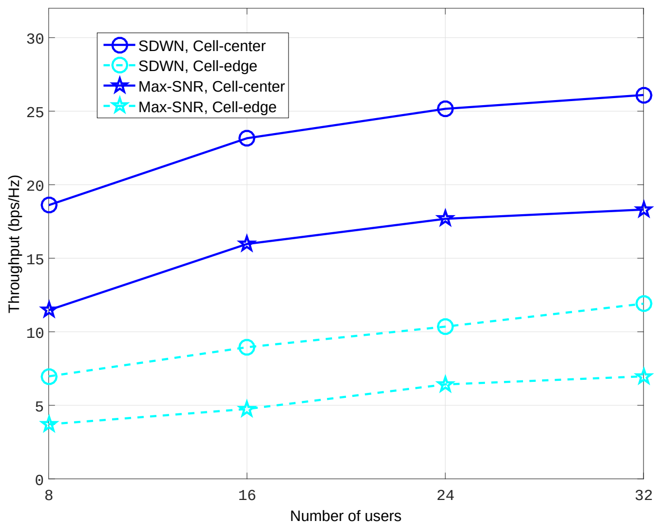
The second example in Fig. 5c investigates the total achieved throughput with respect to the number of users at the cell edge. Obviously, SDWN-enabled association can consistently improve the performance of cell-edge users and maintain the desirable throughput of overall networks regardless of the user deployment density as compared to the Max-SNR. This can be explained by the fact that the SDWN-enabled algorithm solves the optimization problem in (7) and maximizes the total transmission user rates, while max-SNR heuristically connects each user to the BS that offers the



(a) Illustration of virtualized cellular network



(b) Throughput distribution for cell-edge users



(c) Total throughput achieved vs. number of users

Fig. 5: System model and performance evaluation for homogeneous virtualized cellular networks

highest SINR. Moreover, the SDWN-enabled guarantees minimum rate requirements for each slice, applicable to its users located in both the cell-edge or cell-center areas. These two points justify the achieved throughput gain achieved by the SDWN-enabled algorithm.

It should be noted that in the two presented case-studies the network dynamics are reflected by channel dynamics. The proposed user association algorithms need to be run every channel coherence time. Thus, user mobility would automatically be captured and addressed by the change of channel gains. The mobile user who experienced different channel gains toward an AP will be associated with a different AP in a new instance of problem running. The handover procedure is not the focus of this work, but in a virtualized network with a central controller handover can be effectively managed as explained in [22].

5 Conclusion and Future Works

In this paper, a simple, efficient and integrated structure of 5G-SDWN has been proposed based on its three pillars, SDN, SDR, and virtualization. For this networking paradigm, we have provided a general architecture and introduced a framework for converged multi-layer resource management enabled by the 5G-SDWN structure. Via two case-studies, we have highlighted advantages of such resource management for 5G-SDWN. In particular, it is shown how it can increase the coverage and capacity while providing the requested QoS.

We believe that the CML-RM over 5G-SDWN is at the first stage of development, encountering numerous issues and encouraging for future research. From the soft, integrated, centralized and cross-layer structure of 5G-SDWN, it is expected that CML-RM cannot be an evolutionary version of traditional problems in wireless networks, while it needs more revolutionary movement and thoughts.

In addition to the discussed issues, another implementation challenge is deriving all the required information for each user. Clearly, network can hardly obtain the perfect and complete information. Therefore, robust and learning approaches are required to consider the uncertainty on the system parameters and to derive all individual features of users or system information. Solving this type of general optimization problems necessitates more sophisticated mathematical tools and programming algorithms. Last but not least, cooperation and connection of 5G-SDWN with traditional wireless networks cause new challenges.

In the future, we shall focus on the performance evaluation of 5G-SDWN considering other parameters from transmission and control planes. The combination of new radio transmission concepts over access of 5G-SDWN, such as full-duplex, massive MIMO, and device-to-device communications, and more advanced system model, e.g., heterogeneous cloud radio access network (CRAN), will be investigated as the future extension of proposed two case studies. Moreover, integrating the fronthaul and backhaul network segments under the concept of 5G-Crosshaul [39] would also introduce a new degree of freedom which enables full cross-layer and end-to-end CML-RM. Introducing new sets of variables, new constraints and even new objective functions, centralized resource allocation problems can be investigated in a cross-haul and cross-layer structure.

6 References

1 Liang, C., Yu, F., Zhang, X.: 'Information-centric network function virtualization over 5G mobile wireless networks', IEEE Network, 29, (3), pp. 68–74, May 2015.
2 Hossain, E, Hasan, M.: '5G cellular: key enabling technologies and research challenges', IEEE Instrumentation Measurement Magazine, 18, (3), pp. 11–21, June 2015.
3 Sun S., Kadoch, M., Gong, L., Rong, B.: 'Integrating network function virtualization with SDR and SDN for 4G/5G networks', IEEE Network, 29, (3), pp. 54–59, May 2015.
4 Zhang, N., Cheng, N., Gamage, A., Zhang, K., Mark, J., Shen, X.: 'Cloud assisted HetNets toward 5G wireless networks', IEEE Communications Magazine, 53, (6), pp. 59–65, June 2015.
5 Peng, M., Li, Y., Zhao, Z., Wang, C.: 'System architecture and key technologies for 5G heterogeneous cloud radio access networks', IEEE Network, 29, (2), pp. 6–14, 2015.
6 Macedo, D., Guedes, D., Vieira, L., Vieira, M., Nogueira, M.: 'Programmable networks-from software defined radio to software defined networking', IEEE Commun. Surveys Tuts., 17, (2), pp. 1102–1125, 2015.
7 Gerasimenko, M., Moltchanov, D., Florea, R., Andreev, S., Koucheryavy, Y., Himayat, N., Yeh, S.-P., Talwar, S.: 'Cooperative radio resource management in heterogeneous cloud radio access networks', IEEE Access, 3, pp. 397–406, 2015.
8 Sama, M. R., Contreras, L. M., Kaippallimalil, J., Akiyoshi, I., Qian, H., Ni, H.: 'Software-defined control of the virtualized mobile packet core', IEEE Commun. Mag., 53, (2), pp. 107–115, Feb. 2015.
9 Bernardos, C., De La Oliva, A., Serrano, P., Banchs, A., Contreras, L., Jin, H., Zuniga, J.: 'An architecture for software defined wireless networking', IEEE Wireless Commun., 21, (3), pp. 52–61, Jun. 2014.
10 Niephaus, C., Ghinea, G., Aliu, O., Hadzic, S., Kretschmer, M.: 'SDN in the wireless context - Towards full programmability of wireless network elements', in IEEE Conf. Network Softwarization (NetSoft), April 2015, pp. 1–6.
11 Costa-Requena, J., Guasch, V. F., Santos, J. L.: 'Software defined networks based 5g backhaul architecture', in Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication. ACM, 2015, p. 35.
12 Wen, H., Tiwary, P. K., Le-Ngoc, T.: 'Wireless Virtualization', Springer, 2013.
13 Wen, H., Tiwary, P. K., Le-Ngoc, T.: 'Multi-perspective virtualization and software-defined infrastructure framework for wireless access networks', Mobile Networks and Applications, 20, (1), pp. 19–31, 2015.
14 Liang, C., Yu, F.: 'Wireless network virtualization: A survey, some research issues and challenges', IEEE Commun. Surveys Tuts., 17, 1, pp. 358–380, Firstquarter 2015.
15 Kokku, R., *et al.*, 'NVS: A substrate for virtualizing wireless resources in cellular networks', IEEE/ACM Trans. Netw., 20, (5), Oct. 2012.
16 Kim, H., Feamster, N.: 'Improving network management with software defined networking', IEEE Commun. Mag., 51, (2), pp. 114–119, Feb. 2013.
17 Hu, F., Hao, Q., Bao, K.: 'A survey on software-defined network and OpenFlow: From concept to implementation', IEEE Commun. Surveys Tuts., 16, (4), pp. 2181–2206, Fourthquarter 2014.
18 Leon-Garcia, A., Mason, L. G.: 'Virtual network resource management for next-generation networks', IEEE Communications Magazine, 41, (7), pp. 102–109, 2003.
19 Lin, T., Kang, J.-M., Bannazadeh, H., Leon-Garcia, A.: 'Enabling SDN applications on software-defined infrastructure', in IEEE Network Operations and Management Symposium (NOMS), 2014, pp. 1–7.
20 Kang, J.-M., Bannazadeh, H., Leon-Garcia, A.: 'SAVI testbed: Control and management of converged virtual ICT resources', in IFIP/IEEE Intl. Symposium Integrated Network Management, 2013, pp. 664–667.
21 Olwal, T. O., Djouani, K., Kurien, A. M.: 'A survey of resource management toward 5G radio access networks', IEEE Communications Surveys Tutorials, 18, (3), pp. 1656–1686, thirdquarter 2016.
22 Zhang, H., Liu, N., Chu, X., Long, K., Aghvami, A. H., and Leung, V. C. M., 'Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges', IEEE Communications Magazine, 55, (8), pp. 138–145, 2017.
23 Peng, M., Sun, Y., Li, X., Mao, Z., and Wang, C., 'Recent advances in cloud radio access networks: System architectures, key techniques, and open issues', IEEE Communications Surveys Tutorials, 18, 3, pp. 2282–2308, thirdquarter 2016.
24 Wang, W., Chen, Y., Zhang, Q., and Jiang, T., 'A software-defined wireless networking enabled spectrum management architecture', IEEE Communications Magazine, 54, (1), pp. 33–39, January 2016.
25 Liu, D., Wang, L., Chen, Y., Elkashlan, M., Wong, K.-K., Schober, R., and Hanzo, L., 'User association in 5G networks: A survey and an outlook', IEEE Communications Surveys & Tutorials, 18, (2), pp. 1018–1044, 2016.
26 Tseliou, G., Adelantado, F., and Verikoukis, C., 'Scalable RAN virtualization in multitenant LTE-A heterogeneous networks', IEEE Transactions on Vehicular Technology, 65, (8), pp. 6651–6664, August 2016.
27 Parsaeefard, S., Dawadi, R., Derakhshani, M., Le-Ngoc, T., and Baghani, M., 'Dynamic resource allocation for virtualized wireless networks in massive-mimo-aided and fronthaul-limited c-ran', IEEE Transactions on Vehicular Technology, 66, (10), pp. 9512–9520, Oct 2017.
28 Peng, M., Li, Y., Zhao, Z., and Wang, C., 'System architecture and key technologies for 5G heterogeneous cloud radio access networks', IEEE Network, 29, (2), pp. 6–14, Mar. 2015.
29 Perez-Romero, J., Gelabert, X., and Sallent, O., 'Radio resource management for heterogeneous wireless access networks', Heterogeneous Wireless Access Networks Architectures and Protocols, Hossain, E.(Ed.). Springer Science, New York, pp. 133–166, 2008.
30 Khatibi, S., and Correia, L. M., 'Modelling of virtual radio resource management for cellular heterogeneous access networks', in Proc. of PIMRC, Sep. 2014.
31 Caeiro, L., Cardoso, F. D., and Correia, L. M., 'Ondemand virtual radio resource allocation for wireless access', Wireless Personal Communications, 82, (4), pp. 2431–2456, 2015.
32 Demestichas, P., Georgakopoulos, A., Karvounas, D., Tsagkaris, K., Stavroulaki, V., Lu, J., Xiong, C., and Yao, J., '5G on the horizon: Key challenges for the radio-access network', IEEE Vehicular Technology Magazine, 8, (3), pp. 47–53, Sep. 2013.
33 Gotsis, A., Stefanatos, S., and Alexiou, A., 'Optimal user association for massive mimo empowered ultra-dense wireless networks', in Communication Workshop (ICCW), 2015 IEEE International Conference on, Jun. 2015, pp. 2238–2244.
34 Liu, D., Wang, L., Chen, Y., Zhang, T., Chai, K. K., and Elkashlan, M., 'Distributed energy efficient fair user association in massive mimo enabled hetnets', IEEE Communications Letters, 19, (10), pp. 1770–1773, Oct. 2015.
35 Lim, Y.-G., Chae, C.-B., and Caire, G., 'Performance analysis of massive MIMO for cell-boundary users', IEEE Transactions on Wireless Communications, 14, 12, pp. 6827–6842, 2015.

36 Bethanabhotla, D., Bursalioglu, O. Y., Papadopoulos, H. C., Caire, G.: 'Optimal user-cell association for massive MIMO wireless networks', IEEE Transactions on Wireless Communications, 15, (3), pp. 1835–1850, 2016.
37 Derakhshani, M., Wang, X., Le-Ngoc, T., Leon-Garcia, A.: 'Airtime usage control in virtualized multi-cell 802.11 networks', in 2015 IEEE Globecom Workshops (GC Wkshps), Dec 2015, pp. 1–6.
38 Parsaeefard, S., Dawadi, R., Derakhshani, M., Le-Ngoc, T.: 'Joint user-association and resource-allocation in virtualized wireless networks', IEEE Access, 4, pp. 2738 – 2750, April 2016.
39 González, S. Oliva, A., Costa-Pérez, X., Di Giglio, A., Cavaliere, F., Deiß, T., Li, X., Mourad, A.: '5g-crosshaul: An sdn/nfv control and data plane architecture for the 5g integrated fronthaul/backhaul', Transactions on Emerging Telecommunications Technologies, 27, (9), pp. 1196–1205, 2016.