A Survey of ICN Content Naming and In-Network Caching in 5G and Beyond Networks

Oussama Serhane, *Graduate Student Member, IEEE*, Khadidja Yahyaoui, Boubakr Nour, *Member, IEEE*, and Hassine Moungla, *Member, IEEE*

***Abstract*—Internet usability is expanded form just human-to-human interactions toward different communication types, while the communication itself is shifting from the host-centric model to the content-centric paradigm. The 5G and beyond networks promise not only to support such changes but also to provide massive data exchange and connectivity with high reliability. The next-generation networking technologies are the key enabled for 5G that aim at building a new ecosystem. One promising piece of this ecosystem is the information-centric network (ICN), which is a future network architecture that tends to tackle the current host-centric model issues. It natively supports several features, including abstraction content naming and transparent in-network content caching that contribute to improve network performance, reduce traffic, and improve the latency. In this article, we first provide a potential road map by introducing different next-generation active technologies to enable the big picture of 5G, including mobile-edge computing (MEC), software-defined networking (SDN), and network function virtualization (NFV). Then, we discuss the need for ICN and its coexistence within this ecosystem. Later, we present an in-depth review of the recent content naming schemes and a comprehensive review of in-network content caching solutions. We classify these solutions into different classes based on the used technologies and their working principle. Finally, we highlight some research challenges and propose promising directions for the research community.**

***Index Terms*—5G networks, content naming, information-centric network (ICN), in-network functions.**

I. Introduction

TODAY'S Internet is invading all life's areas. It has become impossible to dispense with the services it affords to the various sectors of society, e.g., the economic sector, business, education, etc. Wireless communication, notably cellular communication, has established its meriting due to its ease of use. Driven by each era's requirements, the wireless telecommunications world has evolved from generation to another. The first footprint on wireless communication was the 1G that tended to provide voice communication based on the analogical mode. The 2G came as a successor generation that digitalized the communication to provide better voice quality. The 3G helped in the evolution of the users' devices, including smartphones and tablets, while the ecosystem is witnessing a massive growth in the number of connected devices, exchange content, and rising demands and bandwidth [1]. These challenges led to the development of 4G, which took it upon itself to provide new services and uses cases.

The next 5G standard [2] provides advance wireless communications leveraging technologies to connect people to people, people to data, people to everything, and everything to everything in order to achieve unified connectivity. Indeed, like all new technologies, the 5G is pushed with specific qualifications. According to the 3rd Generation Partnership Project (3GPP) [3], 5G consists of a set of classes including enhanced mobile broadband (EMb) to connect everyone to everything with high coverage area. Besides, the ultrareliable low-latency communication (URLLR) provides not only a high data rate with low-latency links but also extremely ultrareliability and high level of security. The 5G also tends to connect massive thinks, including machines, sensors, and things leveraging the ability to scale down data rates, energy, and mobility, to provide notably low-cost solutions [4].

The 5G network is an application-driven architecture [5] that aims at supporting the network functionalities by providing content caching and computation capabilities at the network near to consumers. In order to enable the next-generation 5G networks, several technologies have been proposed to fill the deficit and reach the needed requirements by offering easier management and low-cost deployment meanwhile ensuring both Quality of Service (QoS) and Quality of Experience (QoE). The mobile/edge computing (M/EC) paradigm [6], one of the most essential technologies in the 5G ecosystem, describes the platform that has the ability to perform computation and executing services close to consumers. This is achieved by enabling computation offloading from the cloud's data centers to the edge servers. On the other hand, software-defined networking (SDN) technology [7] introduces agile management by leveraging a global network's view. The data plane and the control plane in SDN are separated. Therefore, a centralized hardware entity is used to represent the central

Manuscript received April 29, 2020; revised July 24, 2020; accepted September 2, 2020. Date of publication September 7, 2020; date of current version March 5, 2021. This work was supported by the Ministry of Higher Education and Scientific Research of Algeria under Grant C00L07UN290120190001. *(Corresponding author: Khadidja Yahyaoui.)*

Oussama Serhane and Khadidja Yahyaoui are with the Computer Science Department, Université Mustapha Stambouli, Mascara 29000, Algeria (e-mail: oussama.serhane@univ-mascara.dz; khadidja.yahyaoui@univ-mascara.dz).

Boubakr Nour is with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: n.boubakr@bit.edu.cn).

Hassine Moungla is with the LIPADE, Université de Paris, 75006 Paris, France, and also with the UMR 5157 CNRS, Telecom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France (e-mail: hassine.moungla@parisdescartes.fr).

Digital Object Identifier 10.1109/JIOT.2020.3022243

controller, which is directly programmable. SDN is mostly associated with the network function virtualization (NFV) [8]. The latter aims at reducing the deployment cost by decoupling the networks' hardware entities and their dedicated software, then running software on a single platform aiming to provide a single softwarization of network function entity. Each of these technologies tends to provide ultrareliability communication for the 5G network.

Consumers (i.e., end users) are currently motivated by consuming the requested content rather than building a connection with the hoster (i.e., the device or server which owns the content). The communication is now shifting toward the information-centric networking (ICN) paradigm [9]. In the current communication model, the consumer should beforehand know the provider's Internet protocol address (IP) to maintain a session and start the communication. Most of the communication features (e.g., security, mobility, management, etc.) are not built-in [10]. Adding new features and services ends up by designing a complex communication protocol suite that may fail at time and overhead the network resources, which literally affect the network performance and scalability. ICN, on the other hand, uses simple content names to derive communication. ICN decouples the content from its original location, adopts content-based security that secures the content and not the communication channel, and hence applies in-network content caching at the network level. These features make ICN a simple yet efficient communication paradigm for today's Internet [11]. The primary impulse of ICN architecture is to tackle the shortcomings of IP-based networks [12]. ICN changes the concept of host addressing content naming and enables naming abstraction in order to facilitate the integration of new network functions [13], [14]. ICN also adopts a clean communication model where the content is considered the first-class citizen rather than the host or the owner, which brings new opportunities and features by allowing a simple and ubiquitous in-network content caching, mobility, multicast, etc.

The 5G network promises to provide high bandwidth capacity with low latency, massive connecting devices, and high reliability. Indeed, the actual acclaimed evolution in wireless transmission may achieve some of the 5G goals. Yet, several objectives cannot be assured, especially for the long-term perspective, such as ensuring massive content/computation distribution and management, and seamless mobility support. Moreover, the network heterogeneity with different access technologies may degrade the QoE and QoS. Furthermore, reliance on caching only at the access points level may be affected by the future data explosion. While increasing the nodes' caching capabilities is not feasible from both economic and technical perspectives, since it may raise several difficulties in management and delivery, impact the scalability, and costly deployment. ICN can be integrated with 5G networks to overcome different issues [21]–[23]. For instance, the naming abstraction, clean receiver-driven design, built-in in-network content caching, and content-based security can contribute to the realization and development of next-generation 5G networks.

A. Related Surveys

Although very few publications are available in the literature that addresses ICN from the 5G perspective, plenty of solutions are available to address either 5G and their enabling technologies or ICN standalone. Table I summarizes the current existing surveys in the target area. Gupta and Jha [15] presented a review for general 5G architectures, including the ones related to wireless access technologies and device-to-device (D2D) communication. The authors also introduced the NFV and SDN technologies. The prime objective is to study the performance requirements for the 5G network. Nevertheless, the covered solutions are relatively old. Mao *et al.* [17] provided a state of the art about solutions related to MEC-based 5G networks. The authors presented in-depth MEC architecture, solutions, and perspective, by focusing on joint the radio access network and the computation management resource allocation. However, this work considers only the network edge, including the computation offloading efforts.

Wang *et al.* [18] overviewed the next-generation-enabled technologies for 5G networks. The authors discussed the challenges associated with the integration of computing and caching within the 5G networks. However, this work centers only on the edge solutions, besides neglected the ICN-based caching solution. Ahmad *et al.* [19] discussed the security issue related to the next-generation 5G networks, including core/edge network challenges. However, the authors neglect to provide an overview of the pieces that build the 5G ecosystem. Parvez *et al.* [20] surveyed the emerging solutions that enable 5G networks to achieve low-latency requirements. The authors focused on different technologies, including access technologies, such as SDN, NFV, and MEC. However, they neither emphasize on the wireless access efforts nor ICN. Palattella *et al.* [16] surveyed the 5G network focusing on the Internet-of-Things (IoT) use cases. The work targets architectural design, standardization efforts, and business models. However, the authors emphasized more on the IoT use cases rather than the 5G paradigm while neglecting most of the enabling technologies. Afolabi *et al.* [24] present a comprehensive review of the network softwarization and slicing based SDN, NFV, and cloud technologies in order to achieve the end-to-end communication requirements for future 5G networks. The authors also discussed some use cases and emerging applications.

B. Our Contributions

In contrast to the above-listed surveys, our survey focuses on next-generation 5G-enabled ICN networks. Our work provides an extensive and comprehensive up-to-date review of content naming and in-network content caching solutions for 5G-enabled ICN networks. In this regard, the major contributions of our work are summarized as follows.

1) We present an in-depth overview of the leading enabler next-generation network technologies, including 5G networks, MEC, SDN, NFV, and ICN.

TABLE I
COMPARISON OF CONTRIBUTION OF RELATED SURVEY PAPERS

Ref.	Technologies	Recent Work	Future Direction	Covered Solutions	Covered Year	Limitation	Year
[15]	5G, D2D	2015	✗	D2D	2012-2015	• Focus only on the edge access • Do not cover recent solutions	2015
[16]	5G, IoT, D2D	2014	✗	Wireless technologies	2010-2014	• Do not cover 5G enabled technologies	2016
[17]	MEC, 5G, Cloud	2015	✓	Computation offloading	2012-2015	• Focus on computation solutions • Ignore the core network	2017
[18]	MEC, D2D, Mobile cloud computing	2017	✓	Caching and Computing	2015-2017	• Ignore ICN caching efforts	2017
[19]	SDN, NFV, MEC	2017	✓	Security issues	2015-2017	• Missing next-generation technologies • Missing technical clarification	2018
[20]	5G, SDN, NFV, MEC	2016	✓	D2D Caching	2014-2016	• Focus on access solutions • Skip the basic communication model	2018
Our	5G, MEC, SDN, NFV, ICN	2020	✓	Edge/Core Naming and Caching	2010-2020	/	2020

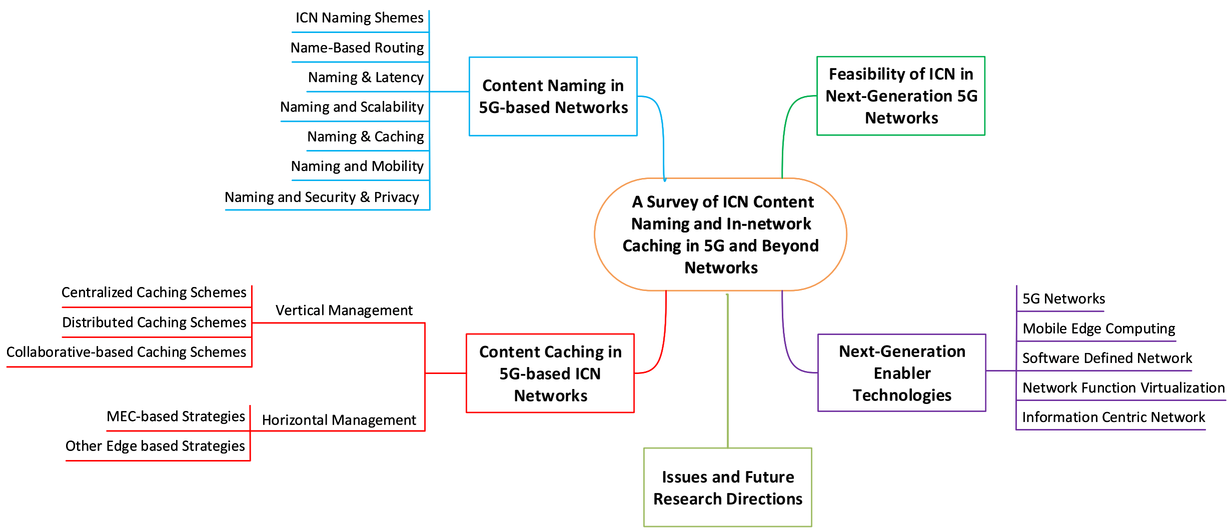

Fig. 1. Survey organization and taxonomy.

2) We highlight the applicability and the feasibility of the ICN paradigm in the next-generation 5G networks.
3) We review in detail the existing content naming solutions in 5G-based ICN networks. We classify these solutions into different classes based on their objectives.
4) We comprehensively survey the existing in-network content caching schemes in 5G networks over ICN. We divide these solutions into two main categories according to their working principles.
5) We highlight challenges and issues faced by the 5G-based ICN and provide some future research directions.

C. Organization

The remainder of the article is organized as shown in Fig. 1. Section II provides a comprehensive overview of most technologies to enable network generation networks. Section III discusses the feasibility of using ICN to enable next-generation 5G networks. Section IV presents the existing content naming schemes in the era of 5G-based ICN networks. Similarly, Section V overviews the state-of-the-art in-network caching solutions in the era of 5G-based ICN networks. Section VI discusses the existing issues and challenges and highlights various guidelines and research directions. Finally, Section VII concludes the article.

II. NEXT-GENERATION ENABLER TECHNOLOGIES: OVERVIEW

Various technologies, as depicted in Fig. 2, have been proposed and developed toward the realization of next-generation networks (also known as future Internet). This section aims to provide a comprehensive overview of these enabler technologies, including 5G networks, MEC, software-defined network, NFV, and ICN.

A. 5G Networks

The massive explosion of the devices and their usage so far has coincided with the increased cravings of bandwidth consumption accompanied by the massive data generation,

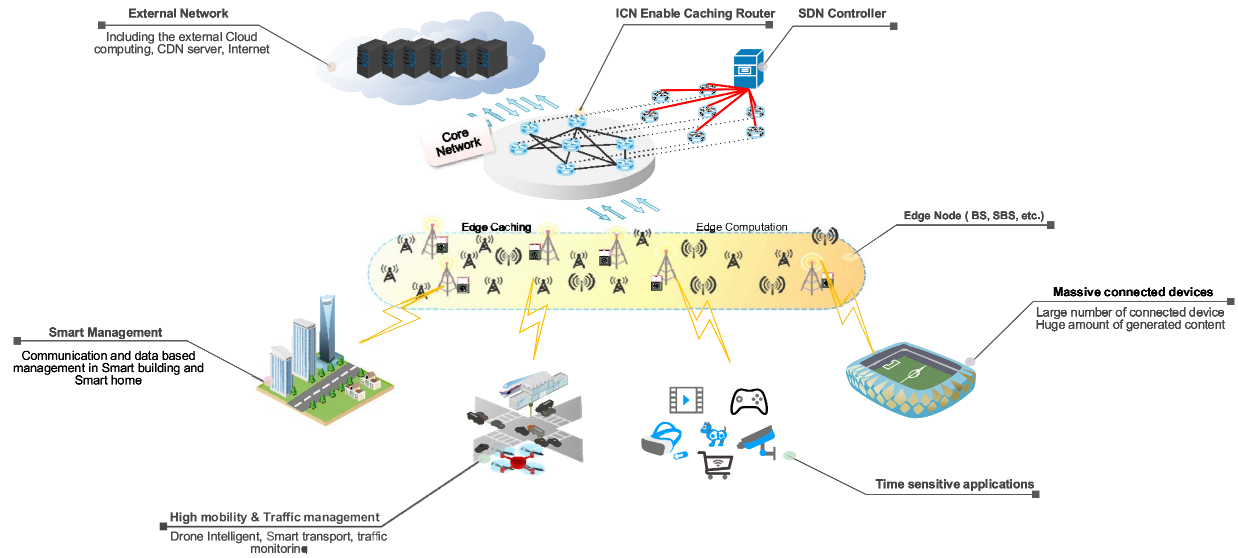

Fig. 2. Overview on the next-generation network.

TABLE II
USE CASES COVERED IN 5G NETWORKS

Use Case	Example	Enabled Technologies
EMb	• High speed video delivery • Massive coverage • Huge device connection	• Small Base Station • Heterogeneous networks • Wi-Fi access
M2M	• IoT device interaction • Intelligent sensors • Vehicle-to-Vehicle	• Millimeter waves • Massive MIMO • Beamforming
URLLC	• Robotics controlling • Tactile Internet • self-driving car	• Core & Edge Network

resulting in poor QoE and QoS. On the other hand, the rapid development of the communication pattern reviles to new communication usage, e.g., machine to machine connectivity (M2M) communication, which requires low-latency access. Researchers are trying to find alternative solutions by pushing the communication to a new dimension with high reliability and high data rate aiming at satisfying all the necessities. Unlike the 4G cellular network, the new 5G network is not only a new wireless network but a global ecosystem that has the ability to provide all the requirements and needs, starting from high coverage—tends to connect everything to everything, to high bandwidth, and low-latency achievement [4].

5G Requirements and Specification: According to International Telecommunication Union (ITU) [25], the fundamental use cases of 5G networks, as presented in Table II, are summarized in three essential points: 1) EMb; 2) M2M; and 3) ultrareliable and low-latency communications (URLLCs).

The EMb describes the different use cases fall into the massive data delivery that requires a high data rate, especially in vast coverage areas. The 5G network highly dependents on different radio access technologies including the higher density of base stations (BSs), small BS (SBS), heterogeneous networks, and WiFi access in order to ensure high area coverage as well as increasing the network throughput [26], [27]. This allows the 5G wireless network to be promised as a technology to replace the current wire access [28].

On the other hand, M2M communication contains the state of communication between multiple machines without human intervention. These use cases introduce a new communication model among different devices equipment, including IoT sensors and actuators, that communicate with each other [29]. Therefore, the 5G network depends on new communication techniques, e.g., millimeter waves and massive MIMO, taking into account the sensitivity of communication between these machines.

Finally, the URLLC category is the fundamental element of 5G networks. It may contain more use cases from other classes. Indeed, this class aims at providing low latency and extremely high reliability. The network task lies to provide low latency for various time-sensitive applications, and hence, leveraging with new radio access technologies and the MEC paradigm and robust core management technologies.

5G Promises and Outlooks: To date, there have been many thoughts of what the 5G network will able to provide. In the following, we list the most important 5G promises.

1) *Massive Devices Connectivity:* According to [30], the 5G network will be handling more than 2.5 billion subscribers. The management of the massive explosion in the connecting device's area is one of the goals behind this new network technology.
2) *Bandwidth:* One of the 5G network trends is to provide a massive throughput and capacity. Indeed, the availability of multiple frequency bands confirmed that transmission speed could reach up to 10-Gb per second. End users are able to download a full high-quality movie in less than two min [31], which drastically improves the users' QoE.

3) *Low Latency:* One of the promises for the 5G network is delivering a high bit rate. The common latency requirement is less than 1 ms for both uplink and downlink in roundtrip transmission [32]. This will revolutionize applications, especially time-sensitive application.
4) *Mobility:* Robust mobility management is one of the pillar requirements of the 5G networks. Indeed, ensuring a seamless users' mobility between different access points is as important as ensuring the basic connectivity which should be provided almost 0 handovers and dropped packets [33].
5) *Security and Reliability:* Regardless of the network heterogeneity, the 5G network promises to provide not only robust connectivity and control but also ensure security, trust, and privacy [34].
6) *Energy Consumption:* The 5G network promises to reduce the energy consumption of all entities embedded at the network, including access points, users' device battery life, and IoT devices that have limitations in energy.

The impact of the 5G networks (compared to the previous technologies) appears in providing better service at a lower cost and supporting the new emerging applications, e.g., augmented reality, virtual reality, gaming, as well as enhancing multimedia services, such as ultrahigh definition display and multiview high definition display. According to [35], the economic contribution of the 5G network is up to $12.3 trillion of the global economic output. In the remainder of this section, we present the key enabling technologies that contribute to the success of the 5G ecosystem.

B. Mobile-Edge Computing

MEC [17], also known as multiaccess edge computing, refers to the platform that performs computation and caching services, close to end users, at the edge of the network [36]. Fig. 3 depicts a high-level design of MEC functionality within a 5G network, including different interactions between the MEC/Could and MEC/Devices. MEC empowers the 5G network to support the ever-increasing amount of mobile applications and the pursuit of endeavors URLLC by offloading some services from the cloud the edge server [37].

In contrast to fog computing [38], [39], MEC is limited to the mobile cellular network and may not provide computing along the entire communication path. In MEC, the radio access network may share its resources to provide fast data delivery, reduce communication latency, and decrease network congestion. MEC is a suitable paradigm for a range of today's Internet applications [40], such as low-latency content delivery, data analytics, and computation offloading [41], [42].

Emergence of MEC: Cloudlet [43] is considered as the initial design of edge computing. It has been proposed by the University of Carnegie Mellon in 2009. The idea behind Cloudlet is to install computer hardware at the edge of the network that provides computing capacity. Thus, the connecting devices use the Cloudlet platform to perform computation rather than reaching the cloud. Although Cloudlet proves its worth, it is not designed for completely mobile infrastructure. In particular, Couldlet is supposed to be mostly accessed by connecting devices through WiFi connection, which impacts the mobility of users. The MEC concept, on the other hand, has been developed by Industry Specification Group (ISG) within the European Telecommunications Standards Institute (ETSI) in 2014 [44]. MEC describes the framework that has the ability to offer cloud computing capability and an Information technology services environment at the edge of the 5G network, by offloading computation tasks from connecting devices to the MEC platform.

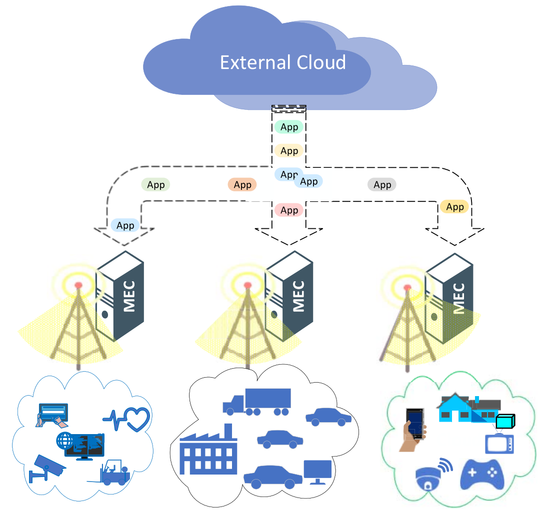

Fig. 3. Overview of MEC paradigm.

Advantages of MEC: In MEC, cloud computing services are expanded to the edge of the network. The computational tasks, from mobile devices to the MEC server, not only enhance QoE and maximize battery life on the user side but also reflect on the whole network. The unique advantages that are potentially offered by MEC can be summarized as follows

1) *Efficiency:* MEC server will be closely integrated with connecting devices systems to enhance overall efficiency and performance. Based on real testbed by the computation offloading to the MEC platform, Dolezal *et al.* [45] demonstrated that the energy consumption of the user device's reach up to 93%, besides a decrease of the latency up to 88%.
2) *Low Latency:* MEC provides powerful computing from the nearest location to user devices, thus reducing the need to transfer processing tasks to the remote cloud through the core network, especially for real-time applications requirement. Thus, MEC reduces communication latency and decreases network congestion.
3) *Security:* The proximity of MEC to users' devices leads to a reduction in the distance that information needs to traverse, which helps in diminishing the chance of eavesdropping.
4) *Cognition:* In MEC, the radio access network may share its resources and pool resources along the edge. MEC can distribute computing power and control functions anywhere in the 5G network and connect devices to take full advantage of the resource's availability.

Benefits of MEC: MEC technology is one of the keys that tends to reduce the backhaul link's congestion while achieving the latency requirement for the 5G network. Thus, there will be diverse prospects and features for the MEC platform. From the end-users' point of view, MEC provides offloading services for different use cases (e.g., virtual reality, M2M, and driving assistance system) to enhance QoE by improving the energy efficiency and achieving the required data delivery deadlines. Moreover, the MEC platform enhances the network performance by providing optimization applications to reduce the load on the backhaul link, especially during peak hours. Furthermore, for the third party services' perspective, MEC empowers the big data interpretation for different treatments. MEC can also contribute to the effectiveness of the network resources, for example, extending the caching capacity of the memory nodes size toward comprising some content [46].

C. Software-Defined Networking

The fundamental equipments of today's network infrastructure are routers and switches. These devices are mainly responsible for packets switching, forwarding, and routing. Managing and controlling a large number of these devices may end up with different issues in terms of scalability and maintenance. SDN technology [47] is a new network paradigm that aims at decoupling the control plane from the data plane and hence providing a centralized controller. In other words, the device entity becomes simple data (packet) forwarding called SDN switch, where the forwarding decisions are flow based. This controller provides an efficient control to the whole network via software-based programming interfaces (i.e., southbound interface and northbound interface). Network devices become simple forwarder switches while all routing/forwarding rules are decided by the centralized controller [48]. The use of SDN makes the network management easier and more efficient, improves the overall performance, and facilitates network evolution.

According to the open networking foundations (ONF)'s white paper [49], a reference SDN architecture may consist of three layers, as shown in Fig. 4 and explained in the following.

1) *Management Plane:* It is the upper layer in the architecture and known as the application layer. It provides an access point in various forms, such as application programming interfaces (APIs). An SDN-application can conveniently access to the global network view. Thus, SDN offers a Platform-as-a-Service (PaaS) model for networking [50].
2) *Control Plane:* It is the middle layer in the architecture and known as the core layer. It serves as a bridge between the two other layers via the Southbound Interface and Northbound interface. The control layer consists of a centralized controller with the responsibility of downward flows, controlling and collect network status from the infrastructure layer, and provide a global view of the network to the application layer. In the upward flow, the controller takes requests from the application layer to manage the network devices in the infrastructure layer [47].
3) *Data Plane:* It is the lowest layer in the architecture and known as the infrastructure layer. It consists of SDN switching devices. These elements are interconnected with each other to formulate a single network aiming at collecting the network status, storing this information temporally, sending them to the controllers, and processing packets based on rules provided by the controller [51].

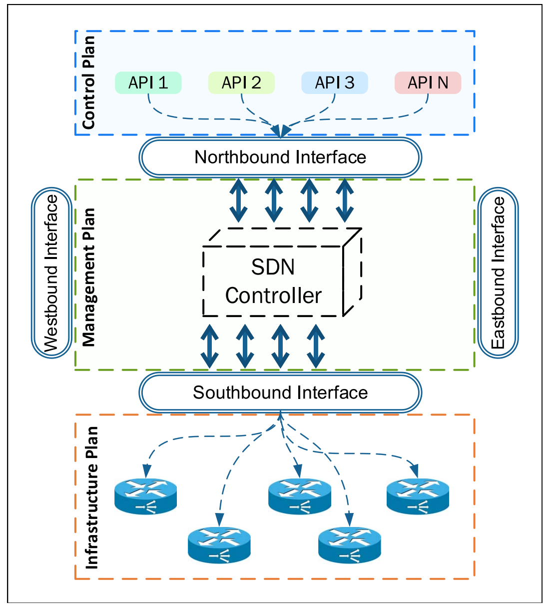

Fig. 4. Overview on the SDN technology.

SDN technology tends to provide flexibility and agility to network management. The key is to allow the control plane to be directly programmable through an open interface. For instance, one of the open standard protocol is OpenFlow [52]. Moreover, each SDN controller maintains two interfaces given as follows.

1) *Northbound Interface:* The SDN controller uses the Northbound Interface to handle transactions with the application layer in order to provide a global view of the network and receive forwarding rules.
2) *Southbound Interface:* The SDN controller uses the Southbound Interface to deal with transactions with the infrastructure layer in order to collect network status and update packet forwarding rules in SDN switch devices.

One of the promises of the 5G network is to enhance mobile broadband. In doing so, the SDN must handle a large geographical area in a distributed fashion. Each SDN controller maintains two additional interfaces: East and Westbound Interface introduced to support multiple SDN controllers. By leveraging these interfaces, the SDN controllers intercommunicate to share network information and coordinate their decision-making process. Various implementations of distributed SDN controller has been proposed, including Onix [53], Disco [54], ElastiCon [55], and ONOS [56]. Each architecture has a dedicated approach. Some of them use a distributed file system in order to maintain a logically centralized and distribute the control plane, while others apply a master/slave communication style.

SDN Features: SDN delivers numerous features to the 5G network since the network entities (switch/router) became a simple packet forwarding devices that flow the injected flow rules. This can help in:

1) SDN not only provides the ability for networks to be programable but also unifies the multivendor and multi-technology environment. This enables network operators and service providers to innovate faster and allow software-based innovation [4];
2) SDN provides high transmission speed since SDN switches do not need to perform any matching or packet inspection to make a forwarding decision for each packet. This simplifies their tasks and provides the efficient use of network resources [50];
3) the network administrator has the ability to alter and manage the needed characteristics to nodes from a central location. This overrides the configuration for each router separately, especially for large-scale networks [51].

In the SDN architecture, the northbound APIs connect the centralized controller with the services and applications running over the network. The main job of this component lies to transfer the needs for each application, e.g., storage, bandwidth, etc., where the network tries to satisfy the request of each application.

D. Network Function Virtualization Among the Essential Components of Today's Network

The embedded hardware appliances that afford essential network functionality. These appliances are also known as Middleboxes as they are integrated to provide network services, such as deep packet inspection (DPI), Firewalls, network address translation (NAT), load balancers, etc. The network operators pay high costs to deploy and manage these Middleboxes in order to meet the network services requirements. However, the new services' innovation raises new challenging problems, such as providing these services in the required dedicate period. Moreover, these appliances still face the manageability problems that result in high expensive deployment cost and time consuming, in which, service providers consider it as failure modes [57]. To overcome the aforementioned issues and problems, NFV has been introduced an emerging paradigm that aims at extracting the software from its installed hardware and therefore providing flexibility and agility to the networks' services [58].

NFV, first introduced in October 2012 [59] authored by over 20 of the world's largest telecommunications operators, including Deutsche Telekom, American Telephone, British Telecom, etc. They formed an ISG within the ETSI to bring the NFV's philosophy to reality. Fig. 5 depicts the transformation model from the physical model toward the NFV system. In traditional mechanisms, the data packet must be handled by multiple dedicated hardware servers that offer multiple services respecting sequential and logical chaining. The NFV-based system aims at accelerating this process based on the virtualization technology that improves the end-to-end interconnection.

NFV, technically, tends to decouple between the software and the dedicated hardware. Therefore, this software becomes a virtual function instance or virtual network function

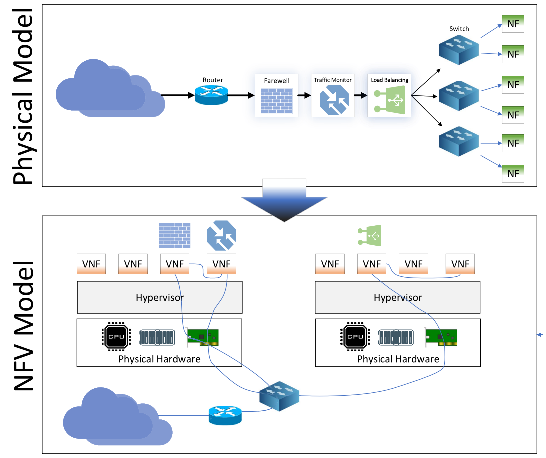

Fig. 5. Overview on NFV.

(VNF) [60], that runs on a virtual machine (VM) in a dedicated server, e.g., standard commercial off-the-shelf servers (COTs). Functions are chained together and orchestrated in order to provide a network service [61]. Hence, the NFV technology brings vital features, such as facilitating network services provisioning and improving network scalability, flexibility, and cost efficiency.

NFV Benefits: The major benefits of NFV can be summarized as follows.

1) NFV contributes not only to reducing the middleboxes density in the network but also in facilitating the creation, deployment, management of services, and reduce the capital expenditure (CAPEX) and operational expenditure (OPEX).
2) By adopting NFV, the physical radio resources can be abstracted and sliced into virtual network resources that hold certain corresponding functionalities. Therefore, multiple parties can share the same infrastructure with isolated network services.
3) NFV tends to reduce end-to-end network interaction, which can reflect on reducing the 5G round-trip latency and decreasing the network delay.

The NFV aims at optimizing the network functions, which bring potential benefits to the network through enhancing service delivery and reduce cost. SDN has the capability to handle decisions of massive infrastructure devices (switches) from a single controller, which facilitates the network administration. Both SDN and NFV contribute to improve network flexibility and provide fast service deployment. The ONF published several white papers that describe the coexistence between the NFV and SDN. For instance, work in [62] illustrates the architecture design and highlights its advantages.

Indeed, the merger of NFV and SDN may contribute to reducing network management, enhancing services provisioning [63], decreasing the complexity, and allowing for software-based innovation. However, the need for agile content distribution highlights the major limits of the IP-based model. Moreover, total lean on the virtualization technologies to deploy the 5G network brings the traditional VM

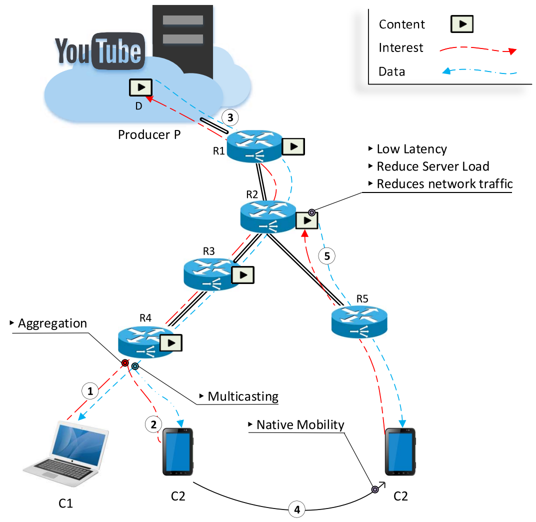

Fig. 6. Overview on ICN communication.

issues, such as VM migration [64] since traditional IP mobility management remains a hard challenge to address.

E. Information-Centric Networking

ICN [65] is a new communication paradigm that promises to replace the current IP concept. ICN is built upon the use of the content name as the primary network element rather than the host address (IP) [66]. To meet the requirements of future Internet architecture (FIA) [67], various architectures have been proposed under the umbrella of ICN, including data-oriented network architecture (DONA) [68], content mediator architecture for content-aware networks (COMET) [69], convergence [70], mobilityfirst [71], scalable and adaptive Internet solutions (SAILs) [72], publish-subscribe Internet technology (PURSUIT) [73], Nebula [74], expressive Internet architecture (XIA) [75], ChoiceNet [76], content-centric network (CCN) [77], named-data network (NDN) [78], and green ICN [79]. Although these architectures have different design principles and implementions, all of them are based on the use of the content name as the main network element to drive the communication.

Name-Based Features: By using the content name as the primary network element, ICN decouples the content from its original location, which enables the network layer to cache and serve content for future demands regardless of the availability of the original content publisher [80]. Moreover, ICN implements content-based security [81], which means all security-related information is applied to the content and not the communication channel. These information travel and cache with the content in all its usage cycle.

Fig. 6 illustrates the ICN communication concept. Assuming the node $C1$ is requesting content D, the ICN concept requires the consumer to specify only the name of the requested content. The content is fetched from the original producer P, and all intermediate nodes in the communication path are able to keep a copy of the content in their content stores (CSs). When another consumer, e.g., $C2$ requests the same content D, the demand will be satisfied by $R2$'s CS without the need to communicate with P. Besides, when there is a content transmission error, the retransmission process is assigned to the closest node (replica node) that has cached the requested content. The consumer does not need to make another request to the original producer P [82]. Similarly, during node mobility, an ICN node is not required to request a new IP address after connecting to the new network. Instead, it sends only the unsatisfied requests using only the content name [83].

Content-Based Security: Unlike IP-based networks, ICN maintains security by applying a content-based security concept [84]–[86]. Since the content is the first citizen in ICN architecture, ICN secures the content itself rather than securing the communication channel. All security-related information is stuck with the content and travels with him during the communication. This helps to provide a fully decoupling of content from the location.

In-Netwok Caching Features: In ICN, each node can cache the frequently used content leveraging their embedded memory [87]. ICN architecture built-in content caching management services that enhance the end-to-end delay. The Interest can be satisfied with intermediate nodes, which decreases the latency and reduces the network conjunction. The in-network content caching maintains two main operations: content placement and replacement [88]. The former consists of selecting which content to cache and at which node (placement). The latter aims at selecting which content to remove from the CS to keep room for the newly arrived popular content (replacement).

Interest-Based ICN Architectures: CCN and NDN, in particular, are the most active ICN projects that use hierarchical human-readable names to identify content. An NDN router maintains three data structure: 1) *CS:* to temporary cache data and serve it for future requests; 2) *pending interest table (PIT):* to keep track of the unsatisfied interest and to deliver the data packets back to consumers; and 3) *forwarding information base (FIB):* to record the suitable interfaces for each reachable name prefix.

NDN is a receiver-driven architecture that implements the Interest-data exchange model. An Interest message is triggered by a consumer to request content. This message carries out the requested content name and is forwarded upstream using name-based routing rules. When the Interest message reaches the original content producer or a replica node that has the requested content, a data message is delivered back carries the requested content.

F. Summary and Insights

The 5G network promises on affording high-quality services with low-latency requirements and the highest reliability. The application-driven nature of 5G networks contributes to sharing the same infrastructure between different services while achieving better resource exploitation and enough QoE and QoS taking into consideration diminishing in the CAPEX and OPEX. A multitude of technologies lies behind the 5G revolution. The purpose is to invest a new network architecture that implements both computing/caching capabilities. Hence, leveraging next-generation networking, including MEC, which is mandatory to improve cloud computing capabilities and enhance the network edge. Besides, SDN and NFV contribute to rendering the agile network management that facilitates innovation and network programmability. Finally, the role of ICN communication is obvious as a clean model that contributes to enhance content distribution and delivery and improve caching management.

III. Feasibility of ICN in Next-Generation 5G Networks

The clean and straightforward design of ICN architecture helps in implementing a simple yet efficient communication platform for 5G by avoiding complex protocols to ensure security, mobility, and data management [89]. On the other hand, the use of content naming instead of host addressing helps in providing a unified platform for different applications and use cases within the same APIs. Finally, employing in-network content caching helps in affording content at the network level regardless of the original producer's reachability and enhancing the QoS and user experience [90], [91].

Different benefits can be achieved in 5G networks if ICN is adopted as a communication enabler. For instance. Li *et al.* [92] proposed an SDN-based orchestration framework that merges MEC functionalities and ICN/CCN capabilities at the edge network. The authors used the SDN technology to support both ICN and IP forwarding by decoupling the data plane and the forwarding plane. The management process is assigned to a centralized SDN controller. The authors also designed an SDN-based forwarding strategy in order to support both ICN name-based forwarding and the conventional packet-level flows. The simulation results show the effectiveness of the proposed architecture that outperforms the conventional structure. However, the proposed solution does not take the case of a large-scale network where the transaction from IP to ICN and *vice versa* become substantial and challenging. Tourani *et al.* [93] proposed an ICN-based mobile converged network (ICN-MCN) architecture. The authors used ICN in order to enhance the device's multihoming when using different embedded interfaces in the users' device. Hence, the network provides fast content retrieval and seamless node mobility via both content and host multihoming. ICN-MCN aims at enhancing the user's QoS and QoE. The authors also discussed some opening challenges to bring novel architecture in real-world applications, such as adopting the native support of multihoming in order to enable multiinterface and ensure fast transfer for the high volume content, the need for an efficient caching and cache replacement strategy, and addressing the producer mobility-related issues. However, such an architecture may need major modifications of heuristics algorithm to be adapted in high network heterogeneity and address the consumers' behavior, such as mobility and the diversity in demand for better bandwidth utilization.

Lertsinsrubtavee *et al.* [94] proposed Picasso, an ICN-based MEC framework. Picasso is designed to adapt within the high network dynamic where service delivery can fail due to links' instability. Therefore, the authors exploited the ability of ICN to not only use for content delivery but also service delivery. The solution is based on the Docker platform to create and deploy services. Picasso takes benefits from ICN features, such as caching the created services and fast service retrieval and delivery. Compared with the traditional IP-based MEC frameworks, Picasso achieves high traffic accomplishments and low latency. However, the proposed solution may violate the simplicity of ICN architecture. It needs additional modifications in the design principle to support additional communication models, such as push and pull traffic. Zhang and Zhu [21] focused on QoS provisioning for content transmission in 5G networks. The authors proposed an architecture that aims at guaranteeing massive data delivery under delay bounded. The architecture combines both NFV, SDN, and ICN in which the latter plays a vital role due to its receiver-driven mechanism and the pervasive in-network caching. NFV is mainly used to encapsulate the physical layer and SDN as global network management. The proposed architecture aims at jointing the needed resources to select the optimal delivery path that fulfills the QoS requirements. In doing so, three scenarios of virtual network selection and transmit power allocation schemes have been presented, including a single requester, multiple requesters, and under noncooperative gaming theory among all requesters. The results prove the efficiency of the solution. The noncooperative game strategy shows a high convergence over the Nash equilibrium. However, the proposed schemes may be foolproof in the wired connection, but the instability of the wireless communication link may pose some issues. Liang *et al.* [14] combined ICN, NFV, and SDN to virtualize wireless architecture in 5G networks. The main idea consists of taking advantage of the virtual network created and managed by the NFV and SDN in order to integrate the ICN-based communication model. The objective is to allow the network to share the same infrastructure (via virtualization) and embed content caching simultaneously, hence, reducing the redundant cached content among different slice. The architecture enables virtual resource allocation (including content slicing, network slicing, and the flow slicing) with the caching strategies as an optimization problem. The objective is to maximize the gain of the virtual network operations to improve the end-to-end network performances. Therefore, the authors suggest the interior point method to solve the problem. Compared with the traditional approaches, the simulation results show that the proposed scheme has a high ability to reduce pressure on backhaul-link. However, it does not take into account user behavior, e.g., mobility. Moreover, the caching strategy may need more investigation since it is based only on content popularity.

Benkacem *et al.* [23] proposed an ICN-based content delivery network (CDN) that aims at enhancing the content delivery in the 5G next-generation networks while reducing the core network conjunction. The proposed architecture leverages MEC and NFV to joint both ICN and CDN to provide an active content distribution service. The architecture consists of a double level of content cache/distribution system. At the same time, the consensus algorithm tends to connect between them (the CDN slice and the ICN slices) at a specific getaway. The simulation results show that the proposed platform outperforms the conventional IP-CDN-based in terms of throughput, publishing time, and content retrieval. However, the major drawback can be the high overload at the ICN/CDN gateway regarding the size of the transferred content. The applicability of this architecture remains open questions. Zhou *et al.* [22] proposed an integrated framework to support caching, computing, and communication within the 5G next-generation networks. The essential components of the proposed architecture are MEC to provide computation, ICN to enable caching capability, and SDN to act as a global virtual network manager. The objective is to provide a virtual network on-demand service with specific requirements, e.g., virtual network with only computing, only caching capability, or with both. To do so, the authors consider all possible elements interaction use cases between MEC and ICN (the framework usability includes MEC-assists ICN and/or ICN-assists MEC). They then adopt a distributed alternating direction method of multipliers (ADMMs) algorithm [95] to tradeoff between caching, communication, and computing resources. The simulation results prove the hypotheses of the need to jointly computation and caching and using them to assist each other. Although the architecture allows better use of resources on the same infrastructure, the security and privacy aspects have not been covered.

Each work of the studies mentioned above attempts to provide a different solution that joins the ICN within the 5G network. Some of these efforts consider the content to be allocated as a network resource, and others tend to optimize the bandwidth based on the in-network content caching feature, etc. Besides, the majority of these works try to enhance the 5G network performance by leveraging ICN as a communication enabler. Therefore, this can be achieved based on its fundamental elements which are *content naming* and *in-network content caching*. In the following two sections, we systematically review the existing content naming and in-network caching schemes in 5G-based ICN networks.

IV. Content Naming in 5G-Based ICN Networks

The content name is the key element in ICN. An ICN name should be a unique identifier and as small as possible. It should also be persistent to validate the content. The used naming scheme must be scalable and allows name aggregation.

A. ICN Basic Naming Schemes

ICN introduces four main types of naming schemes.

1) *Hierarchical Name:* It consists of multiple components to identify the application and describe the services. It has a similar structure to current uniform resource identifiers (URIs) and may originate user friendly and convey meaning names to users. Hierarchical naming enhances scalability since name prefix can be aggregated. However, the main drawback appears in the small content size where the content name becomes much larger than the content itself. Fig. 7(a) depicts an example for content name prefix in a hierarchical format.
2) *Flat Name:* It is mainly obtained through hash algorithms applied to content. There is neither a semantic nor a structure behind a flat name. The generated name is not human friendly and can hardly be assigned to dynamic content that is not published yet. Flat naming has scalability issues since it does not support routing aggregation. Fig. 7(b) shows a flat name for the hierarchical name in Fig. 7(a). This result has been obtained by using the MD2 hashing algorithm.

/app/service/movies/2009/3D/main.mp4

(a)

6a69dfe882ea327b86050a9224b4f6ca

(b)

FileType <String>: mp4
Title <String>: main
Duration <Integer>: 162
Nbr Chunks <Integer>: 200
Organization <String>: century studios
Year <Integer>: 2009

(c)

/app/service/7595fb58/3D/049db7d1

(d)

Fig. 7. ICN naming schemes for the same content.

3) *Attribute-Value-Based Name:* It has a collection of attributes; each attribute has a name, a type, and a set of possible values (creation date/time, content type and location, version, etc.). Collectively, they represent a unique content and its properties. The attribute-value-based naming scheme supports an easy searching process by using known content keywords, which means it is quite possible to find many contents against one search request. Hence, it is hard to ensure naming uniqueness and find unique content in a short period. Fig. 7(c) illustrates attribute naming scheme's example.
4) *Hybrid Names:* It combines at least two of the previously discussed schemes or all of them. The idea is to take advantage of the base scheme's best features to improve network scalability, performance, and enhance security and privacy. For example, taking advantage of name aggregation to enhance the lookup process, and the fixed length of flat names to save space and attribute values to provides keyword searching and security/privacy. Since each content has a unique name identification, the generation of the hybrid names may become a challenging issue, notably for nonstatic and real-time generated content. Fig. 7(d) depicts an example of hybrid name.

Each ICN architecture uses different naming scheme that addresses specific issues. For instance, MobilityFirst [71] adopts flat names and combines broadcast and content based (CBCB) [96] while NDN [78] uses hierarchical names. Zhang *et al.* [97] presented notes about differences and similarities between the naming in both IP and NDN architecture. The paper provides clarifications regarding the prefix components for ICN structure, including location, address, locator, and route. The authors concluded that the ICN naming and routing communication model fits the appropriate communication abstraction than that in the IP model. Adhatarao *et al.* [98]

presented a qualitative and quantitative comparison for both hierarchical and flat names. The authors used different metrics in their study, including name lookup efficiency, name aggregate feature, semantic of produced names, and namespace manageability. The study proves that the forwarding performance is directly impacted by the lookup time. Hence, flat name records a fast lookup, but the explosion of content in real-world scenarios can create a real dilemma since each content is identified with a unique name. On the other hand, the forwarding performance on top of hierarchical names is mainly affected by PIT and FIB's size. However, hierarchical names support the name aggregation feature, which reduces the size of the tables and the lookup time. Hierarchical names also improve multicasting that can send one packet to multiusers. Finally, the study concludes that hierarchical names achieve a higher lookup complexity than flat names since the lookup process requires parse and lookup for each name component to determine the next interface.

B. Name-Based Routing

The use of names to identify the content introduces a name-based routing concept to discover and deliver the content to one or more requester without using any host identification [99]. When a consumer triggers a request asking for specific content. A discovery process will be started to find the content based on its name. The request packet is forwarded hop by hop using name-based rules to specify the next forwarding interface until reaching the original or a replica node that has the requested content, hence the content will be delivered back to the requester [97].

Fig. 8 describes an overview of ICN content discovery and delivery mechanisms. The consumers C1, C2, and C3 send requests to get different content located at the producer P. First, the C1 starts requesting content (e.g., video) from P with interest name `/P/Videos/V1.mp4`. Then, the C2 sends an Interest to get content (e.g., image) using the Interest name `/P/Images/Img1.jpg`. Both Interest packets are routed toward the producer using the FIB table. Indeed, the Interest carries the same name prefix, which allows for the aggregation at router R5 and, thus, reduces the network traffic. The C3 needs a document file by issuing an Interest with the name of `/P/Documents/Book.pdf`. The Interest traverses the network hop by hop until the CS of the router R2 and finds the desired content. Content caching reduces the distance to travel and hence reduces the delay.

In a nutshell, the name prefix is responsible for driving Interest packets to content provider. Any changes in the naming scheme design can directly affect the performance of the lookup, forwarding, and network performance in general. In the following sections, we review the existing naming schemes by focusing on their aim objectives, Table III provides a summary of this discussion.

C. Naming and Latency

During the Interest forwarding process, each node checks whether it has the name prefix on its PIT and FIB tables. This verification mechanism is called the lookup process. The

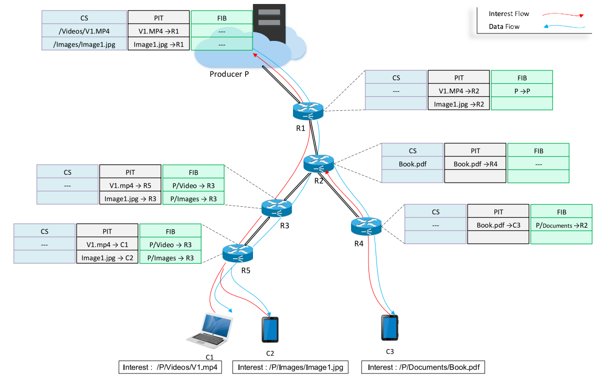

Fig. 8. Content discovery and delivery in ICN.

required time to perform lookup directly impacts the Interest delivery, which increases the network delay due to the verification time. In other words, designing an appropriate naming scheme can reduce the lookup time, and hence reduce the network delay and the overall latency. Nour *et al.* [100], [115] designed a multilayer multicomponent hierarchical attribute-value naming scheme that combines prefix labeling with variable-length encoding methods. The authors use the decimal classification and Fibonacci encoding schemes to identify the hierarchical location of the content. Besides, they incorporate different attributes in each naming level with a set of properties. The proposed scheme consumes less memory and provides a short lookup time and routing process. However, the obtained names are not friendly and use only numbers which are not human readable.

Arshad *et al.* [102] proposed a hybrid naming scheme that combines both hierarchical and flat structure. The overall goal is to reduce the lookup time by minimizing the name length. The scheme also tends to support the push and pull communication models for specific use cases. The hierarchical naming structure is used to describe the domain name, location, and task. The flat name describes both device name and data. The authors also designed specific algorithms for both Interest and Data packets forwarding based on hash functions. The results prove the efficiency of the proposed scheme in different uses cases (mobile and static nodes) in terms of loop avoidance and improve the latency and the Interest aggregation while reducing the lookup delay. However, when an Interest packet reaches a node, the forwarding process time may raise extra upstream delay. Therefore, it affects content retrieval time and QoS. Similarly, Rehman *et al.* [103] proposed a hybrid naming scheme based on hierarchical and flat structure. The hierarchical prefix is mainly used to route the Interest and Data packets. The core idea is to use a query system based on flat names. The authors used a hash function and defined specific keywords to interact with the smart building scenario. The proposed scheme may fill the gap of smart interaction between different devices under different requirements and communication models. The simulation results prove the efficiency of the proposed scheme in terms of energy consumption, average delay, and the number of packets processed per node. However, the applicability of a query system as a naming scheme in general usage is an open question.

TABLE III
SUMMARY OF ICN NAMING SCHEMES

Ref.	Type	Features	Comparison	Limitations	Year
			Latency		
[100]	• Hybrid	• Short name length • Lookup time improvement	• Hierarchical • NDN-HNS [101]	• Not human-readable • Ignore nodes mobility	2020
[102]	• Hybrid	• Avoid communication loop • Support push and pull model	• Flat & Hierarchical	• Complex Interest forwarding process	2018
[103]	• Hybrid	• Introduce Query system	• ISC [104] • HFHN [102]	• Extra Interest processing time	2019
[105]	• Hierarchical	• Introduce naming encoding • Reduce lookup time	• NCT [106] • LNPM [107]	• Lack of push and pull model. • Impact of Interest decoding	2019
			Scalability		
[108]	• Hierarchical	• Flexible naming scheme	• Hierarchical	• Content representativity • Semantic issue	2016
			Caching		
[109]	• Hierarchical	• Caching improvement	• Focus on caching	• Discard naming scheme evaluation	2017
[110]	• Hierarchical	• Content distribution enhancement	• Focus on caching	• No evaluation	2017
			Mobility		
[111]	• Hierarchical	• Flexible naming scheme • Producer's mobility-aware	• Hierarchical	• No evaluation	2014
[112]	• Hybrid	• Producer mobility support	• Hierarchical	• No evaluation • Additional signaling overhead	2018
			Security		
[113]	• Hierarchical	• Security reinforcement	• CCVPN [114]	• Affect on mobility and latency	2019

Khelifi *et al.* [105] focused on the name lookup process by designing a name-to-hash encoding scheme. The authors tend to reduce the length of the hierarchical name by encoding each component separately into a fixed hash using the CRC32 algorithm and hence reducing the name length. The lookup is performed using heuristic name matching based on the Wu–Manber algorithm [116]. Their solution's performance is based on CRC32 (e.g., encoding, collision, etc.) and the heuristic matching process of the Wu–Manber algorithm. Although the scheme reduces the size of the name, it does not support producer mobility. Moreover, the upstream Interest decoding process may affect retrieval delay.

D. Naming and Scalability

The application namespace is generated by the original content producer or an entity that controls names. Any changes or updates in the application namespace can directly affect the performance of content delivery. Jung *et al.* [108] suggested adding an alias name to the application namespace in NDN to enhance the namespace diversity and management, such as updating, modification, joining, and adding. Since the content name is strongly related to the routing in the hierarchical prefix, any management or update operation in application namespace can affect the function of the routing/forwarding. The alias name has a global view of the application's namespaces. When a consumer requests content, the alias name is used as a translator to the corresponding namespace. The proposed scheme eliminates the barrier in the NDN routing, which can appear in multiproducer and multirouting for the same content in the specific application. At the same time, the namespace management becomes more natural under the alias name controlling, hence improving the application scalability and content retrieval. However, the main problem with this alias is that it is not representative of the content and may not carry any semantics.

E. Naming and Content Caching

The content naming in ICN is very flexible, and the design and convention of its semantic return to the application developer to choose what is suitable for its usage. A name does not include only features related to routing and mobility support, but it can also be used to enhance the caching capabilities. Shan *et al.* [109] proposed extending the naming scheme to comply with caching schemes. The authors proposed to extract a group of preferences related to the content name from the Interest packet. The content item is modified to contain the content's format attribute (e.g., text, video, photo, etc.) and the theme attribute (e.g., news, sport, etc.). The edge node is responsible for extracting the content's name attributes and calculating the matching degree for all received Interests. Then, it makes the caching decision based on the extracted information as well as content popularity. However, and from the naming perspective, some of the content may not fall into a common theme that has not been taken into consideration in this work. Moreover, regardless of the increase of content over time, such classification may raise computation overhead at the edge nodes. Kamiyama and Murata [110] applied a hash function to the content name in order to enhance the content diversity in the network. The authors use different naming features to avoid content redundancy in the network's CS. The core idea is to build a matching between the hash value of the content name, and the candidate node name that wants to cache the content. The mechanism is mainly based on the hashing of the name, in which each node enables caching assigned with binary ID. Therefore, the content name is hashed using a binary function. However, the authors did not explain the hashing process, such as hashing collision and the effect on routing/forwarding.

F. Naming and Mobility

The content in ICN is decoupled from its original provider by using a content name. The name is the basic communication element used by the consumer. From the mobility perspective, two main questions occur as follows.

1) *Consumer Mobility:* The consumer sends an Interest packet to fetch content and then moves to another network. The content is sent back to the previous location of the consumer but not successfully delivered.
2) *Producer Mobility:* The content provider moves from a network to another one. The issue is the network cannot locate the new location of the provider and hence cannot deliver the request.

For ICN architectures that use flat names, the node's mobility performance is directly related to the used resolution service mechanism. Thus, it worth noting that ICN, in such architectures, can handle more than 10^{12} content in the near future [117]. Indeed, it can be seen as a drawback since the nodes' mobility management in such a design is directly related to how swift to handle the object at the resolution services [118], [119]. On the other side, consumer mobility may not impede ICN design that uses hierarchical naming. For instance, NDN and CCN support consumer mobility by design but cannot handle producer mobility issues. Wang *et al.* [120] showed that CCN can handle up to 97% of the requests during the consumer's high mobility. Unfortunately, leveraging a seamless consumers' mobility feature increases the producer mobility drawbacks. For this reason, researchers are trying to find costless methods to make additional support for producer mobility. The desirable method should neither be complicated nor require fundamental changes in the naming prefix or the working principle.

Azgin *et al.* [111] proposed a mobility service targeting hierarchical naming in ICN. The solution aims at reducing the time of content retrieval from a mobile producer. The overall idea is to provide a service that allows any entries in the hierarchical naming to be mobile. The proposed system consists of many principal blocks that interact with each other in order to reduce the Interest flooding and redirect the Interest packet toward the current mobile producer attached point. In addition to the high complexity of their working principle, such an approach requires additional material resources (servers), which increases the deployment cost. Zhang *et al.* [112] introduced trace-based mobility support for NDN, namely, KITE. KITE tends to maintain the structure's ease of the NDN paradigm, where the mobile producer acts as a consumer for a fixed rendezvous server. At each location/network changing, the mobile producer sends an interest to the server in order to keep the trace of its new location. The server is responsible for updating the forwarding rules in the FIB table of intermediate nodes in the network. In doing so, the authors suggest injecting a trace attribute in the routing prefix's name to the content generated by the target mobile producer. For the consumer's side, if it finds the trace of the mobile producer during Interest forwarding, it follows the trace to get the content. The simulation results prove the efficiency of KITE in ensuring seamless producer mobility. However, the architecture suffers from the overhead signaling by the producer, which may create a network disaffection. Moreover, time-sensitive content may not fit well with the proposed solution.

G. Naming and Security and Privacy

ICN adopts a content-based security concept by applying security mechanisms to the content itself and using self-certification, such a technique can improve the security in the network, however, using plain text and knowing the semantics of the name may pose different issues. Leshov *et al.* [113] proposed a naming scheme for the NDN architecture in order to improve the security of content naming in sensitive applications and use cases, such as military applications. The main idea is that the requester uses a symmetric key to first encrypt a part of the content name in the Interest packet. This allows a secure information exchange (the Interest and Data packets) only between trusted nodes that are in the path while hiding the real content name. A three-way handshake mechanism is used to transfer the key between the intermediate nodes in order to decrypt the real content name. The comparison with related schemes records a higher interest satisfaction since the proposed approach supports the in-network caching feature. However, such security reinforcement may affect latency and mobility.

H. Summary and Insights

The performance of any network is mainly based on the performance of naming/addressing lookup. For instance, a DNS service lookup is the first step used to resolve a URL to an IP address. However, DNS is only used with address resolution and does not support additional functionalities (e.g., content placement and movement). In contrast, ICN expanses the naming functionalities in order to promote different additional functions and semantics (e.g., mobility support, security, caching, etc.). Indeed, a few efforts have been made in this area as different aspects were not considered in case of designing a new naming application structure. Meanwhile, some field has not yet explored, such as computation-based naming.

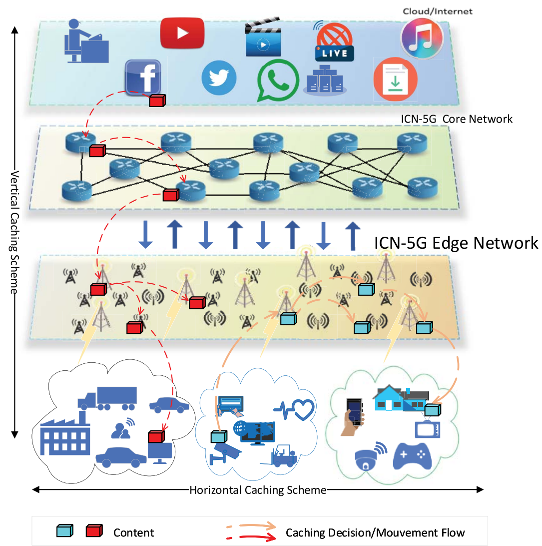

Fig. 9. Content caching in 5G-based ICN networks.

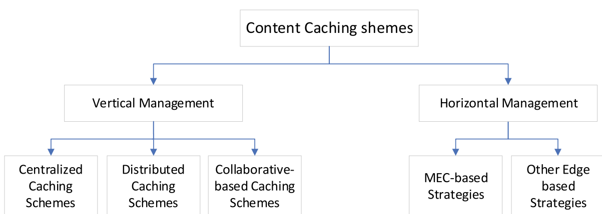

Fig. 10. Taxonomy of existing content caching solutions.

V. Content Caching in 5G-Based ICN Networks

In 5G-based ICN networks, as illustrated in Fig. 9, the network is aware of forwarding the Interest toward any CS that has the requested content [121]. If the content is not cached at the core network, the Interest will be forwarded upstream to the original provider or cloud. During the content delivery phase, each node has the ability to cache the content, and serve it for future demands in a fully transparent way aiming at improving the QoS [122].

In the case of the cache-store runs out of memory, the node evaluates the worthless content in order to keep room for the newly arrived content, this process called the content replacement policy. However, in the literature, these topics have less attention from the research community. The most used content replacement policies are first-in–first-out (FIFO): that evacuates the first entred content to keeping the newcomer one, least recently used (LRU): that attempts to remove the less recently used content, and least frequency used (LFU): that removes the content that has a smaller frequency rate.

The content cache placement is an active research topic where various research efforts attempt to provide seamless and efficient schemes. In this context, we differentiate two main concepts as illustrated in Fig. 10: 1) *vertical management*: tends to select the optimal content cache placement to satisfy a well-defined application scenario (e.g., video streaming) and 2) *horizontal management*: tends to push the popular content toward the edge network to satisfy as many users as possible. In this section, we provide a systematic review of the existing cache placement schemes in the context of the 5G-based ICN networks. Table IV provides a summary of main ICN cache schemes that will be used as a reference in most of the reviewed solutions later.

A. Vertical Management

In the following, we comprehensively review the existing content caching techniques that fit under the vertical management concept for both edge and core networks. Based on the controlled nature that influences the cache decisions, we classify the existing strategies into three main categories: 1) centralized; 2) distributed; and 3) collaborative. Table V provides a summary of these techniques.

1) Centralized Caching Schemes: In centralized caching schemes, a controller entity or a server is used as an imperative network element to decide the cache placement. This entity is responsible for selecting a set of nodes in order to cache a replica version of the content. Centralized schemes provide absolute control over the number of copies of the content, which reduces the cache redundancy and enhances the network [87], [146]. For instance, Li *et al.* [133] proposed a content cache strategy in order to avoid the misutilization of the storage capacity in ICN nodes by decreasing the cache redundancy in the network. The authors designed a probabilistic approach to determine the matching between the content popularity and the node level. The node level is obtained by using different metrics, such as the node's betweenness centrality, node distance, and the node cache capacity. Then, they apply a Gray algorithm [147] to ensure the correlation between the different node's parameters. The proposed scheme is designed to place the content in the appropriate place according to the content popularity and the node level value. Popular content has more probability of being cached at the edge nodes; meanwhile, less-popular content is cached in the core nodes. Although the proposed scheme achieves better efficiency compared to other strategies, the main drawback is the length of the path. In other words, with a larger path with multiple hops, the process will be more complex and require more computation overhead. Moreover, the replacement policy at the node is not considered at all in this work.

Focusing on content distribution, Meng *et al.* [134] proposed a cache strategy that aims at enhancing content distribution in the network while reducing data redundancy. In this strategy, nodes are extended with: 1) a statistical table to collect the frequency demands for each content and 2) a caching mechanism that works as follows: it places all content at a centralized node along the delivery path. The centralized node caches the popular content while storing a copy at the edge node. The popular content is known based on a threshold value (frequency demands). At the same time, less popular content is also kept at the edge node in order to improve the content's availability. The simulation results show the efficiency of the proposed scheme in terms of content diversity, cache hit ratio, and height hops reduction. However, this strategy requires additional information to be stored for each content, which may affect spacial complexity and management.

TABLE IV
MAIN ICN CONTENT CACHING SCHEMES

Ref.	Scheme	Working Principle	Feature	Limitations	Control
[123]	Leave Copy Everywhere (LCE)	All routers cache all passed content	• Reduce server load • Improve retrieval delay • Simple implementation	• Rise cache utilization • High replacement operation • Low diversity ratio	Distributed
[124]	Leave Copy Down (LCD)	Moves the content one hop from the producer	• Reduce server load • Maintain cache utilization • Consider content popularity	• Rise the Interest trip • Increases the retrieval time	Distributed
[125]	Edge Caching(EC)	Only edge nodes cache content	• Decrease latency • Maintain cache utilization • Improve retrieval delay	• Ignore the core enable nodes • Introduce content redundancy	Distributed
[126]	Betweenness	Cache content with maximum betweenness centrality node	• Maintain cache resources • Enhance content delivery	• More overhead on the centralized nodes • Ignore content popularity	Distributed
[127]	PropCache	Decide the cache based on probability value (cache size and path traffic)	• Reduce caching redundancy • Efficient utilization of cache resources	• Ignore content popularity	Centralized
[128]	Prob(α)	Caches the content based on given probability model	• Reduce cache utilization • Reduce replacement operation • Efficient utilization of cache resources	• Ignore content popularity • May keep unpopular content	Centralized
[129]	WAVE	Collaboratively adjusts the content chunk then progressively push them toward the edge	• Consider the chunk level • Increases hit ratio • Consider content popularity	• Introduce inter-traffic overhead • Complex implementation • Rising replacement operation	Collaborative
[130]	Cache Capacity Aware Cache (CCAC)	Estimate the available cache capacity to cache new coming content	• Efficient utilization of cache resources • High cache hit ratio • Cache store awareness	• Ignore content popularity • Low diversity ratio	Distributed
[131]	HighEst Cost Item Caching (HECTIC)	Work at the chunk level, group them based on the chunk-utility then make the caching decision	• Reduce cache utilization • Increases hit ratio	• Algorithm complexity • Low diversity ratio	Centralized
[132]	Random	Cache the content in one or more node in the downstream	• Improve hit ratio • Enhance content delivery	• Unguaranteed efficiency • Ignore content popularity	Centralized

Moreover, the replacement policy is not considered in this strategy. Zheng *et al.* [135] proposed a cache placement strategy, namely, betweenness and edge popularity (BEP) strategy, which is considered as an improvement of the traditional betweenness strategy [126]. The latter selects the node with a higher degree as the appropriate node to cache the content. This decision is justified as most requests pass through the node (e.g., common node in the path). On the other hand, BEP aims at reducing the load on the server by caching the content according to the node's importance and content popularity. The core idea is that the upstream edge nodes compute the content popularity and the node degree. The edge node computes the content popularity with statistic technique taking the old and the current demands of content into account. In the downstream, the BEP algorithm puts the contents according to the node metric and the content popularity. Thus, the popular content will be cached at the highest node degree. The simulation results show that BEP outperforms other strategies in terms of hit ratio and server load reduction. However, by caching the highest requested content at a high degree, more cache hit occurs, and extra overhead is generated, which may diminish the QoS. Moreover, the content replacement policy is not mentioned in this work, which is vital to reduce the overhead on the node. Nguyen *et al.* [136] proposed a progressive popularity-aware caching scheme (PPCS) in order to avoid the cache redundancy of the data cached at the network. PPCS works on the chunk level, which is based on the received demands of content in a period of time. It then decides whether the content is popular or unpopular compared to their access rate with a defined threshold value. For popular content, PPCS caches the first chunk at the edge node then progressively continues to put the rest of chunks at the upstream according to the node level. When the access rate of a chunk at the edge increases over time, PPCS progressively regroups all the chunks of the requested content at the same edge. Besides, the related replacement policy in PPCS tries to remove the least served content obtained from the size of the content over the received demands. The simulation results show that PPCS reduces the load on the server, and outperforms the other strategies in terms of cache utilization and hops ratio. However, in the simulation testbed, the popularity of the content follows Zipf distribution [148]. Zipf represents the access demands for the Web content; it is mentioned that the chunk level can follow the Zipf distribution except that the chunks delivered in the sequential order [149], which is not the case for all content in real-world networks. Moreover, the

TABLE V
SUMMARY OF VERTICAL CONTENT CACHING TECHNIQUES

Ref.	Feature	Comparison	Evaluation	Limitation	Year
			Centralized Caching Schemes		
[133]	• Consider chunk level • Cache resources improvement • Network delay reduction	• LCE, Betweenness, Prob(α)	• Content diversity improvement	• No replacement policy • Computation overhead	2016
[134]	• Improve caching performances • Reduce content redundancy	• ProbCache, CCAC, LCD	• Increase content diversity ratio	• No replacement policy • Considerable network traffic	2019
[135]	• Consider chunk level • Memory resources improvement	• Betweenness • LCD, ProbCache, Probability	• Reduce server load • Cache hit improvement	• No replacement policy • Central node overhead	2019
[136]	• Consider the chunk level • Enhance memory usage • Reduce server load	• WAVE, LCD, ProbCache, LCE • HECTIC	• Latency improvement • Cache hit rate improvement	• No replacement policy • Load on central node	2019
			Distributed Caching Strategies		
[137]	• Bandwidth saving • Reduce power consumption	• LCE, Prob(α), ProbCache	• Reduce bandwidth consumption • Cache hit rate improvement	• Low content diversity ratio	2014
[138]	• Considering missing Interest	• LCE, LCD, ProbCache, Betweenness, WAVE	• Latency improvement • Cache hit rate Improvement	• Algorithm complexity • Extra computation overhead	2017
[139]	• Reduce retrieval delay • Demand prediction	• LCE, LCD, Prob(α)	• Cache hit improvement • Reduce bandwidth consumption	• Consider only local popularity • High content redundancy ratio	2018
[140]	• Improve cache efficiency • Enhance content diversity	• LCE, LCD, EC, ProbCache, WAVE	• Content retrieval enhancement	• Assume popularity known • Unsuitable to multi-provider	2019
			Collaborative Caching Schemes		
[141]	• Improve content diversity • Reduce retrieval delay • Enhance latency	• LCE, ProbCache, Random, LCD	• Cache utilization enhancement • Content retrieval improvement	• Unsuitable to a large scale network • High memory utilization	2016
[142]	• Reduces retrieval delay • Enhance latency	• LCE, ProbCache, Random	• Few cache hit improvement	• Rise data replication • Increase traffic	2017
[143]	• Server load reducing • Congestion Reducing	• LCE, LCD, MCD, MAGIC, ProbCache	• Outperform in term of replacement operation and cache utilization	• Ignore packet processing time	2018
[144]	• consider Bandwidth as metric • Reducing content redundancy	• LCE, LCD, ProbCache	• A few improvements	• Assume popularity known	2018
[145]	• Enhance video streaming delivery	• MPC, LCE, ProbCache	• Reduce bandwidth consumption • Reduce the network delay	• High traffic overhead	2019

exchanged messages between nodes may create extra overhead and congestion.

2) Distributed Caching Schemes: In a distributed caching scheme, there is no centralized entity to control the caching decision. Nodes use local factors and metrics to decide the content caching, which improves network efficiency by reducing the exchanged information between nodes and/or centralized entities.

Ren *et al.* [137] proposed a distributed caching strategy, namely, distributed max-gain (MAGIC). MAGIC aims at reducing bandwidth consumption and avoiding unnecessary content replacement. In this scheme, both Interest and Data packets are extended with gain attributes. During interest forwarding, each node computes the local gain value and then updates the value in the Interest packets only if it is bigger than the existing value in the Interest itself. The local gain is obtained by involving existing local information, such as the demands rate, hop reduction, replacement penalty, etc. In the data delivery process, when a node receives a data packet, it caches the content only if the local gain equals to the maximum gain in the Data packet. The simulation results show that MAGIC has more improvements compared to related schemes in terms of hops reduction ratio, cache utilization, and caching operation. However, the design of this scheme is limited to one replica node selection along the delivery path, which is not recommended in large-scale networks with high content's diversity, besides some content must be replicated to meet the massive demands from users. Banerjee *et al.* [138] designed a greedy cache placement algorithm in order to decrease content delivery latency. The designed strategy runs in a distributed way where each node caches the content based on its popularity. The greedy algorithm constructs a directed acyclic graph by leveraging information from the nodes' routing table. Then, it considers the missed demands on the upstream for each router and extracts the content popularity by computing the missed demands (along the communication path toward the producer). The authors compared their algorithm with other strategies in diverse network topologies and multiple scenarios, including multiproducers. The obtained results show the efficiency of the proposed strategy in terms of cache utilization, cache hit ratio, and hop reduction. However, the latency of the lookup of content in the nodes is not considered. Moreover, the generated graph may take more computational power, especially in the case where the content is far away from the consumers, which may impact the speed of cache decision making. Similarly, Wu *et al.* [139] designed a distributed probabilistic caching strategy in ICN that aims at enhancing the efficiency of in-network content caching. The authors integrated local information to decide the content caching. Each node in the network takes individual decision to either cache the content or not using a probabilistic formula. The probabilistic formula combines the content popularity and the benefit metric, and generates a probabilistic value. The benefit metric refers to the hop reduction between the requester and the cache store's node or the original server. The authors also proposed a content popularity prediction technique that estimates future content's popularity based on the stored historical demands locally at the node. The authors validated their solution in a testbed using diverse network topologies. Results show the proposed solution is better than other strategies by caching popular content. However, the popularity prediction technique is based only on local popularity without considering the network perspective. This may influence the replication decision in other nodes for similar popular content. Moreover, storing historical information (for popularity) for each content in the node excessively affects memory usage and overhead the node.

Kamiyama *et al.* [110], [140] proposed a distributed caching strategy, namely, spatially dispersed caching (SDC), where each node makes an autonomous cache decision. SDC aims at avoiding content duplication among nodes in the same area and caching only the highest popular content. SDC works in four main steps: 1) each router is assigned a single binary ID of X bit, while nearby routers have different values in the higher bit; 2) setting a hash value for each content; 3) if the hash value of the content matches with the router ID, this content is acceptable to the next phase; and 4) based on the content popularity the router classifies the accepted content into multiple classes and then caches the popular one. By applying all these steps, SDC aims at guaranteeing the content's spread over the network. Although the idea is similar to [150], the key addition is in the caching decision. Besides, for each acceptable content, the SDC node classifies them into $X+2$ groups based on the content demands rate, hence, the router differentiates the number of nodes that probability can cache the content based on group popularity then changing the number of bits checked if the node can host the content. The simulation results show that SDC marks a considerable improvement in terms of cache utilization and hop reduction compared with other strategies. However, in the case table of group popularity inputs raising over time, it dramatically affects the node memory as well as the time of the caching decision. The authors do not consider this limit that can degrade the effectiveness of the node retroactively in the real environment with massive content demands. Moreover, the question remains open on its efficiency in the case of multiprovider scenarios.

3) Collaborative-Based Caching Schemes: Nodes in the communication path may build cooperation between each other and exchange information in order to explore more capabilities and better use the network resource. In the following, we review the existing caching schemes using this mechanism. Mick *et al.* [141] proposed a collaborative multihop caching strategy based on a bloom filter technique that aims at enhancing content diversity and latency while reducing intercommunication overhead between collaborative nodes. In this strategy, each node assembles the cache information between n hops reached nodes (neighborhood); then, the caching information is collected and stored in a set of attenuated bloom-filters. Whenever a request reaches a node, the node first checks the stored information regarding the neighborhood. If the content does not exist, it forwards the Interest toward the producer. In the meantime, to achieve content diversity and reduce content redundancy, the eviction policy uses the FIB in order to avoid the caching content in near neighbors. The simulation results show the efficiency of the proposed scheme in terms of the cache hit ratio and latency. However, in large-scale networks, the probability of getting a false positive of the bloom filter technique is high, which may be considered the major drawback of such an approach. Moreover, in the simulation setup, the cache utilization is ignored. Huang *et al.* [142] designed a collaborative cache strategy where nodes in the delivery path collaborate to maximize the use of the network-embedded caching. The authors focused on reducing the cost of traffic by considering both the received demands and the content size. They proposed a payoff solution function least caching utility (LCU), in order to solve the traffic minimization problem. LCU is used as a content replacement policy along the delivery path, which tends to evict the content with the least utility value taking the demand rates and distance between the requester and the original content provider into account. The node exchanges information about the content [time since birth (TSB)] by injecting it into the Data packet. This mechanism improves the collaboration mechanism between the node while reducing the extra overhead of the exchanged messages. Simulation results show the benefits of the proposed solution in terms of low communication overhead (i.e., exchange messages between nodes). However, such a policy may introduce a high redundancy of content caching, especially in large-scale networks. Moreover, with the expectation of increasing the content in the network, checking the TSB for each passed data packet affects the router's processing resource and adds extra delay.

Zhi *et al.* [143] proposed the gain-aware 2-round cooperative (GAC) caching scheme to enhance the content diversity while reducing the data redundancy in the network. In GAC, the node decides the content caching based on the decisions made by other nodes along with the path delivery. GAC's core idea is that the upstream nodes compute the potential gain value and include it in the Interest packet with their node ID. The potential gain value takes the number of received demands and the number of hops reduction if the node caches the content into consideration. In the data delivery path, the node computes the final gain and takes the cache decision accordingly. If the gain value is higher than a predefined threshold, the node updates the hop value to 0 in the Data packet. The hop value describes the number of hops needed to get the content from the replica node or the original producer. The simulation results show the efficiency of the proposed scheme in terms of the cache hit ratio, hop reduction ratio, and cache utilization. However, since the cache decision must be line speed in ICN routers [128], the packet processing for each content may introduce a considerable effect on network delay, especially in the case of multiproducer and massive contents requesting. Moreover, adding the node's ID to each packet is not feasible and may need changes in the nature of ICN packet. Furthermore, defining an ID for each node is a challenging task that needs more investigation to ensure a unique global ID, it also takes the paradigm back to host-centric networks. Liu *et al.* [144] propose a collaborative cache on demand (COD) strategy. COD tends to reduce the number of nodes that cache the content in the network. COD operates based on the number of received demands and the temporal patterns of content popularity. According to the study in [151], the authors relied on the change of content popularity by applying two stages: 1) the early stage, when requesting content in the initiator until up to the peak in popularity and 2) the late stage, the change and degradation of the popularity of the content. Depending on a probabilistic formula, the node chooses to cache the content or push it to the adjacent nodes with more cache size. COD aims to keep the content in the early state close to the producer, and over time, depending on the received demands, change the content to the late stage. The authors provide several measures to show the efficiency of their solution in terms of cache consumption, energy consumption, and content diversity. However, this work is based on the popularity of the content that needs to be known beforehand. The scheme is not applied to dynamic content or content generated based on demands. Besides, each node is obligated to compute the content state periodically, which produces computation overhead. Moreover, the user's mobility may create a major drawback of such a strategy since the demands are not stable. Noh *et al.* [145] proposed a caching strategy for video streaming. The main objective is to provide a seamless video streaming service. First, the authors designed an algorithm to create the metafile for each content that adapts with a multicaching node environment in ICN. The metafile regroups information about content chunks as the priority and the bandwidth requirement. In this strategy, each node selects its range of chunks to cache using the associated content metafile, therefore avoid the redundant caching of the same chunk in the adjacent nodes. Nodes exchange information about cached content using modified Interest/Data packets. At the reception of this information, the proposed algorithm progressively adjusts the range according to chunks' priority hop number and demands rate. The obtained results show that the proposed solution provides seamless video streaming by reducing network stress. Yet, this efficiency is at the expense of the cache utilization, which not considers in this strategy. Moreover, creating metafiles and exchanging information may produce more overhead on the network, especially in profoundly changing content's popularity.

4) Summary and Insights: A centralized caching scheme can manage the global content in the network, which can be installed in a centralized controller, such as an SDN controller. Although it provides better memory resource utilization, the global control of the content replicas is a challenging task and may consume extra resources. It also affects the control entity by exhausting the computation resources, which can reflect network efficiency, especially on the traffic peak-hours. On the other hand, a distributed caching strategy can strongly be adapted to a real network that is distributed in nature. It does not require knowledge of global topology nor the exchanged messages. Thus, it can easily enhance the network scalability. However, the major drawback that nodes are involved in such strategies, focus only on the local information to make the decision, which may complicate the task by raising the contents' redundancy. Finally, a collaborative caching scheme provides a better exploration of the collaborated nodes' memory resources while splitting the burden of decision making. It can improve content diversity in the network. However, exchanged message costs can affect the traffic load, especially in large-scale networks. In short, we believe that designing a distributed scheme with global resources knowledge based on the SDN technology seems to be an efficient solution.

B. Horizontal Content Caching Management

In this section, we explore the management techniques that take advantage of its location (closer to the consumer) to exploit the benefits of the powerful interaction between the edge-enabled caching node and the consumers. Table VI provides a summary of these techniques.

1) MEC-Based Strategies: Several edge caching strategies have been designed that use MEC servers and exploit their robust computation to provide better content distribution at the edge network and enhance QoS for the end users. In this part, we review the recent existing works based on MEC founded in the literature.

Zhang *et al.* [46] proposed an edge caching architecture for 5G networks that combines the caching and computing resources of MEC and ICN-based edge nodes, including a BS, SBS, and user nodes. The proposed architecture aims at enhancing the cache performances within the global edge network. The main idea is to achieve a tradeoff between computing cost and caching gains. In doing so, the authors used MEC computing resources in order to compress the content. This will improve the caching capability at the edge and, by consequence, improve the content availability. The simulation results show a significant improvement in both the

TABLE VI
SUMMARY OF HORIZONTAL CONTENT CACHING STRATEGIES

Ref.	Feature	Comparison	Evaluation	Limitation	Year
			MEC Based Strategies		
[46]	• Improve caching capacity • Enhance content availability	• Cache without MEC	• Latency improvement • Cache hit rate enhancement	• Compression overhead • Computation expenses	2018
[152]	• Context integration • User-behavior aware	• PoPCache, Greedy Algorithm, MPC	• Cache hit improvement • Considerable running time	• Users' privacy • High complexity	2019
[153]	• Recommended system integration	• Random, LCE, SVD Based	• High retrieval delay • Cache hit Ratio enhancement	• Focus only on video contents • High computation overhead	2019
[154]	• User location aware • Consider users' mobility • Content prefetching	• FiFo, LRU, LFU	• Cache hit ratio improvement • Reduce access time	• Users' privacy	2019
[155]	• Enhance MEC resources • Improve user energy	/	/	• Ignore users' behaviors • Lack for the practical part	2019
[156]	• Improve video streaming service	• Non-MEC based	• Data retrieval improvement	• Ignore wireless packets loss	2019
			Other Edge Based Strategies		
[109]	• User preferences aware • Leveragr content naming	• EC, Prob(α)	• Efficient cache hit ratio • Latency improvement	• Edge node overhead • Naming schemes not evaluated	2017
[157]	• Include social network knowledge	• Random caching	• Considerable caching enhancement	• Ignore users behavior • Extra computation overhead	2019
[158]	• Employ deep learning	• EC with different replacement strategies	• Cache hit ratio enhancement	• Learning process	2019

cache hit and the round trip latency. However, the major drawback is the high complexity of scheduling of all resources simultaneously, especially when takes the nodes behaviors as mobility into consideration. Moreover, some nodes are resource-constraints (e.g., IoT devices), hence using them as CSs may affect their performances while ignoring them may lead to overhead computation. Huang *et al.* [159] proposed an online learning framework (OLCB) that aims to reduce the backhaul bandwidth consumption by enhancing the edge nodes' cache efficiency. The proposed strategy initially works to construct the context by capturing the user group access behavior, including connecting period and mobility trace. The exploitation–exploitation machine learning model is adopted to create a relationship between the requested content and their specific context under a distinct time period and estimate the reward of caching the content. The problem is modeled as a Knapsack problem, where each requested content is associated with a specific context that has a particular reward. The objective is to maximize the reward value under specific node storage memory. In another work, Zeng *et al.* [152] proposed an algorithm to solve the knapsack problem leveraging branch and bound approach. A smart caching a heuristic algorithm is proposed that applies the mining on the central effect of consumer behavior in order to extract the consumer content preference in each time slot. Based on the user's content preference, the proposed scheme makes real-time learning-based caching decisions by predicting the variety of content popularity that can share content preferences under specific contexts. The authors simulated the scheme using a real data set of user records to show the efficiency of the cache hit rate. However, the differences in demands for diverse content is questionable. The solution needs more investigation to deal with context management which needs more computation. Moreover, collecting data for learning purposes is challenging due to the users' privacy issues.

Tang *et al.* [154] designed a cache strategy scheme for mobile multimedia devices, namely, edge IoT equipment-assisted caching multimedia (EACM). EACM tends to improve the QoE. The overall idea is to cache the content at the edge node (e.g., SBS), taking the importance of the requested content of the users' group and the users' mobility into consideration. The proposed strategy benefits from the computation capability of the edge by using the recurrent neural network (RNN) to predict the next localization through learning the users' trace. At the same time, the FB-growth algorithm [160] is adopted to extract the common users' group interest. This coordination ensures to prefetching the common content from the remote server and guarantees their distribution along with the edge nodes, according to the users' needs in a specific area. Furthermore, a hybrid LRU and LFU policy are designed to combine and manage respective features, both of them in the double queueing and handle cache-store decisions. The simulation results show that EACM diminishes the load on the server side, while the replacement policy outperforms other policies, reduces the access time, and optimizes the cache hit ratio. However, such an approach may create redundancy for the popular content in the edge nodes, which is not taken into account. Moreover, refetching the most requested content for each area may create overhead demands on the remote server for the same content, which increases the bandwidth consumption. Finally, any changes in the network setting can affect directly on RNN model's precision [153].

Tang [155] tended to enhance the memory of the MEC using the ICN paradigm. In doing so, the authors assumed that each edge node has specific caching resources. The objective is selecting the nearest BS caching node to the MEC. The chosen node is based on consumer proximity to the BS, while the global objective is diminishing the user side's energy consumption. However, the users' behavior (e.g., mobility) has not been considered. Moreover, the proposed solution needs to tradeoff between the content popularity and the memory resource selected by the MEC, which needs more investigation toward real implementation since the authors present only the theory part. Han *et al.* [156] aimed at improving video streaming delivery in the MEC-based ICN paradigm by enhancing the forward error correction (FEC) mechanism of the ICN architecture. The authors proposed a framework running on top of MEC platform, in which FEC generates the packet loos at the network edge instead of the remote server. A deep learning algorithm is then used to learn the received information from FEC and use the FEC rate to monitor network states (by deciding whether to correct the packet error or use the retransmission). Compared with conventional ICN (non-MEC based), the simulation results prove the efficiency of the proposed. However, this framework treats only the error transmission between MEC and the original producer, where some packets may get lost due to the interference.

2) Non-MEC-Based Strategies: In this section, we present an edge caching solution that does not consider the MEC server as a single caching management entity. These works assume that each edge node (BS, SBS, user equipment, etc.) can use the computation and decision.

Shan *et al.* [109] proposed a cache strategy based on the users' group interest preference. The proposed strategy aims at reducing the network delay by caching the most preferred contents in the edge node. The authors used a classification model for the contents under the assumption that the content name contains two attributes theme (e.g., sport and news) and a type (e.g., text and video). The main idea is to ensure the matching between the preferred interest and response data. The edge node extracts the user's group interest from users' requests. When data reaches the edge, it decides whether to cache the content or not according to defined threshold judgment. The simulation results show the efficiency of the proposed strategy in terms of cache utilization and hop reduction. However, this strategy uses only the edge node as the main node to compute, store, and deliver content, which may lead, as a consequence, to overhead and degrade the QoE, especially in the case of highly mobile users. Moreover, due to the less storage memory at the edge node, the computation and communication may congest the backhaul links. Liu *et al.* [157] proposed an analytical framework based on the ICN edge caching scheme. The proposed framework aims at putting the commonly requested content at a nearby location of the users. In the first stage, the framework records the user check-in to get the connected edge node. The authors then leveraged a location-based social network technique that considers: the distance between users, the user interest, and the contact rate. Since the problem is an NP-hard, a metaheuristic simulated annealing algorithm [161] has been adopted to get the near-optimal solution. Compared with other solutions, the proposed framework shows high performance in terms of content propagation ratio. However, this framework needs extra information about users. Moreover, this work considers only the users' common interests and ignore users' behaviors. Wang *et al.* [158] proposed a caching scheme that tends to be adapted to the network environment. In doing so, a deep reinforcement learning algorithm has been presented based on an asynchronous actor–critic agent mechanism. A learning agent is evolved (using the reward or punishment) during interactions with the network environment. Hence, the framework decides whether to keep the requested content at the edge node or not. It also uses the substitution strategy in case the cache is full. The decision is linked to learned network status. The simulation results show the high performance of the proposed strategy, notably in terms of improving the cache hit ratio. However, the results are strongly related to the capacity of the user data in the learning process. Therefore, the relearning of the model at each network may appear as a challenge.

3) Summary and Insights: Horizontal-based caching strategies at the edge are more resilient and flexible as the cache hit is near to users. The content's usefulness is strongly associated with how close it is to the consumer. Therefore, bring content closer to the users comes with many profits, such as reducing retrieval time, decreasing the backhaul traffic, etc. Exploiting the necessary methods to achieve this purpose is the aim of most horizontal-based approaches. Their decisions have a strong dependency on the correlation between the consumer/interests and its behavior. Using advanced techniques to achieve a better user/behavior adaptation may be paved the road to the intelligent edge paradigm. The machine learning approaches are taken place near to the consumer at the network edge and on the top of a powerful computation platform as MEC [162]. However, several issues raised, including the choice of the appropriate model, the learning process, and content privacy.

VI. Issues and Future Research Directions

Although the efforts that have been made in the 5G-based ICN naming and caching aspects, several challenges and issues are open for more investigations. In this section, we describe some challenges that faced 5G-based ICN architecture.

A. Naming-Related Issues

Name-Based Routing in Mobile Networks: User mobility is an essential aspect of the next 5G networks, especially for time-sensitive applications. Although ICN has native support for user mobility, which may provide a seamless user movement, the frequency change of the user attachment point between multiple access technologies may impact the QoE as well as QoS, which leads to service degradation. Furthermore, the main drawback appears if the same user act as a producer (that produces a certain content). In this case, the routing-based naming process uses a flooding Interest for path recovery. Thus, it increases the network traffic, which leads to the network conjunction. In this regard, various works have been done concerning producer mobility issue. However, none of them uses a flexible naming scheme that supports multiple features, e.g., mobility support, security enhancement, scalability, etc.

Naming and Real-Time Applications: Content naming provides content identification and a routing mechanism, which leads to improving content retrieval and enhancing the security level. However, ICN naming may have several issues related to the retrieval of real-time applications, including multimedia that requested by millions of users every day. Potential solutions, such as enhancing the naming with attribute-value-based names can help to solve this issue.

Latency Improvements: The 5G-based ICN networks tend to support multiple applications use cases. From this perspective, a few efforts consider improving the naming schemes targeting specifics 5G applications and use cases. For instance, some IoT actuators and sensors probably request content with a larger name length, which affects the routing tables and lookup performance. Therefore, a tradeoff between naming schemes and the data size should be considered. Applying encoding schemes can be a potential solution without losing name semantic or structure.

B. Content Caching-Related Issues

Caching Improvements With Users' Behavior Support: To perform adequate high-quality horizontal management, matching user preferences, behaviors, and context is indispensable to ensure a reliable caching decision. Thus, system complexity could be very high. Deep reinforcement learning is an advanced machine learning technique that is considered a promising technology to handle a large amount of input data. Such a solution may help to provide distributed and scalable caching schemes.

Popular Content Caching: Caching the frequently requested content (i.e., popular content) without considering additional parameters (e.g., content size) leads to retention and accumulation of small-sized content. Consequently, it impacts the global network caching performance, reduces caching capacity, overloads the lookup process, and adds extra weights to the CS. Deep learning solutions can help to predict content popularity by combing multiple attributes.

Large-Scale Caching-Related Issues: Although several works have been done in vertical caching schemes, further studies should be investigated regarding which the appropriate control caching manner. A qualitative and quantitative comparison could be made in order to clarify the matter.

Service Versus Data Caching: The ICN-assist MEC approach can help to enhance content caching and computation at the same level. A new research question arises regarding the feasibility of cache computation of previously executed services and not only content. This will help to avoid multiple computations and, in result, enhance end-to-end delay. The computation reuse paradigm can be applied at the MEC server without violating ICN primitives.

VII. Conclusion and Summary

The 5G-based ICN network will bring new Internet usage and vision based on a world built on fast communication. 5G can be realized based on multiple new technologies, while ICN helps in realizing a clean and efficient next-generation network. In this article, we provided a comprehensive survey of different enabling technologies that come into this domain. We first highlighted different pieces related to the next generation networks. Then, we argued the necessity and feasibility of the ICN paradigm within the 5G network. We also provided an in-depth survey of the basic component of the ICN architecture, including the naming efforts and the caching strategies. We reviewed recent and essential solutions that have been classified into different categories based on their working principle. Finally, we addressed some challenges and described promising future research directions. We believe that ICN will provide an open view of the 5G implementation by helping to realize the 5G missions and visions.

References

[1] A. V. Bhalla and M. R. Bhalla, "Generations of mobile wireless technology: A survey," *Int. J. Comput. Appl.*, vol. 5, no. 4, pp. 26–32, 2010.

[2] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[3] N. Alliance, "5G white paper," Next Generation Mobile Networks, Frankfurt am Main, Germany, White Paper, 2015.

[4] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.

[5] F. Messaoudi, P. Bertin, and A. Ksentini, "Towards the quest for 5G network slicing," in *Proc. IEEE CCNC Conf.*, 2020, pp. 1–7.

[6] S. Kekki *et al.*, "MEC in 5G networks," ETSI, Sophia Antipolis, France, White Paper, 2018.

[7] D. S. Rana, S. A. Dhondiyal, and S. K. Chamoli, "Software defined networking (SDN) challenges, issues and solution," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 1, pp. 884–889, 2019.

[8] I. Farris, T. Taleb, Y. Khettab, and J. Song, "A survey on emerging SDN and NFV security mechanisms for IoT systems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 812–837, 1st Quart., 2019.

[9] B. Nour, K. Sharif, F. Li, S. Yang, H. Moungla, and Y. Wang, "ICN publisher-subscriber models: Challenges and group-based communication," *IEEE Netw.*, vol. 33, no. 6, pp. 156–163, Nov./Dec. 2019.

[10] G. Xylomenos *et al.*, "A survey of information-centric networking research," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 1024–1049, 2nd Quart., 2014.

[11] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, "A survey of information-centric networking," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 26–36, Jul. 2012.

[12] H. Khelifi *et al.*, "Named data networking in vehicular ad hoc networks: State-of-the-art and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 320–351, 1st Quart., 2020.

[13] R. Ravindran *et al.*, "Deploying ICN in 3GPP's 5G nextgen core architecture," in *Proc. IEEE 5G World Forum (5GWF)*, 2018, pp. 26–32.

[14] C. Liang, F. R. Yu, and X. Zhang, "Information-centric network function virtualization over 5G mobile wireless networks," *IEEE Netw.*, vol. 29, no. 3, pp. 68–74, May/Jun. 2015.

[15] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[16] M. R. Palattella *et al.*, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Mar. 2016.

[17] Y. Mao *et al.*, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017.
[18] C. Wang, Y. He, F. R. Yu, Q. Chen, and L. Tang, "Integration of networking, caching, and computing in wireless systems: A survey, some research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 7–38, 1st Quart., 2018.
[19] I. Ahmad, T. Kumar, M. Liyanage, J. Okwuibe, M. Ylianttila, and A. Gurtov, "Overview of 5G security challenges and solutions," *IEEE Commun. Stand. Mag.*, vol. 2, no. 1, pp. 36–43, Mar. 2018.
[20] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5G: RAN, core network and caching solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3098–3130, 4th Quart., 2018.
[21] X. Zhang and Q. Zhu, "Information-centric virtualization for software-defined statistical QoS provisioning over 5G multimedia big data wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 8, pp. 1721–1738, Aug. 2019.
[22] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Communications, caching, and computing for next generation HetNets," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 104–111, Aug. 2018.
[23] I. Benkacem *et al.*, "Integrated ICN and CDN slice as a service," in *Proc. IEEE GLOBECOM Conf.*, 2018, pp. 1–7.
[24] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network slicing and softwarization: A survey on principles, enabling technologies, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2429–2453, 3rd Quart., 2018.
[25] "M. Series; IMT vision—Framework and overall objectives of the future development of IMT for 2020 and beyond," Int. Telecommun. Union, Geneva, Switzerland, ITU-Recommendation M.2083-0, 2015.
[26] Z. Li *et al.*, "5G URLLC: Design challenges and system concepts," in *Proc. ISWCS Conf.*, 2018, pp. 1–6.
[27] M. Patzold, "5G is coming around the corner [mobile radio]," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 4–10, Mar. 2019.
[28] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[29] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.
[30] (2020). *Forecast Number of 5G Subscriptions Worldwide From 2020 to 2025*. [Online]. Available: https://www.statista.com/
[31] X. You *et al.*, "The 5G mobile communication: The development trends and its emerging key techniques," *Scientia Sinica Informationis*, vol. 44, no. 5, pp. 551–563, 2014.
[32] "Study on scenarios and requirements for next generation access technologies," 3GPP, Sophia Antipolis, France, Rep. 38.913, 2016.
[33] A. Jain *et al.*, "Are mobility management solutions ready for 5G?" 2019. [Online]. Available: arXiv:1902.02679.
[34] B. Nour *et al.*, "A blockchain-based network slice broker for 5G services," *IEEE Netw. Lett.*, vol. 1, no. 3, pp. 99–102, Sep. 2019.
[35] K. Campbell *et al.*, "The 5G economy: How 5G technology will contribute to the global economy," in *IHS Economics and IHS Technology*. San Diego, CA, USA: Qualcomm Technol., 2017.
[36] Z. Piao, M. Peng, Y. Liu, and M. Daneshmand, "Recent advances of edge cache in radio access networks for Internet of Things: Techniques, performances, and challenges," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 1010–1028, Feb. 2019.
[37] B. P. Rimal, D. P. Van, and M. Maier, "Mobile edge computing empowered fiber-wireless access networks in the 5G era," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 192–200, Feb. 2017.
[38] B. Omoniwa, R. Hussain, M. A. Javed, S. H. Bouk, and S. A. Malik, "Fog/edge computing-based IoT (FECIoT): Architecture, applications, and research issues," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4118–4149, Jun. 2019.
[39] J. Pan and J. McElhannon, "Future edge cloud and edge computing for Internet of things applications," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 439–449, Feb. 2018.
[40] B. Varghese and R. Buyya, "Next generation cloud computing: New trends and research directions," *Future Gener. Comput. Syst.*, vol. 79, pp. 849–861, Feb. 2018.
[41] K. Zhang *et al.*, "Predictive offloading in cloud-driven vehicles: Using mobile-edge computing for a promising network paradigm," *IEEE Veh. Technol. Mag.*, vol. 12, no. 2, pp. 36–44, Apr. 2017.
[42] B. Nour, S. Mastorakis, and A. Mtibaa, "Compute-less networking: Perspectives, challenges, and opportunities," *IEEE Netw.*, early access, Aug. 12, 2020, doi: 10.1109/MNET.011.2000180.
[43] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervasive Comput.*, vol. 8, no. 4, pp. 14–23, Oct.–Dec. 2009.
[44] Y. C. Hu *et al.*, "Mobile edge computing–A key technology towards 5G," ETSI, Sophia Antipolis, France, White Paper, 2015.
[45] J. Dolezal, Z. Becvar, and T. Zeman, "Performance evaluation of computation offloading from mobile device to the edge of mobile network," in *Proc. IEEE CSCN Conf.*, 2016, pp. 1–7.
[46] K. Zhang, S. Leng, Y. He, S. Maharjan, and Y. Zhang, "Cooperative content caching in 5G networks with mobile edge computing," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 80–87, Jun. 2018.
[47] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, "A survey on software-defined networking," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 27–51, 1st Quart., 2015.
[48] I. Alam *et al.*, "A survey of network virtualization techniques for Internet of Things using SDN and NFV," *ACM Comput. Surveys*, vol. 53, no. 2, p. 35, Apr. 2020.
[49] "Open networking foundation: Software-defined networking: The new norm for networks," ONF, Palo Alto, CA, USA, White Paper, 2012.
[50] F. Hu, Q. Hao, and K. Bao, "A survey on software-defined network and openflow: From concept to implementation," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2181–2206, 4th Quart., 2014.
[51] H. Kim and N. Feamster, "Improving network management with software defined networking," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 114–119, Feb. 2013.
[52] N. McKeown *et al.*, "OpenFlow: Enabling innovation in campus networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69–74, 2008.
[53] T. Koponen *et al.*, "Onix: A distributed control platform for large-scale production networks," in *Proc. OSDI*, vol. 10, 2010, pp. 1–6.
[54] K. Phemius, M. Bouet, and J. Leguay, "DISCO: Distributed multi-domain SDN controllers," in *Proc. IEEE Netw. Oper. Manag. Symp. (NOMS)*, 2014, pp. 1–4.
[55] A. Dixit, F. Hao, S. Mukherjee, T. V. Lakshman, and R. R. Kompella, "ElastiCon; An elastic distributed SDN controller," in *Proc. ACM/IEEE Symp. Archit. Netw. Commun. Syst. (ANCS)*, 2014, pp. 17–27.
[56] P. Berde *et al.*, "ONOS: Towards an open, distributed SDN OS," in *Proc. Workshop Hot Topics Softw. Defined Netw.*, 2014, pp. 1–6.
[57] J. Sherry *et al.*, "Making middleboxes someone else's problem: Network processing as a cloud service," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 13–24, 2012.
[58] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, "5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges," *Comput. Netw.*, vol. 167, Feb. 2020, Art. no. 106984.
[59] M. Chiosi *et al.*, "Network functions virtualisation: An introduction, benefits, enablers, challenges and call for action," in *Proc. SDN OpenFlow World Congr.*, vol. 48, 2012, p. 202.
[60] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5G," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 84–91, Apr. 2016.
[61] L. Laaziz, N. Kara, R. Rabipour, C. Edstrom, and Y. Lemieux, "FASTSCALE: A fast and scalable evolutionary algorithm for the joint placement and chaining of virtualized services," *J. Netw. Comput. Appl.*, vol. 148, Dec. 2019, Art. no. 102429.
[62] ONF Solution Brief, "OpenFlow-enabled SDN and network functions virtualization," *Open Netw. Found.*, vol. 17, pp. 1–12, 2014.
[63] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "NFV and SDN—Key technology enablers for 5G networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2468–2478, Nov. 2017.
[64] F. Zhang, G. Liu, X. Fu, and R. Yahyapour, "A survey on virtual machine migration: Challenges, techniques, and open issues," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1206–1243, 2nd Quart., 2018.
[65] B. Nour *et al.*, "A survey of Internet of Things communication using ICN: A use case perspective," *Comput. Commun.*, vols. 142–143, pp. 95–123, Jun. 2019.
[66] M. F. Bari, S. R. Chowdhury, R. Ahmed, R. Boutaba, and B. Mathieu, "A survey of naming and routing in information-centric networks," *IEEE Commun. Mag.*, vol. 50, no. 12, pp. 44–53, Dec. 2012.
[67] J. Pan, S. Paul, and R. Jain, "A survey of the research on future Internet architectures," *IEEE Commun. Mag.*, vol. 49, no. 7, pp. 26–36, Jul. 2011.
[68] T. Koponen *et al.*, "A Data-oriented (and beyond) network architecture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 4, pp. 181–192, 2007.

[69] G. García *et al.*, "COMET: Content mediator architecture for content-aware networks," in *Proc. Future Netw. Mobile Summit (FutureNetw)*, 2011, pp. 1–8.

[70] N. B. Melazzi *et al.*, "Publish/subscribe over information centric networks: A Standardized approach in CONVERGENCE," in *Proc. Future Netw. Mobile Summit*, 2012, pp. 1–8.

[71] I. Seskar *et al.*, "Mobilityfirst: Future Internet architecture project," in *Proc. Asian Internet Eng. Conf.*, 2011, pp. 1–3.

[72] M. Brunner, "Scalable & Adaptive Internet solutions (SAIL)," in *Proc. Future Internet Assembly*, 2010, pp. 1–2.

[73] N. Fotiou *et al.*, "Developing information networking further: From PSIRP to PURSUIT," in *Proc. EAI BROADNETS Conf.*, 2010, pp. 1–13.

[74] T. Anderson *et al.*, "The nebula future Internet architecture," in *Proc. Future Internet Assembly*, 2013, pp. 16–26.

[75] A. Anand *et al.*, "XIA: An architecture for an evolvable and trustworthy Internet," in *Proc. ACM Workshop Hot Topics Netw.*, 2011, p. 2.

[76] T. Wolf *et al.*, "ChoiceNet: Toward an economy plane for the Internet," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 3, pp. 58–65, 2014.

[77] F. Oehlmann, "Content-centric networking," in *Proc. Seminar FI IITM Netw. Archit. Services*, vol. 43, 2013, pp. 11–18.

[78] L. Zhang *et al.*, "Named data networking," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 3, pp. 66–73, Jul. 2014.

[79] *Greenicn Architecture and Applications of Green Informationcentric Networking*. Accessed: Jan. 14, 2020. [Online]. Available: http://www.greenicn.org/

[80] A. Ioannou and S. Weber, "A survey of caching policies and forwarding mechanisms in information-centric networking," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2847–2886, 4th Quart., 2016.

[81] R. Tourani, S. Misra, T. Mick, and G. Panwar, "Security, privacy, and access control in information-centric networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 566–600, 1st Quart., 2018.

[82] H. Khelifi, S. Luo, B. Nour, and H. Moungla, "LQCC: A link quality-based congestion control scheme in named data networks," in *Proc. IEEE WCNC Conf.*, Apr. 2019, pp. 1–6.

[83] B. Nour *et al.*, "Internet of Things mobility over information-centric/named-data networking," *IEEE Internet Comput.*, vol. 24, no. 1, pp. 14–24, Jan./Feb. 2020.

[84] B. Nour, K. Sharif, F. Li, and Y. Wang, "Security and privacy challenges in information-centric wireless IoT networks," *IEEE Security Privacy*, vol. 18, no. 2, pp. 35–45, Mar. 2020.

[85] H. Khelifi *et al.*, "Security and privacy issues in vehicular named data networks: An overview," *Mobile Inf. Syst.*, vol. 2018, pp. 1–11, Sep. 2018.

[86] H. Khelifi *et al.*, "A blockchain-based architecture for secure vehicular named data networks," *Comput. Elect. Eng.*, vol. 86, Sep. 2020, Art. no. 106715.

[87] B. Nour *et al.*, "NCP: A near ICN cache placement scheme for IoT-based traffic class," in *Proc. IEEE GLOBECOM Conf.*, Dec. 2018, pp. 1–6.

[88] H. Khelifi *et al.*, "A QoS-aware cache replacement policy for vehicular named data networks," in *Proc. IEEE GLOBECOM Conf.*, Dec. 2019, pp. 1–6.

[89] O. Serhane *et al.*, "A label-based producer mobility support in 5G-enabled ICN networks," in *Proc. IWCMC Conf.*, 2020, pp. 1–6.

[90] O. Serhane *et al.*, "CnS: A cache and split scheme for 5G-enabled ICN networks," in *Proc. IEEE ICC Conf.*, Jun. 2020, pp. 1–6.

[91] M. Arumaithurai *et al.*, "Exploiting ICN for flexible management of software-defined networks," in *Proc. ACM ICN Conf.*, 2014, pp. 107–116.

[92] H. Li, K. Ota, and M. Dong, "ECCN: Orchestration of edge-centric computing and content-centric networking in the 5G radio access network," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 88–93, Jun. 2018.

[93] R. Tourani *et al.*, "IC-MCN: An architecture for an information-centric mobile converged network," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 43–49, Sep. 2016.

[94] A. Lertsinsrubtavee *et al.*, "Information-centric multi-access edge computing platform for community mesh networks," in *Proc. ACM SIGCAS Conf. Comput. Sustain. Soc.*, 2018, pp. 1–12.

[95] S. Boyd *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[96] A. Carzaniga *et al.*, "A routing scheme for content-based networking," in *Proc. IEEE INFOCOM Conf.*, vol. 2, 2004, pp. 918–928.

[97] Y. Zhang *et al.*, "A note on routing scalability in named data networking," in *Proc. IEEE ICC Conf. Workshops*, 2019, pp. 1–6.

[98] S. S. Adhatarao *et al.*, "Comparison of naming schema in ICN," in *Proc. IEEE Int. Symp. Local Metropolitan Area Netw. (LANMAN)*, 2016, pp. 1–6.

[99] B. Nour *et al.*, "A collaborative multi-metric interface ranking scheme for named data networks," in *Proc. IWCMC Conf.*, Jun. 2020, pp. 1–6.

[100] B. Nour *et al.*, "A unified hybrid information-centric naming scheme for IoT applications," *Comput. Commun.*, vol. 150, pp. 103–114, Nov. 2019.

[101] S. Arshad *et al.*, "Towards information-centric networking (ICN) naming for Internet of Things (IoT) the case of smart campus," in *Proc. ICFNDS Conf.*, 2017, pp. 1–6.

[102] S. Arshad *et al.*, "Hierarchical and flat-based hybrid naming scheme in content-centric networks of things," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1070–1080, Apr. 2018.

[103] M. A. U. Rehman *et al.*, "NINQ: Name-integrated query framework for named-data networking of things," *Sensors*, vol. 19, no. 13, p. 2906, 2019.

[104] S. S. Adhatarao, M. Arumaithurai, D. Kutscher, and X. Fu, "ISI: Integrate sensor networks to Internet with ICN," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 491–499, Apr. 2018.

[105] H. Khelifi, S. Luo, B. Nour, and H. Moungla, "A name-to-hash encoding scheme for vehicular named data networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, Jun. 2019, pp. 603–608.

[106] S. Feng, M. Zhang, R. Zheng, and Q. Wu, "A fast name lookup method in NDN based on hash coding," in *Proc. ICMII Conf.*, 2015, pp. 575–580.

[107] D. Saxena, V. Raychoudhury, N. Suri, C. Becker, and J. Cao, "Named data networking: A survey," *Comput. Sci. Rev.*, vol. 19, pp. 15–55, Feb. 2016.

[108] E. Jung, "A design of alias naming scheme for namespace in named data networking," in *Proc. ICACT Conf.*, 2016, pp. 537–540.

[109] S. Shan, C. Feng, T. Zhang, and Y. Liu, "A user interest preferences based on-path caching strategy in named data networking," in *Proc. IEEE/CIC ICCC Conf.*, 2017, pp. 1–6.

[110] N. Kamiyama and M. Murata, "Dispersing content over networks in information-centric networking," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 2, pp. 521–534, Jun. 2019.

[111] A. Azgin *et al.*, "A scalable mobility-centric architecture for named data networking," 2014. [Online]. Available: arXiv:1406.7049.

[112] Y. Zhang *et al.*, "KITE: Producer mobility support in named data networking," in *Proc. ACM ICN Conf.*, 2018, pp. 125–136.

[113] N. Leshov *et al.*, "Content name privacy in tactical named data networking," in *Proc. ICUFN Conf.*, 2019, pp. 570–572.

[114] I. O. Nunes *et al.*, "Namespace tunnels in content-centric networks," in *Proc. IEEE LCN Conf.*, 2017, pp. 35–42.

[115] B. Nour *et al.*, "M2HAV: A standardized ICN naming scheme for wireless devices in Internet of Things," in *Proc. WASA Conf.*, 2017, pp. 289–301.

[116] Y. Lu *et al.*, "A data-deduplication-based matching mechanism for URL filtering," in *Proc. IEEE ICC Conf.*, 2018, pp. 1–6.

[117] A. Ghodsi *et al.*, "Information-centric networking: Seeing the forest for the trees," in *Proc. ACM Workshop Hot Topics Netw.*, 2011, pp. 1–6.

[118] G. Tyson *et al.*, "A survey of mobility in information-centric networks: Challenges and research directions," in *Proc. ACM Workshop Emerg. Name Oriented Mobile Netw. Design Archit. Alg. Appl.*, 2012, pp. 1–6.

[119] G. Tyson *et al.*, "A survey of mobility in information-centric networks," *Commun. ACM*, vol. 56, no. 12, pp. 90–98, 2013.

[120] J. Wang *et al.*, "DMND: Collecting data from mobiles using named data," in *Proc. IEEE VNC Conf.*, 2010, pp. 49–56.

[121] B. Nour *et al.*, "A distributed cache placement scheme for large-scale ICN networks," *IEEE Netw.*, early access, Jul. 8, 2020, doi: 10.1109/MNET.011.2000081.

[122] H. Khelifi, S. Luo, B. Nour, and H. Moungla, "In-network caching in ICN-based vehicular networks: Effectiveness & performance evaluation," in *Proc. IEEE ICC Conf.*, Jun. 2020, pp. 1–6.

[123] V. Jacobson *et al.*, "Networking named content," in *Proc. CoNEXT Conf.*, 2009, pp. 1–12.

[124] N. Laoutaris *et al.*, "The LCD interconnection of LRU caches and its analysis," *Perform. Eval.*, vol. 63, no. 7, pp. 609–634, 2006.

[125] A. Dabirmoghaddam *et al.*, "Understanding optimal caching and opportunistic caching a 'the edge' of information-centric networks," in *Proc. ACM ICN Conf.*, 2014, pp. 47–56.

[126] W. K. Chai *et al.*, "Cache 'less for more' in information-centric networks (extended version)," *Comput. Commun.*, vol. 36, no. 7, pp. 758–770, 2013.

[127] I. Psaras *et al.*, "Probabilistic in-network caching for information-centric networks," in *Proc. ACM Workshop ICN*, 2012, pp. 55–60.

[128] S. Arianfar *et al.*, "Packet-level caching for information-centric networking," in *Proc. ACM SIGCOMM ReArch Workshop*, 2010, pp. 1–16.

[129] K. Cho *et al.*, "WAVE: Popularity-based and collaborative in-network caching for content-oriented networks," in *Proc. IEEE INFOCOM Workshops*, 2012, pp. 316–321.

[130] Q. Chen, R. Xie, F. R. Yu, J. Liu, T. Huang, and Y. Liu, "Transport control strategies in named data networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 2052–2083, 3rd Quart., 2016.

[131] L. Wang, S. Bayhan, and J. Kangasharju, "Optimal chunking and partial caching in information-centric networks," *Comput. Commun.*, vol. 61, pp. 48–57, May 2015.

[132] K. Suksomboon *et al.*, "PopCache: Cache more or less based on content popularity for information-centric networking," in *Proc. IEEE LCN Conf.*, 2013, pp. 236–243.

[133] Y. Li *et al.*, "Content popularity and node level matched based probability caching for content centric networks," in *Proc. IEEE/CIC ICCC Conf.*, 2016, pp. 1–6.

[134] Y. Meng *et al.*, "EHCP: An efficient hybrid content placement strategy in named data network caching," *IEEE Access*, vol. 7, pp. 155601–155611, 2019.

[135] Q. Zheng *et al.*, "A cache replication strategy based on betweenness and edge popularity in named data networking," in *Proc. IEEE ICC Conf.*, 2019, pp. 1–7.

[136] Q. N. Nguyen *et al.*, "PPCS: A progressive popularity-aware caching scheme for edge-based cache redundancy avoidance in information-centric networks," *Sensors*, vol. 19, no. 3, p. 694, 2019.

[137] J. Ren *et al.*, "Magic: A distributed max-gain in-network caching strategy in information-centric networks," in *Proc. IEEE INFOCOM Workshops*, 2014, pp. 470–475.

[138] B. Banerjee *et al.*, "Greedy caching: A latency-aware caching strategy for information-centric networks," in *Proc. IFIP Conf.*, 2017, pp. 1–9.

[139] H. Wu *et al.*, "Design and evaluation of probabilistic caching in information-centric networking," *IEEE Access*, vol. 6, pp. 32754–32768, 2018.

[140] N. Kamiyama *et al.*, "Spatially-dispersed caching in information-centric networking," in *Proc. IEEE ICC Conf.*, 2018, pp. 1–7.

[141] T. Mick *et al.*, "MuNCC: Multi-hop neighborhood collaborative caching in information centric networks," in *Proc. ACM ICN Conf.*, 2016, pp. 93–101.

[142] L. Huang *et al.*, "On-path collaborative in-network caching for information-centric networks," in *Proc. IEEE INFOCOM Workshops*, 2017, pp. 175–180.

[143] J. Zhi *et al.*, "GAC: Gain-aware 2-round cooperative caching approach in information-centric networking," in *Proc. IEEE IPCCC Conf.*, 2018, pp. 1–8.

[144] W.-X. Liu *et al.*, "COD: Caching on demand in information-centric networking," *Telecommun. Syst.*, vol. 69, no. 3, pp. 303–319, 2018.

[145] H. Noh *et al.*, "Progressive caching system for video streaming services over content centric network," *IEEE Access*, vol. 7, pp. 47079–47089, 2019.

[146] H. Khelifi *et al.*, "An optimized proactive caching scheme based on mobility prediction for vehicular networks," in *Proc. IEEE GLOBECOM Conf.*, Dec. 2018, pp. 1–6.

[147] J.-L. Deng *et al.*, "Control problems of grey systems," *Syst. Control Lett.*, vol. 1, no. 5, pp. 288–294, 1982.

[148] L. Breslau *et al.*, "Web caching and Zipf-like distributions: Evidence and implications," in *Proc. IEEE INFOCOM Conf.*, vol. 1, 1999, pp. 126–134.

[149] G. Zhang *et al.*, "Caching in information centric networking: A survey," *Comput. Netw.*, vol. 57, no. 16, pp. 3128–3141, 2013.

[150] H. Balakrishnan *et al.*, "Looking up data in P2P systems," *Commun. ACM*, vol. 46, no. 2, pp. 43–48, 2003.

[151] J. Yang *et al.*, "Patterns of temporal variation in online media," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2011, pp. 177–186.

[152] Y. Zeng *et al.*, "Smart caching based on user behavior for mobile edge computing," *Inf. Sci.*, vol. 503, pp. 444–468, Nov. 2019.

[153] Z. Zhang *et al.*, "Smart caching: Empower the video delivery for 5G-ICN networks," in *Proc. IEEE ICC Conf.*, 2019, pp. 1–6.

[154] Y. Tang *et al.*, "A smart caching mechanism for mobile multimedia in information centric networking with edge computing," *Future Gener. Comput. Syst.*, vol. 91, pp. 590–600, Feb. 2019.

[155] Y. Tang, "Minimizing energy for caching resource allocation in information-centric networking with mobile edge computing," in *Proc. IEEE Conf. DASC/PiCom/CBDCom/CyberSciTech*, 2019, pp. 301–304.

[156] L. Han *et al.*, "Deep learning based loss recovery mechanism for video streaming over mobile information-centric network," *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 9, pp. 4572–4586, 2019.

[157] Y. Liu, A. Liu, N. N. Xiong, T. Wang, and W. Gui, "Content propagation for content-centric networking systems from location-based social networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 10, pp. 1946–1960, Oct. 2019.

[158] C. Wang *et al.*, "Content-centric caching using deep reinforcement learning in mobile computing," in *Proc. IEEE HPBD&IS Conf.*, 2019, pp. 1–6.

[159] G. Huang *et al.*, "Online Learning Framework based on user-centric access behavior," in *Proc. IEEE HPCC/SmartCity/DSS Conf.*, 2018, pp. 290–297.

[160] J. Han *et al.*, "Mining frequent patterns without candidate generation: A frequent-pattern tree approach," *Data Mining Knowl. Discovery*, vol. 8, no. 1, pp. 53–87, 2004.

[161] P. Czyżżak and A. Jaszkiewicz, "Pareto simulated annealing—A metaheuristic technique for multiple-objective combinatorial optimization," *J. Multi Criteria Decis. Anal.*, vol. 7, no. 1, pp. 34–47, 1998.

[162] H. Khelifi *et al.*, "Bringing deep learning at the edge of information-centric Internet of Things," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 52–55, Jan. 2019.

Oussama Serhane (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in computer science from Djillali Liabes University, Sidi Bel Abbes, Algeria, in 2016 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Université Mustapha Stambouli, Mascara, Algeria.

His research interests include next-generation networking, information-centric network, and 5G networks.

Khadidja Yahyaoui received the Engineer degree in soft engineering and the M.Tech. degree in industrial computing from the Computer Science Department, Oran University, Oran, Algeria, in 2000 and 2006, respectively, and the Ph.D. degree in computer science from Oran University in 2013.

She is currently a Research Member of the Computer Science Laboratory, Mascara University, Mascara, Algeria. Her main research interests include optimization, artificial intelligence, Internet new architecture, and reconfigurable embedded systems.

Boubakr Nour (Member, IEEE) received the Ph.D. degree in computer science and technology from the Beijing Institute of Technology, Beijing, China, in 2020.

His research interests include next-generation networking and Internet.

Dr. Nour is a recipient of the Best Paper Award at IEEE GLOBECOM in 2018 and the Excellent Student Award at Beijing Institute of Technology in 2016, 2017, and 2018.

Hassine Moungla (Member, IEEE) received the Ph.D. degree from the University of Paris XIII, Villetaneuse, France, in 2006.

He is an Associate Professor with the Université de Paris, Paris, France, where he was a Member of the Paris Descartes Computer Science Laboratory since 2008. He was a Researcher with INRIA, Cedex, France, until 2008 and a Research Fellow with CNRS-LIPN Laboratory, Paris Nord University, Paris, France, until 2007. His research interests are in the field of wireless area body networking, WSNs, middleware for 5G, and sensor networks.

Dr. Moungla serves on editorial board of several conferences.