Photonic Network Communications (2023) 45:49–66
https://doi.org/10.1007/s11107-022-00991-x

ORIGINAL PAPER

OTN-over-WDM optimization in 5G networks: key challenges and innovation opportunities

Abdallah Moubayed[1,2] · Dimitrios Michael Manias[1] · Abbas Javadtalab[3] · Mahdi Hemmati[3] · Yuren You[3] · Abdallah Shami[1]

Received: 11 March 2022 / Accepted: 20 December 2022 / Published online: 31 January 2023

Abstract

The continued growth of both mobile broadband and fixed broadband subscriptions as well as the added deployment of Internet of Things devices has led to making 5G networks a reality. More specifically, 5G networks are expected to support a diverse set of new applications/services in addition to existing applications/services from previous generations (2G/3G/4G). The COVID-19 pandemic has further increased the demand for such services which has resulted in a further surge in the Internet usage. Thus, 5G networks are expected to have a highly flexible architecture at all levels including at the radio, core, and transport levels. Optical Transport Networks (OTN) have been proposed as one potential and promising supporting technology for 5G networks at the transport level, particularly for next generation transport networks featuring large-granule broadband service transmissions. This is because it allows for more flexible, efficient, and dynamic networks. However, adopting and deploying OTNs in 5G networks comes with its own set of challenges including control, management, and orchestration of such networks as well as their security. Accordingly, this paper overviews 5G networks along with their requirements and provides a brief summary of OTNs and the corresponding optimization mechanisms. Additionally, this work discusses the challenges facing OTNs and their optimization within the context of 5G. Moreover, it outlines some of the key research areas and opportunities for innovation stemming from the data-driven intelligent networking paradigm using Machine Learning techniques.

Keywords OTN-over-WDM optimization · 5G networks · Challenges · Innovation opportunities

1 Introduction

The continued growth of both mobile broadband and fixed broadband subscriptions as well as the added deployment of Internet of Things (IoT) devices has led to making 5G networks a reality [1]. More specifically, 5G networks are expected to support a diverse set of new applications/services such as Augmented/Virtual Reality (AR/VR), smart factories, and autonomous vehicles. These applications/services typically fall under one of three major 5G use cases defined by the International Telecommunication Union (ITU), namely: Enhanced Mobile Broadband (eMBB), Massive Machine Type Communication (mMTC), and Ultra-Reliable Low Latency Communications (URLLC) respectively [2]. This is in addition to existing applications/services from previous generations (2G/3G/4G) such as mobile voice, messaging, and Internet access [3]. The COVID-19 pandemic has further increased the demand for such services which has resulted in a further surge in the Internet

✉ Abdallah Moubayed
amoubaye@uwo.ca; abdallah.moubayed@lau.edu.lb

Dimitrios Michael Manias
dmanias3@uwo.ca

Abbas Javadtalab
abbas.javadtalab@huawei.com

Mahdi Hemmati
mahdi.hemmati@huawei.com

Yuren You
yuren.you@huawei.com

Abdallah Shami
abdallah.shami@uwo.ca

1 Western University, London, Ontario, Canada

2 Lebanese American University, Blat Area, Byblos, Lebanon

3 Huawei Technologies Canada Research Center, Ottawa, Ontario, Canada

usage. For example, Austria reported that the percentage of individuals using the Internet to make audio or video calls increased from 41 to 60% between 2019 and 2020 [1]. Thus, 5G networks are expected to have a highly flexible architecture at all levels including at the radio, core, and transport levels [3]. Moreover, this entails having a high automation level in the deployment and maintenance of networks, parts of a network. or single resources (e.g., network slices).

Optical Transport Networks (OTN) have been proposed as one potential and promising supporting technology for 5G networks at the transport level, particularly for next generation transport networks featuring large-granule broadband service transmissions [4–8]. This is due to the fact that OTN allow for more productive multiplexity and switching of high-bandwidth signals and provide the capability of cross connect dispatching of wavelengths and sub-wavelengths resulting in efficient wavelength utilization [8]. This allows for the decoupling of the clients from the Dense Wavelength Division Multiplexing (DWDM) line interfaces. In turn, this improves the network efficiency because the utilization rate of the more costly DWDM links is significantly improved which means that we do not have any stranded bandwidth [9]. Moreover, it builds on previous technologies such as SONETs and SDH, providing an added level of transparency as well as end-to-end connection and networking capabilities for WDM with reduced complexity [10, 11]. Additionally, this offers enhanced protection capabilities through the use of Forward Error Correction (FEC). OTN also provides more flexible service protection functions based on electrical and optical layers [10].

However, adopting and deploying OTNs in 5G networks comes with its own set of challenges. This is exacerbated by the 5G radio access network (RAN) deployment architecture and the corresponding function splits adopted. One such challenge is the added complexity, particularly in terms of managing these networks. The control, management, and orchestration of such networks continues to evolve for fast provisioning of light-paths, fast restoration and high availability [12]. This is because such networks have to deal with dynamic elasticity requirements of new applications/services. Moreover, they need to effectively and efficiently scale with both planned (predictable traffic) and unplanned (unpredictable traffic surges) traffic loads that can significantly degrade the services' QoS performance. Since the manual configuration of such networks can lead to slow service provisioning, automated configuration and management of optical networks allows carrier networks based on optical paths to become more scalable, dynamic, and manageable [12]. In turn, this would allow the networks that are being built upon OTNs to reach their full performance potential and meet the 5G service requirements by reducing latencies, increasing throughput, and improving the reliability of the overall architecture [13]. Another challenge is the security of OTNs, particularly against both passive and active optical attacks. Moreover, with the introduction of new services and applications, the potential attack surfaces will also increase, resulting in additional security threats. Thus, the security of OTNs need to be considered when optimization decisions are being made.

Therefore, there are multiple challenges to consider when proposing OTN networks as part of future 5G deployments that need to be addressed. Accordingly, this paper focuses on summarizing the challenges facing the deployment of OTN networks in 5G and discussing potential research opportunities to address these challenges using data-driven paradigms and Machine Learning (ML) techniques.

The remainder of this paper is as follows: Sect. 2 provides a brief background about 5G in terms of its use cases and general architecture. Similarly, Sect. 3 succinctly describes the operation and functionality of OTN networks. Section 4 presents the major challenges facing the adoption and deployment of OTNs in 5G networks. Consequently, Sect. 5 discusses some potential research opportunities to address these challenges. Lastly, Sect. 6 presents the paper's conclusion.

2 5G background

5G promises to support new applications and services as well as improved network performance, flexibility and reliability. This is in an effort to deliver an improved end-user experience and provide ubiquitous connectivity to everything and everyone [2]. The goal is to build upon the success of previous generations (2G/3G/4G) by supporting new services and business models, resulting in significant projected economic benefits. In what follows, a brief overview of the 5G use cases and the corresponding 5G architecture (including the RAN architecture functional splits) are presented.

2.1 Use cases

The ITU grouped the new applications and services that 5G networks are expected to support into three main use cases, namely: Enhanced Mobile Broadband (eMBB), Massive Machine Type Communication (mMTC), and Ultra-Reliable Low Latency Communications (URLLC) respectively [2]. Each use case has its own set of requirements and target metrics. Figure 1 summarizes these three use cases and briefly presents some of the applications/services within each.

1. *Enhanced mobile broadband (eMBB)* The first use case is the enhanced mobile broadband, also commonly referred to as eMBB. As the name suggests, this use case focuses on applications and services that require high data rates and high mobility support. This includes

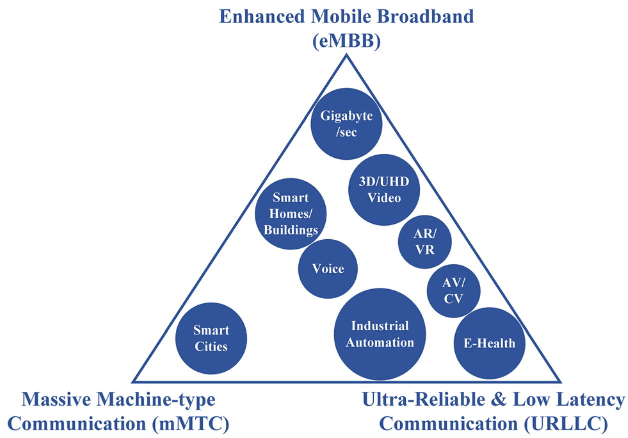


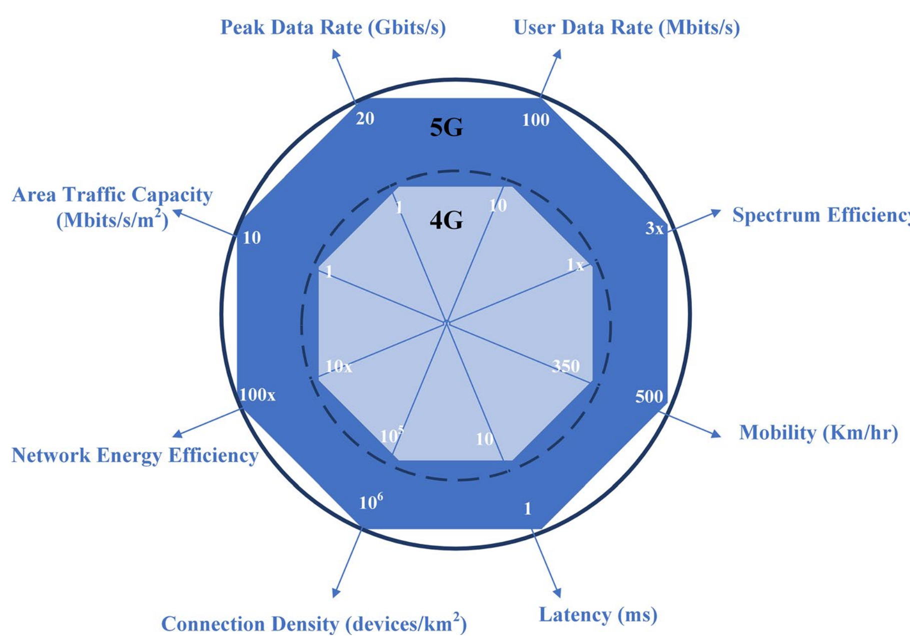
Fig. 1 5G use cases

applications such as 3D or Ultra High Definition (UHD) video and streaming [14]. To illustrate the performance requirement of such applications, it is expected that 5G networks are able to provide a mobility interruption time less than 4 msec and support video resolutions such 1080p, 2K, 4K and also 8K video [14, 15].

2. *Massive machine-type communication (mMTC)* The second use case defined by the ITU is the massive Machine-type Communication, also referred to in the literature as mMTC. This use case focuses on supporting the communication between devices such as smart meters, sensor nodes, and home appliances without human intervention [16, 17]. This will result in the creation of new vertical sector and business models such as smart cities, smart utilities, and smart homes [16]. Since mMTC is expected to provide massive access to a large number of often low-complexity and low-power Machine Type Devices (MTDs), it is crucial to carefully design and deploy the required technologies and infrastructure while considering the varying delay, reliability, energy consumption, complexity, security, and throughput requirements of the different applications/services within this use case category [16, 17].
3. *Ultra-reliable & low latency communication (URLLC)* The third use case proposed by ITU for 5G networks is the ultra-reliable & low latency communication, commonly referred to as URLLC. As part of this use case, new applications and services such as Autonomous/Connected vehicles (AV/CV), industrial automation, and tactile Internet for telemedicine/e-Health have emerged that introduced new stringent requirements with the main performance metrics being latency, reliability, and availability [18, 19]. For example, the latency can range between 1 and 10 msec for industrial automation mission control messaging and AV/CV communication in Intelligent Transportation Systems (ITSs) respectively [19]. Similarly, the reliability requirement of such applications and services can range between 99.9 and 99.9999% for ITS-related applications [19].

Figure 2 shows the expected capability enhancements from 5G networks in terms of the different performance metrics such as peak data rate, spectrum efficiency, energy efficiency, and latency.

Fig. 2 5G expected capability enhancement



2.2 5G architecture

As mentioned earlier, one of the main changes brought on by 5G is in terms of the adopted architecture. It is worth noting that when talking about 5G architectures, two varying concepts emerge, namely the deployment architecture and the core architecture. As the name suggests, the deployment architecture refers to how 5G infrastructure is deployed, particularly in terms of its interaction with previous generation infrastructure(s). On the other hand, the 5G core architecture refers to the manner in which the 5G network is built and how its elements interact to offer the service/application. A brief summary of both the deployment architecture and core architecture is provided.

1. *RAN deployment architecture* The 3GPP defined two main deployment architectures for 5G networks, namely the Non-Standalone (NSA) and the Standalone (SA) architectures [21, 22]. The NSA architecture refers to the scenario where the 5G RAN and its corresponding New Radio (NR) interface are used in conjunction with the existing LTE and EPC infrastructure core network [21]. This allows NR technology to be available without the need to have any network replacements. Consequently, the previously supported 4G services would be able to enjoy the capacities offered by the 5G NR [22]. In contrast, the SA refers to the scenario where the 5G RAN and its corresponding NR interface are only connected to the 5G core network [21]. This configuration thus allows all the 5G services to be supported. The advantage of the SA deployment architecture is that it is fully virtualized, facilitates network slicing, and allows network operators to offer new services with a variety of deployment models (e.g. on-premises private cloud, public cloud, or hybrid) [22].

Figure 3 illustrates the possible RAN architectures including the function splits for the gNBs [20]. Such splits offer multiple economic and operational benefits including increased flexibility and modularity, simplified radio site engineering with reduced space and power demands, more efficient radio resource coordination, and more efficient fiber plant usage [20].

However, depending on the adopted RAN architecture and the corresponding function split, the requirements on the xHaul (i.e. fronthaul, midhaul, and backhaul) may differ. For example, adopting higher-layer functional splits reduces the transport capacity requirements. One example of such a split is the Distributed 4G/5G RAN architecture. In this case, all baseband processing functionalities are placed within the RU [23]. As such, more processing occurs at the RU which means that less data has to travel between the RU and the CU. In turn, this leads to lower bandwidth requirements from the fronthaul with higher latency tolerances. In contrast, if the function splits are adopted at lower layers, higher transport capacity will be required [23]. An example of such a high demanding functional split is the centralized 5G RAN architecture. In this case, more data has to flow between the CU and RU since lower layer functions are more data intensive [23]. Additionally, the communication between some of the lower layer functionalities (e.g. Physical layer) and higher layer processes (e.g. hybrid ARQ) have stringent latency requirements (as low as 5ms in some cases [23]). As a result, the bandwidth and latency demands are higher and more stringent for such architectures. Therefore, the 5G RAN architecture adopted needs to be considered when designing and optimizing transport networks as it has a direct impact on the expected performance requirements.

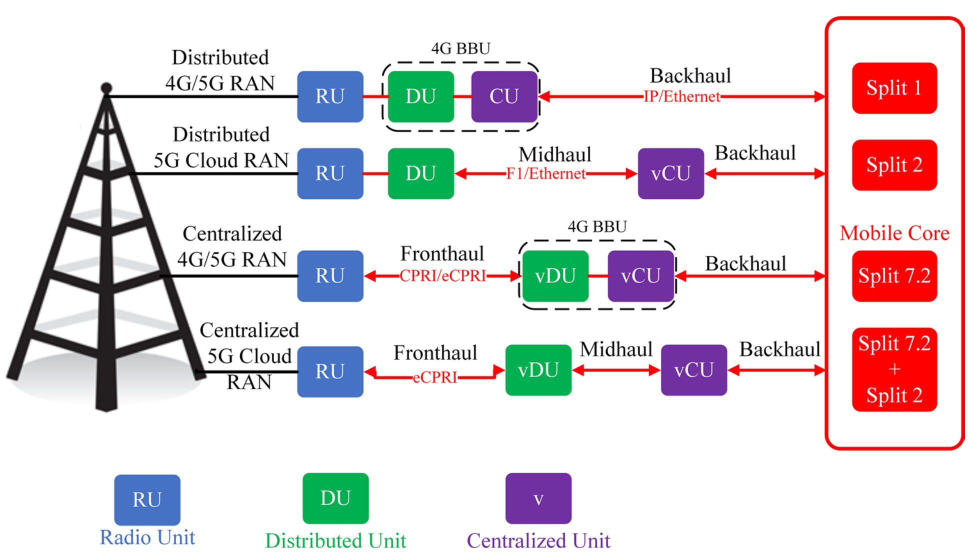


Fig. 3 5G RAN architecture (adapted from [20])

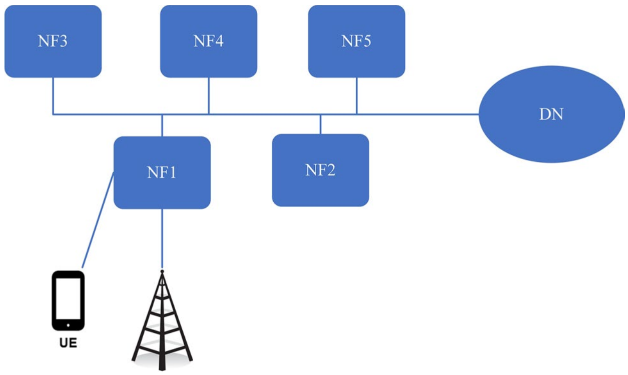


Fig. 4 5G SBA core architecture—global view

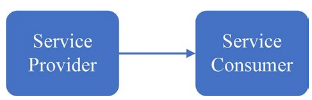


Fig. 5 5G SBA core architecture—granularity view

2. *Core architecture* 5G core networks follow what is known as a Service-Based Architecture (SBA). This architecture relies on the concept of microservices, an approach in which an application is comprised of smaller independent services that interact with each other [24]. As illustrated in Fig. 4, this architecture assumes that each network element performs a particular Network Function (NF), with the different NFs communicating via well-defined APIs [25]. Moreover, these NFs can also communicate with the Data Network (DN) using these interfaces. Each of these NFs is composed of a set of microservices that can be executed/called on-demand. The 5G core SBA architecture can also be viewed in terms of a services "consumer" and a service "provider" as illustrated in Fig. 5. In this case, one NF can act as the service provider by offering a set of microservices to be called/requested by other NFs (service consumers).

   This SBA architecture again builds on the foundations of SDN and NFV technologies/paradigms to offer better network management, flexibility, and scalability that allows 5G networks to support the different emerging applications and services.

3 OTN optimization background

3.1 OTN preliminaries & technologies

Optical Transport Networks (OTNs) were defined as a standard by the International Telecommunications Union in report ITU-T Rec. G.709 [26]. This standard proposes using Wavelength Division Multiplexing (WDM) to better multiplex a substantial number of signals onto a single fiber [26]. This had several advantages. More specifically, OTNs reduced the complexity for transport applications and incorporated overhead that is optimized for transporting signals over WDM-based networks [27]. In turn, this reduced the corresponding transport network operations expenses. Moreover, OTNs provided a more scalable and cost-efficient solution for carrying high-speed Wide Area Network (WAN) data clients including Ethernet and Storage Area Network (SAN) protocols [27]. As such, optimizing OTN networks is a crucial component to allow providers to reap its benefits.

One of the fundamental components of OTN networks are optical Data Units (ODUs). ODUs are basically the data structures generated and monitored by OTNs from one source node to the destination node [28]. This structure offers multiple functionalities such as tandem connection monitoring (TCM) support, path monitoring (PM) and end-to-end path supervision and client adaptation [29]. However, due to the emergence of more diverse 5G applications and to better support diverse client services, OTNs have progressed towards the use of service-based constructs referred to as Optical Service Units (OSUs) [30–33]. OSUs allow for greater flexibility and more fine-grained resource scheduling [30–33]. This is because they create flexible containers that can be either multiplexed into lower order ODUs and then multiplexed into higher order ones or directly be multiplexed into higher order ODUs [30–33]. For example, it would be inefficient to carry a client request with a bit rate less than 1 Gbps using ODUs since the smallest ODU has a rate of approximately 1.25 Gbps. However, using OSUs, flexible bandwidth allocations with approximately 2 Mbps granularity can be achieved, thus increasing the OTN bandwidth utilization when carrying requests with bit rates ranging between 2 and 500 Mbps [30–33]. This makes OSUs a crucial building block for OTNs in 5G networks due to the diverse nature of service requests expected in such networks.

One important technology currently being used in OTNs is space-division multiplexing (SDM). SDM refers to the set of technologies that allow for the transmission of individual data signals over different paths spatially using a shared optical fiber channel. The benefit of these technologies is that they multiply the information carrying capacity of optical fibers [34]. Additionally, SDM can help decrease the energy consumption of the optical nodes as well as improve the overall network efficiency [34]. The utilization of SDM technologies gained more popularity within the last decade due to the fact that single-mode fiber systems based on WDM started to reach their theoretical capacity limits [35]. As a result of the surge in research focusing on developing SDM technologies, it has been illustrated (for example by Soma et al.) that it is possible to achieve a combined data rate greater than 10 Pbps using systems with more than 100 spatial channels with each channel supporting hundreds of WDM channels [36]. Based on the aforementioned benefits, SDM has been increasingly proposed in the literature as a crucial component of OTNs [37, 38]. This is attributed to the fluctuating nature of the traffic that OTNs need to transport. This requires an increase in the capacity and an improvement in the efficiency of the currently deployed WDM systems [38]. Hence, SDM addresses those challenges by significantly increasing the capacity of the fiber and improving the transmission efficiency [38].

3.2 OTN optimization problem description

In general, OTN optimization refers to the problem of determining the number and placement location of the OTN interfaces needed along the available physical links to satisfy the incoming aggregate traffic. The goal is often to minimize the number of interfaces (connecting to OTN nodes) in order to reduce the corresponding cost. This is illustrated in Fig. 6 in which the OTN layer receives the aggregate traffic and tries to find the least number and optimal location for the OTN interfaces to be deployed on the WDM layer below it that is capable of meeting the incoming traffic requirement. Hence,

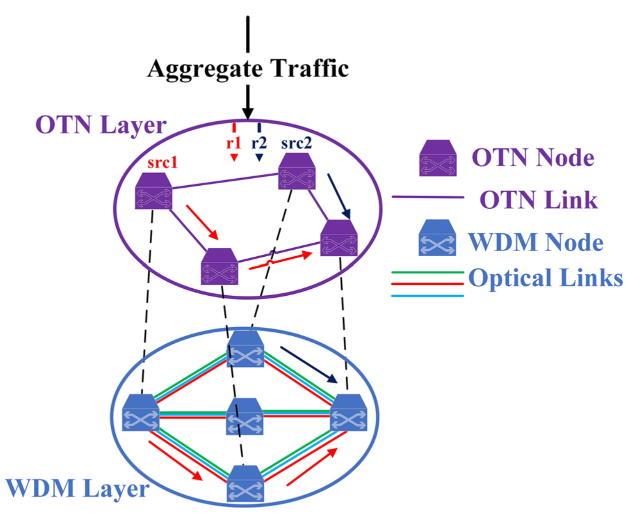


Fig. 6 OTN over WDM general optimization problem

it is assumed that the physical topology of the WDM layer and the incoming traffic matrix of services is known a priori. The goal is to determine the least number of OTN interfaces needed to serve all requests within the traffic matrix.

Mathematically speaking, this can be modeled as follows: Assume we have a traffic matrix consisting of R requests to be served, a set of OTN nodes O, a set of WDM nodes W, a set of OTN links/interfaces L_O with capacity C_O, and a set of optical links L_W with capacity C_w. Using this information, the optimization problem becomes:

- Objective function: $\min_x / \max_x f(x)$ where $f(x)$ is the objective function to be minimized/maximized. This function can be a single function or a multi-objective function such as (but not limited to):
  - Capex cost
  - Resiliency/Survivability
  - Resource Utilization
- Constraints: The optimization problem is subject to a variety of constraints including (but not limited to):
  - Link capacity: $\sum_l \text{demand}_l \leq C_l$ where demand $_l$ is the demand/traffic routed on link l and C_l is the total capacity on the link where $l \in L_O$ or $l \in L_W$. In other words, the aggregate demand along each link (whether at the OTN layer or WDM layer) should not exceed the link's total capacity
  - Flow conservation: $\sum_n \text{flow}_n = -\text{demand}_r$ (if n is source node of request r) or $\sum_n \text{flow}_n = \text{demand}_r$ (if n is destination node of request r) or $\sum_n \text{flow}_n = 0$ (if n is an intermediate node for the path of request r) where flow $_n$ is the aggregate (incoming and outgoing) flow for node $n \in O$ or $n \in W$ and $r \in R$. This means that the flow needs to be conserved at each node (whether at the OTN or WDM layer) between source and destination.

These problems are often formulated and modeled as Integer Linear Programming (ILPs) problems as per the description.

3.3 Related works

Based on the aforementioned benefits of deploying OTNs, significant interest and effort has been given towards optimizing OTNs with multiple works from the literature tackling it [39–49]. As illustrated in Fig. 7, these works fall into one of three main categories, namely: mathematical optimization-based solutions, metaheuristic solutions, and

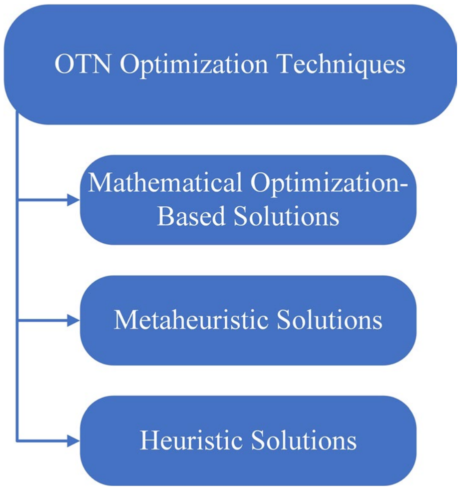


Fig. 7 OTN optimization techniques from the literature

Table 1 Summary of OTN optimization related works

OTN optimization methodology	List of related works
Mathematical optimization-based solutions	[39, 40, 42–45, 48, 49]
Metaheuristic solutions	[46–49]
Heuristic solutions	[39–41, 43, 44]

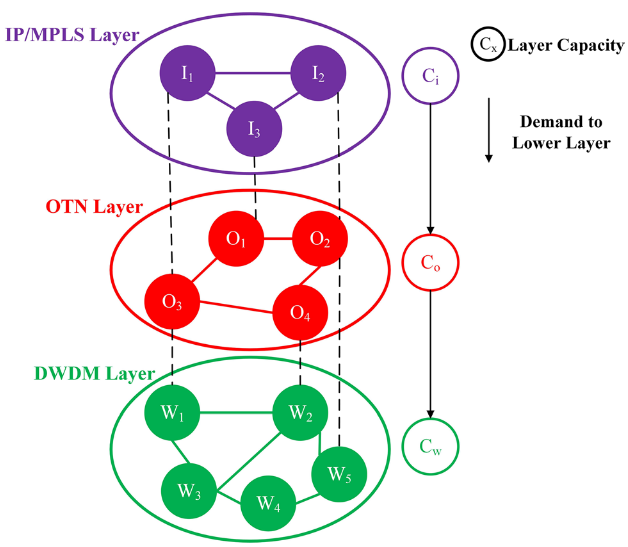


Fig. 8 IP/MPLS over OTN over DWDM network architecture and design approach [39, 40]

low complexity heuristic solutions. Table 1 summarizes the OTN optimization techniques/solutions from the literature.

Katib and Medhi proposed a three-layer architecture consisting of the IP/MPLS layer, the OTN layer, and the Dense Wave Division Multiplexing (DWDM) layer [39]. As illustrated in Fig. 8, the authors presented an integrated multilayer optimization model that aimed at reducing the total network planning cost while adhering to multiple constraints such as tunneling, Internet Protocol/Multiprotocol Label Switching (IP/MPLS) layer capacity, IP/MPLS flow, OTN layer capacity, OTN flow, and DWDM layer capacity constraints [39]. The authors also developed a low complexity heuristic algorithm to solve the optimization model for large networks. Their simulation results showed the impact of each layer's resources and corresponding costs on the neighboring lower layers. The authors extended their work by providing a protection mechanism at each layer [40]. The proposed optimization design model guaranteed the multilayer network survivability when facing three simultaneous link failures (i.e., one single failure at each layer) [40]. In addition to the optimization model formulated, the authors again proposed a three-phase low-complexity algorithm that provided protection at each layer. Simulation results showed that the protection capacity was larger than its regular capacity. This was particularly evident at the DWDM layer because of the longer protection paths at that layer and its larger granularity [40].

In a similar fashion, Govardan et al. proposed a heuristic algorithm to reduce the capital expenditure of a multilayer OTN over DWDM network architecture [41]. In this architecture, the end-to-end services are provisioned while considering the functionalities of both OTN and DWDM technologies. The authors adopted an Integrated OTN-DWDM system which eliminated the fiber interconnections [41]. The heuristic algorithm consisted of five modules. The Input module was made up of the network topology. The topology consists of the input traffic matrix (the demands between OTN nodes) as well as the set of fiber-connected nodes at Layer 0 (DWDM). The Shortest Path module implemented Yen's K-shortest path algorithm. The Amplifier Placement and OSNR Computation module arranged the K shortest paths for each source and destination pair based on the OSNR value. More specifically, the paths are sorted in decreasing value of OSNR computed. The Multilayer Switching Implementation module realized the end-to-end traffic demands on OTN over DWDM network. Finally, the Cost computation module determined the cost of a demand traversing the OTN over DWDM system [41]. Simulation results showed that the proposed heuristic found a balanced path between opaque and transparent network over all traffic volumes. Moreover, they illustrated the suitability of this heuristic for metro networks where the demand relative to the network is typically high.

Zefreh et al. also focused on optimizing IP/optical networks [42]. To that end, the authors developed a Mixed Integer Linear Programming (MILP) problem that aimed at minimizing the CAPEX cost while satisfying the different constraints. This includes flow conservation, number and configuration of input/output ports, number and configuration of routers, number of transponders and wavelengths, number and capacity of multiplexers and demultiplexers, and network cost constraints. Moreover, the authors developed a sub-optimal model that separated the routing from the provisioning to reduce the complexity of the optimization model. Simulation results showed that the proposed sub-optimal model closely followed the optimal solution for a wide range of traffic scaling factors, highlighting its effectiveness.

On the other hand, Papanikolaou et al. hypothesized that following shorter upgrade cycles and adopting a multiperiod network planning approach is more suitable modern, flexible, and software-driven networks [43]. To that end, the authors formulated the incremental multilayer planning problem of an IP over elastic optical network as an Integer Linear Programming (ILP) optimization model. The model's objective was to minimize both the CAPEX and OPEX by deploying the minimum number of additional network equipment (to deal with traffic changes) while simultaneously minimizing the transitioning changes [43]. Simulation results highlighted the impact of each layer's reconfiguration capabilities on the total network cost. Furthermore, they illustrated that short network periods can result in significant reductions in costs due to being able to closely capture the dynamic nature of the traffic [43].

Xing et al. proposed a low-complexity heuristic algorithm that aims at minimizing the network cost by effectively provisioning the resources of the multi-layered OTN network [44]. The authors modeled the network traffic streams as a combination of constant bit-rate and Variable Bit-Rate (VBR) traffic streams. Moreover, the authors adopted a shortest path routing-based solution at each layer in order to address the architectural and traffic requirements [44]. The proposed heuristic followed an iterative approach by provisioning the resources link-by-link in each layer. Simulation results showed that the proposed heuristic algorithm can provision the available resources while considering efficient traffic management. Moreover, it provided insights that telecommunications providers can use as guidance with regard to their choice of technologies for future multi-layered networks [44].

Moniz et al. proposed a network design framework that is dependent on the real-time performance monitoring of the OTN over DWDM network [45]. More specifically, the focus is on effectively operating the OTN network with smaller performance margins. To that end, the authors developed an Integer Linear Programming (ILP) problem that aims to minimize the capital expenditures and reduce the traffic disruption [45]. Simulation results showed that the proposed ILP model was capable of reduce the number of interfaces by up to 32% and the number of mandatory client rerouting by up to 68%, highlighting its effectiveness in meeting its desired objective.

In contrast, Da Silva et al. proposed a metaheuristic-based methodology for OTN over DWDM network planning [46]. Specifically, the authors focused on minimizing the cost of OTN interfaces (the element with the most significant cost) in the network while simultaneously meeting the performance as well as resiliency requirements [46]. To that end, the authors proposed a Multi-Objective Evolutionary Algorithm (MOEA)-based solution, namely Non-dominated Sorting Genetic Algorithm II (NSGAII). The proposed algorithm aimed at minimizing two conflicting objectives. The first is the number of network interfaces and the second is the number of failures in the restoration processes while considering all the possible combination of double failures [46]. Simulation results results illustrated the promise of the proposed NSGAII solution as they showed that it has the potential to obtain optimized solutions for multi-layer scenarios.

The authors further extended their work by proposing a heuristic-based solution for OTN over DWDM network planning [47]. To that end, the authors developed two variations of their heuristic. The first variant delivered a robust solution but did not guarantee that the minimum number of resources is used. The second variant simultaneously minimized the number of resources used and maximized the resilience by better reusing the available components [47]. Simulation results showed that the proposed heuristics closely matched those of the exhaustive-based solution for all traffic conditions in small network sizes. Moreover, in the case of large networks, the heuristics achieved considerable network cost savings (up to 20%) in the different traffic scenarios considered.

In a similar fashion, De Oliveira et al. also proposed a meta-heuristic solution for the OTN over DWDM optimization problem [48]. More specifically, the authors considered the case where the traffic matrix between the demand nodes is known beforehand. The authors first formulated the problem as an ILP with the objective of minimizing the cost. Then, the Firefly algorithm, a swarm-based meta-heuristic algorithm, was proposed to solve the aforementioned ILP problem. Simulation results showed that the proposed Firefly algorithm outperformed the standard genetic algorithm and had close-to-optimal performance.

The authors extended their work further by proposing a hybrid meta-heuristic solution that combines the Firefly algorithm with the standard genetic algorithm [49]. The goal again was to minimize the cost of the OTN over DWDM network planning. Simulation results showed that the proposed hybrid firefly-genetic algorithm outperformed the standalone firefly algorithm and standalone genetic algorithm [49]. Moreover, the hybrid meta-heuristic achieved close-to-optimal performance while having a significantly lower execution time, particularly for large network sizes. This highlighted the proposed algorithm's effectiveness and efficiency in solving the OTN over DWDM network planning problem.

4 Challenges facing OTN optimization in 5G networks

Although OTN networks provide significant advantages such as reduced complexity and network expense cost, optimizing them also comes with its own set of challenges that need to be considered and addressed [50–52]. These challenges are further exacerbated when using OTNs in the context of 5G networks due to the heterogeneous and dynamic nature of such networks. Figure 9 highlights some of the major challenges facing OTN networks and their optimization, particularly in the context of 5G networks.

4.1 Control & management

One of the main challenges facing OTN optimization in 5G networks is how to control and manage the OTNs [51]. In layman terms, this refers to the question "*who will control and manage the network resources?*". This is emphasized further when dealing with the SDN paradigm, a pillar of 5G networks. In this case, different potential control architectures can be considered such as centralized, distributed, and hybrid [51]. However, each of these architectures has its merits and limitations. For example, centralized architectures provide a more comprehensive view of the network and thus can lead to more optimal solutions. However, this comes at the expense of higher computational complexity and signaling overhead. Additionally, a centralized controller represents a single point of failure [53]. In contrast, adopting a distributed control architecture reduces the computational complexity and signaling overhead since each controller would only communicate with a portion of the network nodes. However, the resulting management decisions may be sub-optimal due to the controller having a reduced view of the network status [51, 53]. Moreover, the synchronization between controllers and the exchange of control data between them is an added issue to consider.

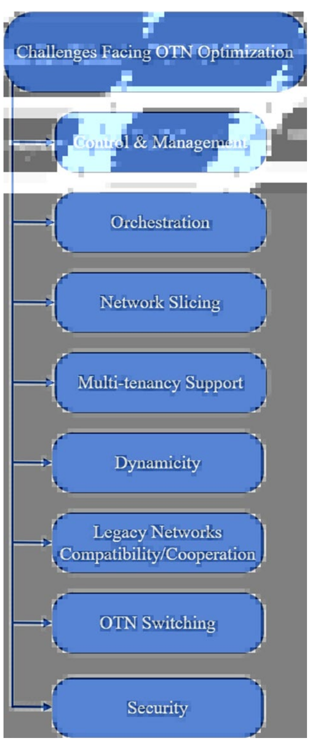


Fig. 9 Challenges facing OTN optimization in 5G networks

4.2 Orchestration

A second challenge facing OTN optimization is the effective and efficient orchestration of the available resources. Orchestration can be defined as "*the selection of resources to satisfy service demands in an optimal way, where the available resources, the service demands and the optimization criteria are all subject to change*." as per the Open Network Foundation [54]. Within the context of NFV, ETSI defines orchestration as "*the coordination of the resources and networks needed to set up cloud-based services and applications*" [55]. Simply speaking, this refers to the question "*how to allocate the network resources?*". In the case of OTN, orchestration decisions refer to the allocation of lightpaths and wavelengths in such a manner that a particular objective is maximized or minimized [56]. For example, the objective can be to minimize the request rejection rate or to maximize the resource utilization to avoid bandwidth fragmentation [50]. The challenge lies in the potentially conflicting objectives and stringent performance constraints/requirements of 5G traffic as well as the highly elastic nature of 5G applications/services. As such, this is a significant challenge that needs to be addressed.

4.3 Network slicing

In addition to the control, management, and orchestration challenges mentioned, a new related challenge exists. More specifically, dealing with the different network slices is a critical factor for OTNs [57, 58]. Network slicing can be defined as the process of creating multiple logical (virtual) and isolated networks on top of a common physical network infrastructure with each logical network/slice being tailored for a particular use case [59]. A network slice represents the "set of run-time network functions, and resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required" [60]. Each slice may be fully or partially, logically or physically, isolated from other network slices and is comprised of a set of physical and logical resources [60]. Additionally, each slice has its own distinct policies and configurations based on the application/service/use case it is created/tailored for [60].

Since network slicing is a key feature of 5G networks, the challenges that this concept poses in terms of network management and resource monitoring/provisioning need to be addressed to guarantee the end-to-end service delivery [57]. This is particularly important when considering that different network slices have different Quality of Experience (QoE) and Quality of Service (QoS) requirements. Moreover, network slice deployments typically operate across several domains and network segments, requiring continuous monitoring and inter-slice isolation as well as effective coordination and precise synchronization to guarantee the QoE/QoS requirements [58]. Thus, the implications of supporting the network slicing concept need to be considered when optimization the OTNs.

4.4 Multi-tenancy support

A related challenge to the network slicing challenge discussed above that OTNs, as an enabling technology for 5G networks, have to address is supporting multi-tenancy [61, 62]. More specifically, the logical separation between different network slices as well as between different virtual network operators within the same network slice is challenging. This is because this requires the efficient and effective monitoring and allocation of the available resources [62]. Moreover, ensuring that the traffic is separated between the different tenants and different slices is a must. Thus, it is important to take into consideration the need for multi-tenancy support when optimizing the OTN network.

4.5 Dynamicity

Another challenge facing OTN optimization is the dynamic and heterogeneous nature of 5G networks and traffic. As shown in many of the previous work on OTN optimization, network architectures were assumed to be static with a fixed number of nodes and links. Moreover, traffic was typically assumed to be static and known a priori [41, 44]. However, 5G networks are expected to be dynamic both in terms of architecture and traffic [63, 64]. This is due to the supporting technologies enabling 5G including SDN and NFV as well as the new services and applications being offered such as Intelligent Transportation Systems (ITSs), Content Delivery Networks (CDNs), e-Health, and industrial automation. The resulting traffic from these new services and applications is expected to be highly dynamic and continuously changing [65]. As such, this dynamicity of network status (in terms of network architecture and traffic) needs to be considered when making OTN optimization decisions.

4.6 Legacy networks cooperation/compatibility

A sixth challenge facing OTNs is being compatible/cooperative with legacy networks [66–69]. This is because the move to 5G is a gradual process. This is further highlighted by the massive amount of 4G infrastructure/equipment currently available/deployed [67]. These equipment and corresponding infrastructure are performing very well and therefore are expected to remain in function for some number of coming years [69]. Hence, an essential guiding principle in 5G networks is being able to co-exist and work with the legacy systems [69]. As such, as part of the continued 5G roll-out and deployment efforts, OTNs need to be compatible and cooperative with legacy networks to ensure that the transition is seamless. This means OTNs need to effectively manage both legacy networks as well as 5G networks including different network tenants and slices.

4.7 OTN switching

Another challenge facing OTNs is the complexity of the OTN switching process, particularly when dealing with rates higher than 100 Gbps [70]. OTN switching refers to the process of facilitating wavelength and sub-wavelength switching within an OTN node through fast optical-electrical-optical (OEO) conversion [4]. As a result, OTN switching helps improve the bandwidth utilization rate and reduces the number of wavelengths needed to transport packets [4]. However, as the required data rates grow higher, this switching process becomes more complex due to the technical limitations of creating digital wrappers [70]. This issue is exacerbated in the case of 5G traffic, especially for eMBB traffic type in which extremely high data rates are required [14, 15]. Consequently, other mechanisms such as Layer 3 traffic grooming become essential to support these higher data rates [70].

To address the complexity problem for OTN switching, several switching architectures have been previously proposed in the literature [4–6]. For example, Eramo et al. proposed a scalable low-complexity switch core. The switch core uses space switching fabric to route at higher rate (high-order optical channel data unit (ODU)) and an OTN time-space switching fabric to route both signals at a lower rate (low-order ODU) and a higher rate (high-order ODU) (lower rate ODUs carried by higher rate ODUs) [4]. The underlying assumption was that the number of OTN switches was known apriori. The authors then extended their work by considering the problem of minimizing the number of required OTN switches [5]. The authors' simulation results showed that if the traffic requiring OTN switching is less than 45% of the total traffic, the proposed integrated WDM/OTN architecture achieved up to 25% of savings. The authors further extended their work by considering the same integrated WDM/OTN switch architecture and studying the impact of switch blocking on the network performance in the case of static and dynamic traffic [6]. The simulation results conducted showed that the switch blocking ratio increases in the case of dynamic traffic scenarios, particularly in networks with large average node degrees [6].

Despite the efforts from previous works, they mostly focused on separating the traffic based on the underlying ODUs carrying them (higher-order and lower-order ODUs). However, they did not propose architectures that address the main technical limitation introduced by the digital wrapping process at extremely high rates (>100 Gbps). Accordingly, this needs to be explored in the case of 5G traffic due to the high data rates expected for different 5G applications/services.

4.8 Security

Last but not least, security is a major concern in OTNs [52, 71]. This is because OTNs are prone to a variety of attacks such as passive eavesdropping attacks, in-band jamming, out-of-band crosstalk attacks, and amplifier attacks [52]. There is also the potential for signal insertion attacks, which may lead to service disruption [71]. This issue is further emphasized with the introduction of new devices that are connected to the network. These new connected devices introduce a new set of potential attack surfaces [72]. As a result, ensuring the security and resiliency of the OTN network needs to be considered as part of any optimization process.

5 Data driven-based opportunities for innovation

As next-generation networks take shape, the QoS requirements of new and emerging use cases get increasingly stringent. Consequently, Network Service Providers (NSPs) are tasked with adhering to these new requirements while simultaneously ensuring their internal business objectives are met. These conflicting objectives, coupled with the increasing complexity of future networks, create a unique opportunity for innovation in network Management and Orchestration (MANO). Specifically, the transition from conventional networking to the paradigm-shifting data-driven intelligent networking is a promising solution to address many of the challenges faced by NSPs. To support this transition, Machine Learning (ML) has been identified as a candidate technology due to its ability to learn patterns and policies from network-generated data and reduce the run-time complexity of traditional solutions [73, 74]. The following section will outline some of the key research areas and opportunities for innovation stemming from this emerging paradigm, particularly in the context of OTN in 5G networks. These opportunities are summarized in Fig. 10.

5.1 Traffic prediction

One of the main advantages of ML implementations in networks is the ability to forecast network traffic. This ability is especially advantageous as it allows NSPs to plan resource allocation and service provisioning in advance and take appropriate measures to ensure end-user needs are met. Moreover, traffic prediction can also assist NSPs "anticipate capacity exhaustion and degradation, or to predict and localize failures" [75]. Additionally, given the new and emerging types of network traffic generated by next-generation use cases, predicting traffic can help optimize the future aggregation of network traffic to maximize network performance and reduce operational costs. This should be done while considering the impact of features such as network slicing and multi-tenancy support including:

- Traffic separation between network slices and multi-tenant networks
- Tenant authorization levels
- Service level agreement requirements
- Data isolation and authenticity

Regarding traffic forecasting, it is critical to examine and understand the characteristics of both legacy networks and 5G network traffic as well as consider the constantly changing end-user behaviour. Once the domain knowledge is built and understood, accurate traffic prediction models can be

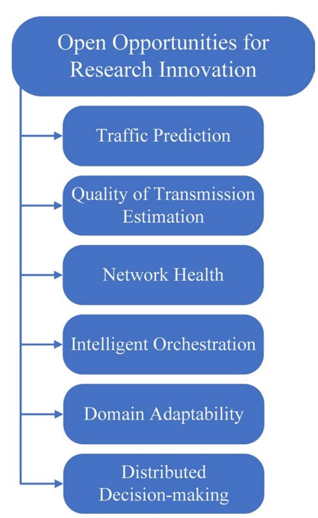


Fig. 10 Research opportunities for innovation

built to assist in network MANO tasks. The benefit of traffic prediction in optical-based networks has been illustrated in [76] in which the author proved through simulations that traffic prediction achieved closed to optimal performance by reducing the request blocking probability to optimal levels. Similarly, the authors in [77] showed that accurate traffic prediction improved the resource allocation process and consequently, reduced the operational expenditure (OPEX) of the network.

The traffic prediction process can be performed at multiple levels within the network. For example, as shown in Fig. 11, the ML-enabled traffic prediction module can be deployed at different locations within the network, namely at the network edge level or at the core level. More specifically, when deployed at the network edge, 5G gNBs can collect the data from the users/devices generating requests for different applications and services belonging to different slices. Using this data, the gNB can predict the incoming traffic resulting from the user requests using various supervised ML techniques. Based on this predicted traffic, the gNB can better allocate the wireless resources to the users. Additionally, it can share the aggregate of these predictions with the core network to help in the allocation of the optical resources at the OTN layer. On the other hand, another potential architecture is to deploy the ML-enabled traffic prediction module at the core. This is done to reduce the computational requirement at the gNBs and offload it to the core. In this case, the module would collect the aggregate traffic from multiple gNBs and locations to train the ML model. Then, using this model, the OTN layer can better plan how to allocate its resources based the predicted incoming traffic volume.

The advantage of the edge level traffic prediction is that it allows for the prediction to be done at a finer granularity and provides better planning and allocation of the wireless resources. However, this prediction may not allow for optimal planning at the OTN level since the gNBs are often unaware of the traffic patterns in adjacent and other locations. In contrast, the advantage of the core level traffic prediction is that it can provide more accurate prediction of the traffic that will is expected to be handled by the OTN. This is because it has a more global view of the network. More specifically, it is better able to estimate the spatial dimension characteristics of the traffic while also analyzing and accounting for the time dimension characteristics. Accordingly, it can better plan the allocation decisions of the optical resources for the whole network. However, this comes at the expense of added computational complexity due to the larger size of data to deal with.

5.2 Quality of transmission estimation

To ensure the efficient and effective management of the network, in addition to the aforementioned incoming traffic, a second perspective to consider is the channel characteristics and the corresponding quality of transmission. This is because having a better understanding and insights into the expected channel conditions can help with improved network resource allocation and management based on both the traffic as well as the channel conditions. To that end, ML algorithms and paradigms can play a pivotal role in accurately estimating the quality of transmission for better resource allocation and management.

Despite some earlier efforts to incorporate ML algorithms for quality of transmission estimation [78–81], further work can be done. This is because many of the previous works relied on simple or traditional classifiers. Moreover, previous

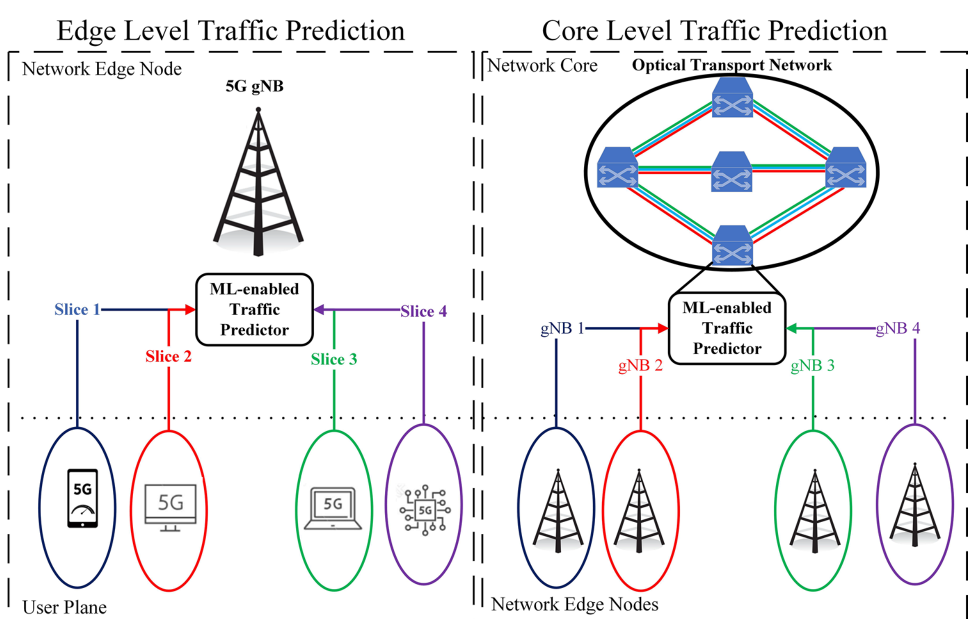


Fig. 11 Potential ML-enabled traffic prediction architectures

works often used synthetic datasets to train the models. This can lead to biased models that may not be indicative of the true quality of transmission observed/experienced in real-world networks. Therefore, to have more effective and efficient network management decisions, more accurate models need to be considered. To that end, more complex ML algorithms can be investigated. For example, deep neural network learning models such as convolutional neural networks can be developed to explore their estimation capabilities with regard to optical transmission quality. In a similar fashion, algorithms such as recurrent neural networks and long-short term memory (LSTM) can also be investigated as they are able to account for the time-varying nature of the optical transmission quality. Hence, they have great potential in providing accurate optical transmission quality estimation over the time dimension. Again, in a similar vein to the traffic prediction case, these models can be deployed at both the edge or network core depending on the transmission quality estimation granularity desired.

5.3 Network health

One of the main properties distinguishing 5G and beyond networks from previous network generations is the notion of self-healing networks. This property entails the sensing of network conditions and the mitigation of any faults or failures that emerge as well as any potentially contradicting configuration rules (due to multi-tenancy). ML will be at the forefront of this effort as its ability to make quick and effective decisions will allow NSPs to ensure the continual performance of their networks while reducing the impact of outages on the end-users. Recently, Spark, in association with Ciena, has started the deployment of a self-healing OTN, which enables the automation of light signal path changing after a fault [82]. The ability to provide self-healing capabilities to networks will ultimately improve network health and result in more resilient and adaptable networks.

Additionally, ML has a significant role to play in providing added security to next-generation networks, particularly for transport networks [83]. ML has already shown great potential as an effective and efficient approach for different network security applications including network anomaly detection and intrusion detection [84–91]. Thus, similar approaches can be adopted at the transport network layer to further improve the security and resiliency of the OTN, particularly with the added potential attack surfaces due to 5G and beyond networks. ML models can be developed to detect the various OTN attacks and alert NSPs to take actions to mitigate the impact of these attacks. Moreover, ML models can also be used to detect and isolate contaminated insider traffic. In turn, this again will help improve the network health as it becomes more secure and resilient to failures as well as attacks.

5.4 Intelligent orchestration

While many implementations for ML-assisted intelligent network MANO have been explored in recent studies, not all ML is made equal. Ideally, for a highly dynamic network that is constantly prone to changes, an ML technique capable of realizing and adapting to these types of changes is necessary. This includes being able to account for the various QoS and prioritization levels in addition to the tunneling protocols that accurately reflect these parameters (due to network slicing and multi-tenancy). This can be done by the continuous resource slicing and monitoring as well as logging and reporting various networks service metrics. To solve this problem, the use of advanced intelligence techniques, such as federated and reinforcement learning, has been proposed [92, 93]. Reinforcement learning has been identified as a candidate specifically for its ability to learn policies through experience. This type of learning can be used to learn optimal network MANO decisions and execute them with a much lower run-time complexity compared to traditional methods. Federated learning, on the other hand, has been praised for its ability to leverage intelligence from highly distributed systems and provide a decentralized and privacy-preserving method of developing a global system model. As a result, the network management and orchestration process can be automated and delegated accordingly.

As an illustrative example, Fig. 12 shows how reinforcement learning can be deployed and utilized as effective learning mechanism of the environment. In this case, the learning agent can learn the optimal policy for OTN optimization based on the data collected and experience gained from the OTN network. More specifically, the agent would collect the data corresponding to the optical resource allocation decisions and the associated reward as well as the status of the OTN network. Using this data, the learning agent can

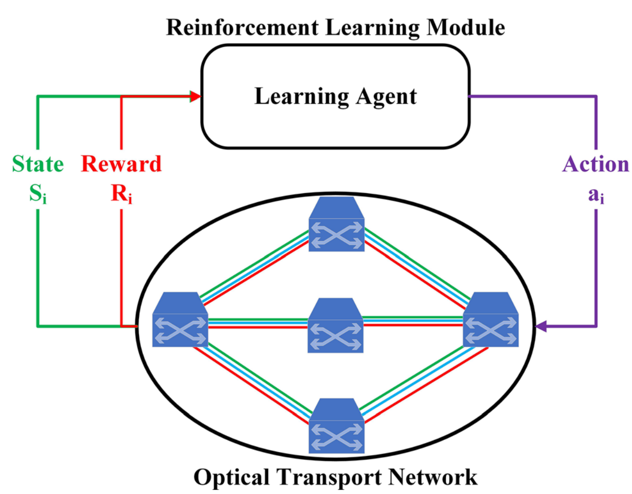


Fig. 12 Potential RL-enabled network orchestration architecture

determine the optimal network MANO policy and implement it for future decisions.

This builds on the two earlier opportunities, namely "traffic prediction" and "Quality of Transmission Estimation". More specifically, RL can be deployed/implemented as illustrated in Fig. 12 to perform both traffic prediction and quality of transmission estimation tasks. Combining these two tasks and having an accurate estimate of both will result in a more efficient, effective, and intelligent orchestration as it will have more information about the current network state. Consequently, it can make better informed decisions and allocate the available resources in an optimized manner to maximize the utility by maximizing the network resource utilization and throughput while minimizing the network cost. It is worth noting that such an RL-enabled system/framework can be deployed/implemented regardless of the OTN supporting technology used. This means whether the OTN uses a traditional ODU-based architecture or a more advanced OSU-based one, the RL module can perform both the traffic prediction and quality of transmission estimation tasks to take efficient orchestration decisions by allocating the appropriate ODUs or OSUs to the client demands. Hence, the described RL system/framework is agnostic to the OTN supporting technology used.

5.5 Domain adaptability

As with any ML application, a changing domain has the potential to impact the performance of a model. This is especially pertinent to the next-generation networks and use cases as they are highly dynamic systems. When implementing intelligence in such a system, extreme caution must be taken to ensure the detection and mitigation of performance degradation caused by a changing domain (known as model drift) that can impact the system's stability. To this end, ML implementations should include model drift detection mechanisms and provide a course of action to remediate a detected drift. This is especially pertinent in the case of network security as new threats emerge, the landscape of known threats changes effectively, causing many models to drift. Novel research into the ability of a system to detect and mitigate model drift, especially in security applications [94], while ensuring constant performance, is a clear indication of the future direction of this field and its applicability to network systems.

5.6 Distributed decision-making

The final opportunity is based on the concept of distributed decision-making. As the network size and complexity increases, the ability to perform rapid intelligent decisions to improve the network's performance is required. To this end, leveraging insights from multiple regions in a communication-efficient way is critical. The aforementioned technique of federated learning is a critical step towards achieving this as it efficiently enables the cooperation of multiple intelligent agents without the need for sharing or transferring large amounts of data. As such, by placing multiple agents throughout the network, they can both learn the individual properties of their region while also leveraging intelligence and insights from other regions. Developing and deploying intelligent agents throughout the network and leveraging cooperative insights will enable increased levels of automation and improve performance and the efficiency of network MANO activities.

Figure 13 provides a potential FL-enabled OTN architecture. In this case, the local models and parameters trained at each location about the status of the network and is shared with the FL global learning agent. Using these models and parameters, the agent develops a global model through the aggregation of the local models and parameters. This allows each gNB to leverage the insights gained both by itself and that of other gNBs. In turn, this can help them make better local network management and orchestration decisions.

6 Conclusion

5G networks have become a reality with added development and deployment efforts due to the continued growth of both mobile broadband and fixed broadband subscriptions as well as the added deployment of Internet of Things (IoT) devices. 5G networks are expected to support a diverse set of new applications/services, in addition to existing applications/services from previous generations (2G/3G/4G). The

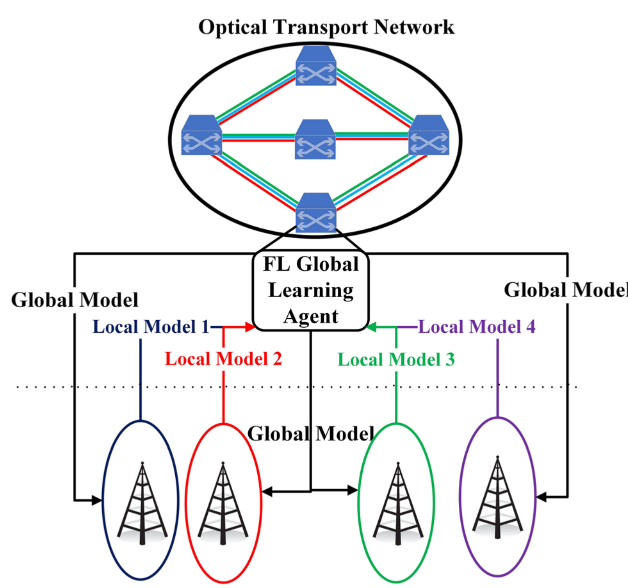


Fig. 13 Potential FL-enabled OTN architecture

COVID-19 pandemic has further increased the demand for such services which has resulted in a further surge in the Internet usage. Therefore, 5G networks are expected to have a highly flexible architecture at all levels (radio, core, and transport levels). Moreover, this requires a high automation level in the deployment and maintenance of networks, parts of a network, or single resources (e.g. network slices).

Optical Transport Networks (OTNs) have been proposed as one potential and promising supporting technology for 5G networks at the transport level, particularly for next generation transport networks featuring large-granule broadband service transmissions. However, this introduces a fresh set of challenges. One such challenge is the added complexity, particularly in terms of managing these networks. The control, management, and orchestration of such networks continues to evolve for fast provisioning of light-paths, fast restoration and high availability. Another challenge is the security of OTNs against the different passive and active optical attacks. Additionally, new security threats have emerged due to the introduction of new services and applications resulting in additional potential attack surfaces. Thus, OTN security is a prime concern to consider.

To that end, this work provided a brief overview of 5G networks and their requirements. It also summarized the advantage of OTNs and surveyed some of the previous work proposed for optimizing such networks. Additionally, this work presented the challenges facing OTNs and their optimization within the context of 5G. Moreover, it outlined some of the key research areas and opportunities for innovation stemming from the data-driven intelligent networking paradigm using ML techniques. This includes opportunities such as the use of ML to predict OTN network traffic, improve network health, offer intelligent orchestration, and provide distributed decision-making processes.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. International Telecommunication Union, Measuring digital development: facts and figures 2020. Technical report, International Telecommunication Union, Geneva, Switzerland (2020). https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx
2. International Telecommunication Union: Setting the Scene for 5G: Opportunities & Challenges. Technical report, International Telecommunication Union, Geneva, Switzerland (2018)
3. Covell, B., Liebhart, R.: Drivers and Motivation for 5G, 5G for the Connected World, pp. 1–34 (2019)
4. Eramo, V., Listanti, M., Sabella, R., Testa, F.: Definition and performance evaluation of a low-cost/high-capacity scalable integrated OTN/WDM switch. J. Opt. Commun. Netw. **4**(12), 1033 (2012). https://doi.org/10.1364/JOCN.4.001033
5. Eramo, V., Listanti, M., Sabella, R., Testa, F.: Integrated OTN/WDM switching architecture equipped with the minimum number of OTN switches. J. Opt. Commun. Netw. **6**(2), 138 (2014). https://doi.org/10.1364/JOCN.6.000138
6. Eramo, V., Listanti, M., Lavacca, F.G., Testa, F., Sabella, R.: Performance evaluation of integrated OTN/WDM metropolitan networks in static and dynamic traffic scenarios. J. Opt. Commun. Netw. **7**(8), 761 (2015). https://doi.org/10.1364/JOCN.7.000761
7. Eramo, V., Listanti, M., Lavacca, F., Iovanna, P.: Dimensioning of OTN/WDM rings for the transport of ethernet/CPRI flows in 5g scenario. In: 2018 20th International Conference on Transparent Optical Networks (ICTON) (2018), pp. 1–6. https://doi.org/10.1109/ICTON.2018.8473863
8. Katib, I., Medhi, D.: A network optimization model for multi-layer IP/MPLS over OTN/DWDM networks. In: International Workshop on IP Operations and Management, pp. 180–185. Springer (2009)
9. Schmidt, A.: Integrated OTN switching virtualizes optical networks. Technical Report, Infonetics Research, California, United States (2012). https://www.yumpu.com/en/document/read/49760130/integrated-otn-switching-virtualizes-optical-networks-infinera
10. RS Tech, What is OTN, Is it Similar with ONT? Technical Report, Router-Switch, California, United States (2020). https://blog.router-switch.com/2020/07/what-is-otn-is-it-similar-with-ont/
11. Vaez-Ghaemi, R.: Next-generation optical transport networks. Technical Report, Viavi Solutions, California, United States (2015). https://www.viavisolutions.com/en-us/literature/next-generation-optical-transport-networks-white-papers-books-en.pdf
12. Cavalcante, J., Patel, A., Celestino, J.: Performance management of optical transport networks through time series forecasting, In: 2017 IEEE 31st international conference on advanced information networking and applications (AINA), pp. 152–159 (2017). https://doi.org/10.1109/AINA.2017.117
13. Alcatel-Lucent, Multi-layer network optimization: a pragmatic approach for delivering better quality at lower cost. Technical Report, Alcatel-Lucent, Paris, France (2013). https://www.tmcnet.com/tmc/whitepapers/documents/whitepapers/2014/10741-multi-layer-network-optimization.pdf
14. Erel-Özçevik, M., Canberk, B.: Road to 5G reduced-latency: a software defined handover model for EMBB services. IEEE Trans. Veh. Technol. **68**(8), 8133 (2019). https://doi.org/10.1109/TVT.2019.2925393
15. Ericsson, Ericsson mobility report. Technical Report, Ericsson, Stockholm, Sweden (2020). https://www.ericsson.com/4adc87/assets/local/mobility-report/documents/2020/november-2020-ericsson-mobility-report.pdf
16. Dawy, Z., Saad, W., Ghosh, A., Andrews, J.G., Yaacoub, E.: Toward massive machine type cellular communications. IEEE Wirel. Commun. **24**(1), 120 (2017). https://doi.org/10.1109/MWC.2016.1500284WC
17. Guo, J., Durrani, S., Zhou, X., Yanikomeroglu, H.: Massive machine type communication with data aggregation and resource scheduling. IEEE Trans. Commun. **65**(9), 4012 (2017). https://doi.org/10.1109/TCOMM.2017.2710185
18. Pandey, S.R., Alsenwi, M., Tun, Y.K., Hong, C.S.: A downlink resource scheduling strategy for urllc traffic. In: 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 1–6 (2019). https://doi.org/10.1109/BIGCOMP.2019.8679266
19. Li, Z., Uusitalo, M.A., Shariatmadari, H., Singh, B.: 5G URLLC: design challenges and system concepts. In: 2018 15th international

symposium on wireless communication systems (ISWCS), pp. 1–6 (2018). https://doi.org/10.1109/ISWCS.2018.8491078
20. Stein, Y., Gittik, Y., Insler, R.: Everything You Need To Know About 5G xHaul. Technical Report, RAD, New Jersey, USA (2020). https://www.rad.com/resources/white-papers/everything-you-need-know-about-5g-xhaul
21. 3GPP, TR 21.915. Technical Report, 3GPP (2018). https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3389
22. Liu, G., Huang, Y., Chen, Z., Liu, L., Wang, Q., Li, N.: 5G deployment: standalone vs. non-standalone from the operator perspective. IEEE Commun. Mag. **58**(11), 83 (2020). https://doi.org/10.1109/MCOM.001.2000230
23. Olli Andersson, Benetel, Functional splits: the foundation of an open 5G RAN. Technical Report, 5G Technology World (2021). https://www.5gtechnologyworld.com/functional-splits-the-foundation-of-an-open-5g-ran/
24. Mazzara, M., Bucchiarone, A., Dragoni, N., Rivera, V.: Size matters: microservices research and applications, In: Microservices, pp. 29–42, Springer (2020),
25. Brown, G.: Service-based architecture for 5G core networks. Technical Report, Huawei Technologies Co. Ltd. (2017)
26. International Telecommunication Union, ITU-T G.709: Optical Transport Networks. Technical Report, International Telecommunication Union, Geneva, Switzerland (2010). https://www.itu.int/rec/T-REC-G.709/en
27. Gorshe, S.: A tutorial on ITU-T G.709 optical transport networks (OTN). Technical Report, PMC-Sierra (2010). https://www.all-electronics.de/wp-content/uploads/migrated/article-pdf/85811/417-2081250-otn-tutorial-053751.pdf
28. Smith, D.: Odu/oduk—optical data unit (2016)
29. Joseph, V., Mulugu, S.: 6-emerging trends in packet optical convergence. In: Network Convergence, Joseph, V., Mulugu S. (eds.), pp. 525–552. Morgan Kaufmann, Boston (2014). https://doi.org/10.1016/B978-0-12-397877-6.00006-0. https://www.sciencedirect.com/science/article/pii/B9780123978776000060
30. Bai, L.: Optical service unit (osu)-based next generation optical transport network (ng otn) technology and verification. In: MATEC Web of Conferences, vol. 336, p. 04014. EDP Sciences (2021)
31. Li, X., Liu, Y., Zhao, Y., Li, Z., Li, Y., Zhang, J.: All optical service network for f5g. In: 2021 IEEE 21st International Conference on Communication Technology (ICCT), pp. 710–714 (2021). https://doi.org/10.1109/ICCT52962.2021.9657914
32. Yang, H., Wang, Y., Chi, H., Yuan, D.: Transmission delay in optical fiber communication system of power transmission and transformation. In: International Conference on Applications and Techniques in Cyber Security and Intelligence, pp. 50–57. Springer (2021)
33. Liu, X.: Enabling optical network technologies for 5G and beyond. J. Lightwave Technol. **40**(2), 358 (2022). https://doi.org/10.1109/JLT.2021.3099726
34. Puttnam, B.J., Rademacher, G., Luís, R.S.: Space-division multiplexing for optical fiber communications. Optica **8**(9), 1186 (2021). https://doi.org/10.1364/OPTICA.427631
35. Essiambre, R.J., Kramer, G., Winzer, P.J., Foschini, G.J., Goebel, B.: Capacity limits of optical fiber networks. J. Lightwave Technol. **28**(4), 662 (2010). https://doi.org/10.1109/JLT.2009.2039464
36. Soma, D., Wakayama, Y., Beppu, S., Sumita, S., Tsuritani, T., Hayashi, T., Nagashima, T., Suzuki, M., Yoshida, M., Kasai, K., Nakazawa, M., Takahashi, H., Igarashi, K., Morita, I., Suzuki, M.: 1016-Peta-B/s dense SDM/WDM transmission over 6-mode 19-core fiber across the c+l band. J. Lightwave Technol. **36**(6), 1362 (2018). https://doi.org/10.1109/JLT.2018.2799380
37. Moghaddam, E.E., Beyranvand, H., Salehi, J.A.: Resource allocation in space division multiplexed elastic optical networks secured with quantum key distribution. IEEE J. Sel. Areas Commun. **39**(9), 2688 (2021). https://doi.org/10.1109/JSAC.2021.3064641
38. Zhu, R., Samuel, A., Wang, P., Li, S., Oun, B.K., Li, L., Lv, P., Xu, M., Yu, S.: Protected resource allocation in space division multiplexing-elastic optical networks with fluctuating traffic. J. Netw. Comput. Appl. **174**, 102887 (2021). https://doi.org/10.1016/j.jnca.2020.102887
39. Katib, I., Medhi, D.: IP/MPLS-over-OTN-over-DWDM multilayer networks: an integrated three-layer capacity optimization model, a heuristic, and a study. IEEE Trans. Netw. Serv. Manag. **9**(3), 240 (2012). https://doi.org/10.1109/TNSM.2012.12.110124
40. Katib, I., Medhi, D.: Network protection design models, a heuristic, and a study for concurrent single-link per layer failures in three-layer networks. Comput Commun **36**(6), 678 (2013). https://doi.org/10.1016/j.comcom.2012.09.008
41. Govardan, C., Sri Krishna Chaitanya, K., Naik, B. Krishna Kumar, Rao, D.S. Shreesha, Jagadeesh, C., Gowrishankar, R., Sai, S. Siva Sankara., Beherey, P., Kishorez, B.S.: A heuristic algorithm for network optimization of OTN over DWDM network. In: 2015 IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS), pp. 1–6 (2015). https://doi.org/10.1109/ANTS.2015.7413606
42. Zefreh, M.S., Tizghadam, A., Leon-Garcia, A., Elbiaze, H., Miron, W.: Capex optimization in multi-chassis multi-rate IP/optical networks. In: 2016 International Conference on Computing, Networking and Communications (ICNC), pp. 1–6 (2016). https://doi.org/10.1109/ICCNC.2016.7440611
43. Papanikolaou, P., Christodoulopoulos, K., Varvarigos, E.: Optimization techniques for incremental planning of multilayer elastic optical networks. IEEE/OSA J. Opt. Commun. Netw. **10**(3), 183 (2018). https://doi.org/10.1364/JOCN.10.000183
44. Xing, C., Addie, R.G., Peng, Y., Lin, R., Li, F., Hu, W., Abramov, V.M., Zukerman, M.: Resource provisioning for a multi-layered network. IEEE Access **7**, 16226 (2019). https://doi.org/10.1109/ACCESS.2019.2894396
45. Moniz, D., Pedro, J., Pires, J.: Multi-layer network optimization efficiently exploiting real-time performance monitoring. In: 2019 IEEE Global Communications Conference (GLOBECOM), pp. 1–6 (2019). https://doi.org/10.1109/GLOBECOM38437.2019.9013474
46. da Silva, J.C., Bastos-Filho, C.J.A., Almeida-Júnior, R.C., Chaves, D.A.R., Araújo, D.R.B., Xavier, A.V.S., Martins-Filho, J.F.: OTN network planning over DWDM using computational intelligence. In: 2019 21st International Conference on Transparent Optical Networks (ICTON), pp. 1–4 (2019). https://doi.org/10.1109/ICTON.2019.8840191
47. Xavier, A.V.S., da Silva, J.C., Martins-Filho, J.F., Bastos-Filho, C.J.A., Chaves, D.A.R., Almeida-Júnior, R.C., Araujo, D.R.B., Martins, J.A.: Heuristic planning algorithm for sharing restoration interfaces in OTN over DWDM networks. Opt. Fiber Technol. **61**, 102426 (2021). https://doi.org/10.1016/j.yofte.2020.102426
48. de Oliveira, B.Q., de Sousa, M.A., Vieira, F.H.T.: Firefly algorithm based optimization model for planning of optical transport networks. Adv. Electr. Comput. Eng. **20**(2), 55 (2020)
49. de Oliveira, B.Q., de Sousa, M.A., Vieira, F.H.T.: A hybrid firefly-genetic algorithm for planning of optical transport networks. J. Commun. Inf. Syst. **35**(1), 243 (2020)
50. Oki, E., Chatterjee, B.C.: Design and control in elastic optical networks: issues, challenges, and research directions. In 2017 International Conference on Computing, Networking and Communications (ICNC), pp. 546–549 (2017). https://doi.org/10.1109/ICCNC.2017.7876187
51. Casellas, R., Martínez, R., Vilalta, R., Muñoz, R.: Control, management, and orchestration of optical networks: Evolution, trends, and challenges. J. Lightwave Technol. **36**(7), 1390 (2018). https://doi.org/10.1109/JLT.2018.2793464

52. Dahan, D., Mahlab, U.: Security threats and protection procedures for optical networks. IET Optoelectron. **11**(5), 186 (2017)
53. King, D., Farrel, A., King, E.N., Casellas, R., Velasco, L., Nejabati, R., Lord, A.: The dichotomy of distributed and centralized control: Metro-haul, when control planes collide for 5g networks. Opt. Switch. Netw. **33**, 49 (2019)
54. Open Network Foundation, SDN Architecture, Issue 1.1 - ONF TR-521. Technical Report, Open Network Foundation, California, United States (2016)
55. European Telecommunications Standards Institute, Network functions virtualisation (NFV); Architectural framework. Technical Report, European Telecommunications Standards Institute, Sophia Antipolis, France (2014)
56. Castoldi, P., Fichera, S., Gharbaoui, M., Giorgetti, A., Martini, B., Paolucci, F.: Disaggregated optical network control and orchestration of heterogeneous domains. In: 2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC), pp. 1–3 (2019). https://doi.org/10.23919/PS.2019.8817856
57. Montero, R., Pagès, A., Agraz, F., Spadaro, S.: Supporting QoE/QoS-aware end-to-end network slicing in future 5g-enabled optical networks. In: Metro and Data Center Optical Networks and Short-Reach Links II, vol. 10946,, p. 109460F. International Society for Optics and Photonics (2019)
58. Huang, S., Guo, B., Liu, Y.: 5G-oriented optical underlay network slicing technology and challenges. IEEE Commun. Mag. **58**(2), 13 (2020). https://doi.org/10.1109/MCOM.001.1900583
59. Ordonez-Lucena, J., Ameigeiras, P., Lopez, D., Ramos-Munoz, J.J., Lorca, J., Folgueira, J.: Network slicing for 5g with SDN/NFV: Concepts, architectures, and challenges. IEEE Commun. Mag. **55**(5), 80 (2017). https://doi.org/10.1109/MCOM.2017.1600935
60. Alliance, N.G.M.N.: Description of network slicing concept. Technical Report, Next Generation Mobile Networks Alliance, Frankfurt, Germany (2016)
61. Li, X., Casellas, R., Landi, G., de la Oliva, A., Costa-Perez, X., Garcia-Saavedra, A., Deiss, T., Cominardi, L., Vilalta, R.: 5G-crosshaul network slicing: Enabling multi-tenancy in mobile transport networks. IEEE Commun. Mag. **55**(8), 128 (2017). https://doi.org/10.1109/MCOM.2017.1600921
62. Oladejo, S.O., Falowo, O.E.: 5G network slicing: a multi-tenancy scenario. In: 2017 Global Wireless Summit (GWS) (2017), pp. 88–92. https://doi.org/10.1109/GWS.2017.8300476
63. Kapassa, E., Touloupou, M., Stavrianos, P., Kyriazis, D.: Dynamic 5G slices for IoT applications with diverse requirements. In: 2018 Fifth International Conference on Internet of Things: Systems, Management and Security, pp. 195–199 (2018). https://doi.org/10.1109/IoTSMS.2018.8554386
64. Raza, M.R., Fiorani, M., Rostami, A., Öhlen, P., Wosinska, L., Monti, P.: Dynamic slicing approach for multi-tenant 5G transport networks. J. Opt. Commun. Netw. **10**(1), A77 (2018)
65. Golkarifard, M., Chiasserini, C.F., Malandrino, F., Movaghar, A.: Dynamic VNF placement, resource allocation and traffic routing in 5G. Comput. Netw. **188**, 107830 (2021)
66. González, S., de la 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. Trans. Emerg. Telecommun. Technol. **27**(9), 1196 (2016)
67. Agiwal, M., Roy, A., Saxena, N.: Next generation 5G wireless networks: a comprehensive survey. IEEE Commun. Surveys Tutor. **18**(3), 1617 (2016). https://doi.org/10.1109/COMST.2016.2532458
68. Costa-Perez, X., Garcia-Saavedra, A., Li, X., Deiss, T., de la Oliva, A., di Giglio, A., Iovanna, P., Moored, A.: 5G-crosshaul: an SDN/NFV integrated fronthaul/backhaul transport network architecture. IEEE Wirel. Commun. **24**(1), 38 (2017). https://doi.org/10.1109/MWC.2017.1600181WC
69. Zaidi, Z., Friderikos, V., Yousaf, Z., Fletcher, S., Dohler, M., Aghvami, H.: Will SDN be part of 5G? IEEE Commun. Surveys Tutor. **20**(4), 3220 (2018). https://doi.org/10.1109/COMST.2018.2836315
70. Janson, C.: Why OTN keeps trucking on. Available at: https://www.nokia.com/blog/why-otn-keeps-trucking-on/ (2021). Accessed 14 Nov 2022
71. Furdek, M., Natalino, C., Schiano, M., Di Giglio, A.: Experiment-based detection of service disruption attacks in optical networks using data analytics and unsupervised learning. In: Metro and data center optical networks and short-reach links II, vol. 10946, p. 109460D, International Society for Optics and Photonics (2019)
72. Yang, L., Moubayed, A., Shami, A.: Mth-ids: a multi-tiered hybrid intrusion detection system for internet of vehicles, IEEE Internet Things J. pp. 1–1 (2021). https://doi.org/10.1109/JIOT.2021.3084796
73. Buda, T.S., Assem, H., Xu, L., Raz, D., Margolin, U., Rosensweig, E., Lopez, D.R., Corici, M.I., Smirnov, M., Mullins, R. et al.: Can machine learning aid in delivering new use cases and scenarios in 5G?. In: NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium, pp. 1279–1284. IEEE (2016)
74. ETSI, White paper: Zero-touch network and service management. Technical Report, ETSI (2017). https://portal.etsi.org/TBSiteMap/ZSM/OperatorWhitePaper
75. Velasco, L., Piat, A.C., Gonzlez, O., Lord, A., Napoli, A., Layec, P., Rafique, D., D'Errico, A., King, D., Ruiz, M., Cugini, F., Casellas, R.: Monitoring and data analytics for optical networking: Benefits, architectures, and use cases. IEEE Network **33**(6), 100 (2019). https://doi.org/10.1109/MNET.2019.1800341
76. Aibin, M.: Traffic prediction based on machine learning for elastic optical networks. Opt. Switch. Netw. **30**, 33 (2018)
77. Vinchoff, C., Chung, N., Gordon, T., Lyford, L., Aibin, M.: Traffic prediction in optical networks using graph convolutional generative adversarial networks. In: 2020 22nd International Conference on Transparent Optical Networks (ICTON), pp. 1–4 (2020). https://doi.org/10.1109/ICTON51198.2020.9203477
78. Tremblay, C., Aladin, S.: Machine learning techniques for estimating the quality of transmission of lightpaths. In: 2018 IEEE Photonics Society Summer Topical Meeting Series (SUM), pp. 237–238 (2018). https://doi.org/10.1109/PHOSST.2018.8456791
79. Morais, R.M., Pedro, J.: Machine learning models for estimating quality of transmission in DWDM networks. J. Opt. Commun. Netw. **10**(10), D84 (2018). https://doi.org/10.1364/JOCN.10.000D84
80. Morais, R.M., Pedro, J.: Evaluating machine learning models for qot estimation. In: 2018 20th International Conference on Transparent Optical Networks (ICTON), pp. 1–4 (2018). https://doi.org/10.1109/ICTON.2018.8473941
81. Rottondi, C., di Marino, R., Nava, M., Giusti, A., Bianco, A.: On the benefits of domain adaptation techniques for quality of transmission estimation in optical networks. J. Opt. Commun. Netw. **13**(1), A34 (2021). https://doi.org/10.1364/JOCN.401915
82. Ciena, Spark starts roll out of new faster, 'self healing' network technology. Technical Report, Ciena (2020). https://www.ciena.com/about/newsroom/press-releases/Spark-starts-roll-out-of-new-faster-self-healing-network-technology.html
83. Furdek, M., Natalino, C., Di Giglio, A., Schiano, M.: Optical network security management: requirements, architecture, and efficient machine learning models for detection of evolving threats. J. Opt. Commun. Netw. **13**(2), A144 (2021)
84. Moubayed, A., Injadat, M., Shami, A., Lutfiyya, H.: DNS typosquatting domain detection: a data analytics & machine learning based approach. In: 2018 IEEE Global Communications

Conference (GLOBECOM), pp. 1–7 (2018). https://doi.org/10.1109/GLOCOM.2018.8647679
85. Moubayed, A., Aqeeli, E., Shami, A.: Ensemble-based feature selection and classification model for DNS typo-squatting detection. In: 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–6 (2020). https://doi.org/10.1109/CCECE47787.2020.9255697
86. Moubayed, A., Aqeeli, E., Shami, A.: Detecting DNS typo-squatting using ensemble-based feature selection classification models Détection du typosquattage DNS à l'aide de modèles de sélection et de classification basés sur un ensemble de caractéristiques, IEEE Can. J. Electr. Comput. Eng. 1–11 (2021). https://doi.org/10.1109/ICJECE.2021.3072008
87. Injadat, M., Moubayed, A., Shami, A.: Detecting botnet attacks in IoT environments: an optimized machine learning approach. In: IEEE 32nd International Conference on Microelectronics (2020)
88. Moubayed, A., Injadat, M., Shami, A.: Optimized Random Forest Model for Botnet Detection Based on DNS Queries, in IEEE 32nd International Conference on Microelectronics (2020)
89. Injadat, M., Moubayed, A., Nassif, A.B., Shami, A.: Machine learning towards intelligent systems: applications, challenges, and opportunities, Artificial Intelligence Review pp. 1–50 (2021)
90. Injadat, M., Salo, F., Nassif, A.B., Essex, A., Shami, A.: Bayesian optimization with machine learning algorithms towards anomaly detection. In: 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1–6 (2018). https://doi.org/10.1109/GLOCOM.2018.8647714
91. Injadat, M., Moubayed, A., Nassif, A.B., Shami, A.: Multi-stage optimized machine learning framework for network intrusion detection, IEEE Transactions on Network and Service Management pp. 1–14 (2020)
92. Manias, D.M., Shami, A.: The need for advanced intelligence in NFV management and orchestration, IEEE Network (2020)
93. Manias, D.M., Shami, A.: Making a case for federated learning in the internet of vehicles and intelligent transportation systems. IEEE Netw. **35**(3), 88 (2021)
94. Yang, L., Shami, A.: A lightweight concept drift detection and adaptation framework for IoT data streams, IEEE Internet of Things Magazine (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abdallah Moubayed received his Ph.D. in Electrical & Computer Engineering from the University of Western Ontario in August 2018, his M.Sc. degree in Electrical Engineering from King Abdullah University of Science and Technology, Thuwal, Saudi Arabia in 2014, and his B.E. degree in Electrical Engineering from the Lebanese American University, Beirut, Lebanon in 2012. He was a Postdoctoral Associate in the Optimized Computing and Communications (OC2) lab at University of Western Ontario between 2018 and 2022. Currently, he is an Assistant Professor in the Electrical & Computer Engineering Department at the Lebanese American University. His research interests include wireless communication, resource allocation, wireless network virtualization, performance & optimization modeling, machine learning & data analytics, computer network security, cloud computing, and e-learning.

Dimitrios Michael Manias is a Ph.D. student in the Optimized Computing and Communications (OC2) Lab at Western University, Ontario, Canada. He received his B.ESc. degree in Electrical and Computer Engineering and his M.ESc. degree in Electrical and Computer Engineering with a Vector Institute Accredited Graduate Collaborative Specialization in Artificial Intelligence from Western University. His current research interests include machine learning, advanced analytics, computer networks, and next-generation communication systems.

Abbas Javadtalab received his Ph.D. in Electrical and Computer Engineering from University of Ottawa, Canada, in 2014, his M.Sc. and B.Sc. degrees in Computer Engineering from Sharif University of Technology, and Azad University of Arak, Iran, in 2002 and 2005, respectively. He is currently a project manager at Huawei Canada. His research interests include quality of service management, congestion prediction, congestion control and network delivery, and network optimization.

Mahdi Hemmati received the M.Sc. degree in electrical engineering (systems and control) from the Sharif University of Technology, Iran, and the Ph.D. degree in electrical and computer engineering from the University of Ottawa in 2017. He is currently a Senior Research Engineer with the Huawei Technologies Canada Research Center, Ottawa, ON, Canada. His research interests include distributed control and optimization, fair resource allocation, and intelligent network automation. He received the NSERC Postgraduate Scholarship for his doctoral studies. He was a co-recipient of the Best Paper Award at IEEE/ACM/IFIP CNSM 2019.

Yuren You is a Senior Principal Researcher at Huawei Canada. His research interests include machine learning, neural networks, optical fibre amplifiers, telecommunication computing, and wavelength division multiplexing.

Abdallah Shami is currently a Professor in the Electrical and Computer Engineering Department and the Acting Associate Dean (Research) of the Faculty of Engineering, Western University, London, ON, Canada, where he is also the Director of the Optimized Computing and Communications Laboratory. He is currently an Associate Editor for IEEE Transactions on Mobile Computing, IEEE Network, and IEEE Communications Tutorials and Survey. Dr. Shami has chaired key symposia for IEEE GLOBECOM, IEEE ICC, IEEE ICNC, and ICCIT. Dr. Shami was the elected Chair of the IEEE Communications Society Technical Committee on Communications Software.