electronics

Article

Overview of Prospects for Service-Aware Radio Access towards 6G Networks

Zixiao Zhao [1,2], Qinghe Du [1,2,*], Dawei Wang [3], Xiao Tang [3,4] and Houbing Song [5]

1 School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China; zzx67120787@stu.xjtu.edu.cn
2 Shaanxi Smart Networks and Ubiquitous Access Research Center, Xi'an 710049, China
3 School of Electronics and Information, Northwestern Polytechinical University, Xi'an 710072, China; wangdw@nwpu.edu.cn (D.W.); tangxiao@nwpu.edu.cn (X.T.)
4 National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
5 Electrical Engineering and Computer Science Department, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA; houbing.song@erau.edu
* Correspondence: duqinghe@mail.xjtu.edu.cn

Abstract: The integration of space–air–ground–sea networking in 6G, which is expected to not only achieve seamless coverage but also offer service-aware access and transmission, has introduced many new challenges for current mobile communications systems. Service awareness requires the 6G network to be aware of the demands of a diverse range of services as well as the occupation, utilization, and variation of network resources, which will enable the capability of deriving more intelligent and effective solutions for complicated heterogeneous resource configuration. Following this trend, this article investigates potential techniques that may improve service-aware radio access using the heterogeneous 6G network. We start with a discussion on the evolution of cloud-based RAN architectures from 5G to 6G, and then we present an intelligent radio access network (RAN) architecture for the integrated 6G network, which targets balancing the computation loads and fronthaul burden and achieving service-awareness for heterogeneous and distributed requests from users. In order for the service-aware access and transmissions to be equipped for future heterogeneous 6G networks, we analyze the challenges and potential solutions for the heterogeneous resource configuration, including a tightly coupled cross-layer design, resource service-aware sensing and allocation, transmission over multiple radio access technologies (RAT), and user socialization for cloud extension. Finally, we briefly explore some promising and crucial research topics on service-aware radio access for 6G networks.

Keywords: radio access networks; cloud-based RAN; 6G; service-aware

check for updates

Citation: Zhao, Z.; Du, Q.; Wang, D.; Tang, X.; Song, H. Overview of Prospects for Service-Aware Radio Access towards 6G Networks. *Electronics* **2022**, *11*, 1262. https://doi.org/10.3390/electronics11081262

Academic Editor: Christos J. Bouras

Received: 28 February 2022
Accepted: 14 April 2022
Published: 16 April 2022

1. Introduction

The 5th Generation (5G) mobile communications system has attracted global attention from the Information and Communication Technology (ICT) sector in the past few years [1–13] and has also been used for increasingly sophisticated commercial applications. Discussions on the prospects of the 6th Generation (6G) mobile communications system have inspired new research [14–23]. There is a consensus in the research community that the future 6G network will evolve towards the integration of space, air, ground, and sea, providing communication services with much wider coverage and more satisfactory experiences compared with 5G [18,19,24,25]. Based on the three major application scenarios of 5G—i.e., massive machine-type communications (mMTC), ultra-reliability, and low-latency communications (uRLLC)—and the enhanced mobile broadband communications (eMBB), new types of services will emerge and be defined in the 6G network. These will have more comprehensive but also more diversified quality-of-service (QoS) requirements. However, the current 5G architecture may not be adequate to support the newly appearing applications due to the following reasons: (1) The heterogeneousness of both the networks and resources is constantly increasing due to the presence of diverse networking architectures, simultaneous multi-scale optimization, fragmentized resources, and a highly distributed topology. This urgently requires the development of a complete and universal AI-driven management structure [26]. (2) With the appearance of new radio access technologies (RAT) and the expansion of the air interface access capability, the current network inevitably shows a deficiency in coordination across concurrent services. Though there have been discussions on the coexistence of LTE and 5G, the 5G architecture mainly supports network switching, rather than fulfilling different service performance needs with a collaboratively unified framework. (3) There has been dramatic growth in physical technologies, e.g., the bandwidth, processing speed, and flexibility [27]. However, the upper layers of the protocol stack have lacked timely evolution compared with other layers. The closer to the top of the protocol stack a layer is, the longer the interaction delay introduced is, and this prevents the network from catching up with the faster change in the transmission environment at the physical layer. (4) In spite of the great convenience introduced by enabling multi-RAT capability in single devices, the new problem of co-site interference, as occurs when accessing 4G and 5G air interfaces simultaneously, severely degrades the quality of service. (5) Security strategies also need to evolve with new physical layer resources [28,29].

While the RAT required to remedy the aforementioned issues will no doubt cover a wide spectrum of topics and motivate massive research efforts, this article concentrates on potential techniques that may significantly benefit service-aware access, either for issues newly encountered in 5G networks or for those that have not yet been thoroughly studied for 6G networks. Specifically, we explain the concept of service awareness for 6G networks from three perspectives: (1) understand users' quality of experience (QoE)/QoS requirements; (2) the ability to sense and analyze past/current service statuses and be capable of conducting timely and effective network optimization; (3) the awareness of the resource occupation statuses, utilization patterns, and varying tendencies, and the capability of effective discovery, allocation, coordination, aggregation of resources.

We initiate our discussion on the intelligent Radio Access Network (RAN) architecture, which is targeted at effectively balancing the computation load, fronthaul burden, and service-awareness demands. The distributed AI technologies in each layer not only assist with the prediction of network behaviors prospectively and optimize the network to achieve faster service responses, they also identify heterogeneous resources and coordinate various user requirements.Service awareness requires the 6G network to be aware of the demands of diverse services as well as complicated resource occupation and utilization tasks. This will give it the potential to provide services with more intelligent and effective resource-configuration strategies. In particular, we analyze the challenges and potential solutions for the heterogeneous resources configuration, including a tightly coupled cross-layer design, resource service-aware sensing and allocation, transmission over multiple radio access technologies (RAT), and user socialization for cloud extension. Though local sensing and optimization may lead to a better performance for a single or group of devices, this distributed strategy will increase the interaction delay, especially when working for services with a highly varied status. Thus, we propose the use of a tightly coupled cross-layer design for the protocol stack as powerful support for the aforementioned architecture. Determining how to sense and monitor heterogeneous resources, including the spectrum, computing, and caching resources, is crucial to achieve the expected network performance while reducing unnecessary collisions and assuring diverse QoS requirements are met. There are multiple RATs in 6G networks, allowing users to obtain more suitable communication resources and faster network responses flexibly. We discuss how to fully utilize these conveniences via multi-path transmission and also address the newly encountered problem of co-site interference. User socialization, in which an idle device provides services for others through resource sensing and computing, could be a potential solution that allows the achievement of a seamless service.

In this article, we start with a brief preview on the evolution of cloud-based RAN architectures from 5G to 6G. Then, to address the aforementioned problems, we present an intelligent cloud RAN architecture that provides service-aware radio access in 6G networks. Using this architecture, we further discuss the major challenges and potential solutions involved in integrating service awareness with radio access in 6G networks with emphasis on heterogeneous-RAT access, tight coupling across protocol layers, resource sensing and allocation, and user socialization. We conclude the article after discussing current issues and future research directions for radio access in 6G networks.

2. The Evolution of Cloud-Based RAN from 5G to 6G

The cloud-based RAN was originally proposed to effectively reduce the capital and operational expenditure [8,9]. Furthermore, it also motivates the strong desires of wireless communications engineers to globally optimize resource allocation with service-awareness access over the RAN. Initially, when the mobile communications systems evolved to the 5G network, users were expected to take advantage of 5G's unique features such as the high RAT diversity allowing the fulfillment of QoS requirements for latency, reliability, throughput, mobility, etc. [3,30,31]. However, the multiple heterogeneous coexisting RATs, together with different types of serving stations deployed in an ultra-dense fashion, severely complicate the entire system, leading to severe cross interference. Therefore, it is harder than ever to efficiently manage the network as well as effectively adapt the network to the highly varied demands and environments [4,9] in a more distributive manner.

To address this problem, the cloud-based RAN emerged as a revolutionary architecture, where one of the representative designs is the Cloud RAN (C-RAN) proposed by China mobile [8]. Rather than undergoing simple separation of the radio servers and radio units, as was done in previous network architectures, the cloud-based RAN pools the BaseBand Units (BBU) such that they can be conveniently and flexibly accessed by a large number of remote radio units, called remote radio heads (RRH) or remote radio units (RRU). In the cloud-based RAN, BBUs are integrated and viewed as one entity, namely the BBU pool, in order to support the vast computing burden of remotely and densely deployed RRHs, which are connected to the BBU pool via fronthaul links.

The 6G networks will be designed to fully connect four major geographical domains, i.e., space–air–ground–sea. In particular, satellite networks contain satellites and space stations with different orbital altitudes; unmanned aerial vehicle (UAV) networks are assisted by UAVs and airplanes; terrestrial networks evolve on the basis of cellular systems; and maritime networks provide wireless connections and the ability to communicate with devices on and below the surface. A typical RAN architecture for a 6G network is depicted in Figure 1. As shown in this figure, though we mainly emphasize the new RATs proposed for 6G networks, some typical networks and RATs, including 2G–5G, WLAN, WPAN, and even millimeter-wave networks and visible-light networks, will also evolve and be employed flexibly. Such RATs, along with diverse terminals integrated under space–air–ground–sea scenarios and the central pool, form a cloud structure. The RATs connect to the cloud and their demands for computing and caching are served by the cloud in a centralized manner. The cloud pool allocates a subset of computing and caching resources and integrates necessary information that is requested by services or reported from peer access terminals via resource sensing. Then, it completes the control functions, such as baseband signal processing, scheduling, and radio resource allocation.

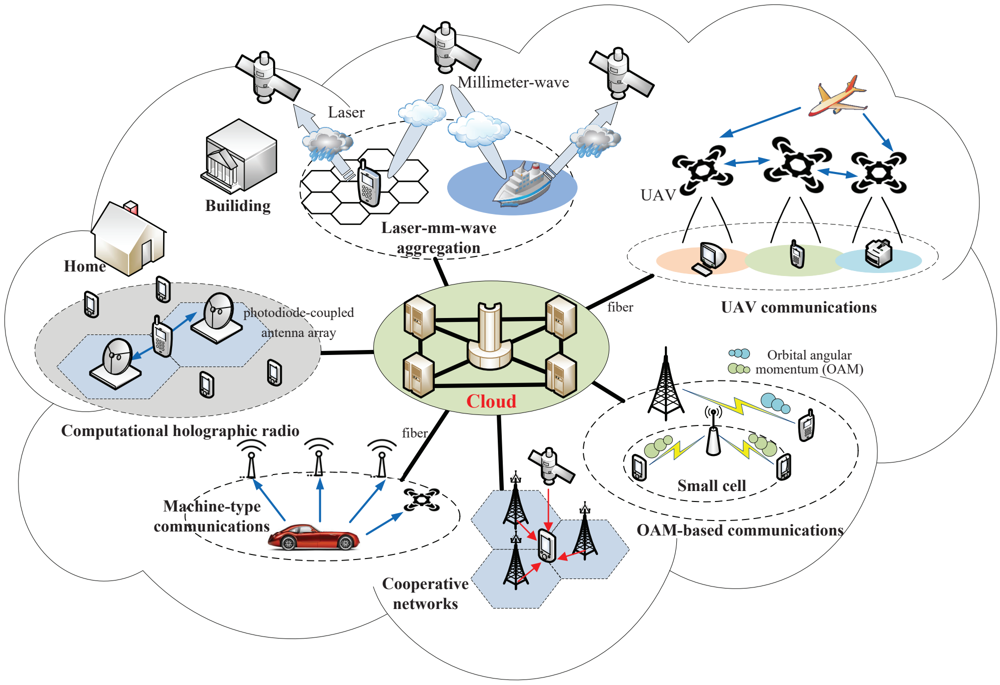

Figure 1. Cloud-based radio access network architecture for 6G mobile communications systems.

In particular, laser-mm-wave aggregation can balance the shortness of laser and millimeter-waves on cloudy/foggy and rainy days, respectively [16]. A benefit of the co-operation between the two waves is that the satellites can support communications even in complex weather conditions. Orbital angular momentum (OAM)-based communication utilizes orthogonal beams with different OAM modes and offers a new method of multiple access with less interference and higher reliability compared with the traditional frequency/time/code-domain [32]. The computational holographic radio extracts available information from the received interference to achieve precise control of the whole communication environment [16]. Correspondingly, new RATs will, no doubt, inject much vigor into the 6G RAN, which will have more capability to support the connection requirements initiated by different devices and users in globally integrated scenarios.

Based on the novel RATs, the 6G network will extend the current applications and also support new applications. We depict these in Figure 2. However, higher standards in terms of the network's connection, computing, and sensing capability will be required.

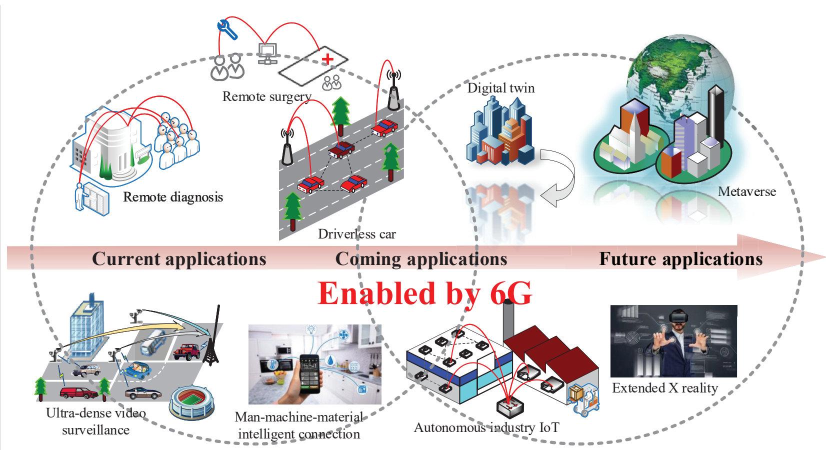

Figure 2. Coming and future applications enabled by 6G.

- Current and Coming Applications

(1) Ultra-dense video surveillance: In the past, network video surveillance has played a major role in security systems. However, limited by bandwidth, cost, and transmission rate, the coverage area and video clarity are unsatisfactory. The 6G network will better support ultra-dense monitors with higher speed and a lower overhead, especially in IoT scenarios.
(2) Telemedicine: Enabled by 5G, patients have been able to receive remote diagnoses and treatments. The problem is determining how to assure real-time imaging and a synchronous conversation, especially for remote surgeries. Our tightly coupled cross-layer framework design, which is introduced below, will make this vision possible from the perspective of an appropriate protocol.
(3) Driverless cars: As the typical use case for URLLC in 5G, the idea of driverless cars has raised widespread concern from the perspective of appropriate technologies; however, another equally important issue is scientific ethics. The 6G era may provide a balanced solution while fully ensuring security.
(4) Man-machine-material intelligent connection: Many theoretical conclusions have been derived about cognitive functioning between humans and machines [33,34], and some realistic functions have already been developed, e.g., human–machine conversations, smart homes, wearable smart devices, etc. However, there is still quite a long way to go toward the achievement of comprehensive applications.

- Coming and Future Applications

(1) Extended X reality: In the 5G era, some applications of virtual reality (VR) and augmented reality (AR) have shown attractive business prospects. In the future, the 6G network will assist with the extension of limited virtual experiences into borderless and immersive worlds generated by the use of wearable devices.
(2) Autonomous industry IoT: The organization and management of industrial devices is achieved through a central platform without manual intervention. In [35], a generalized cognitive model for autonomous IoT, which fully considers the reasoning, learning, and planning processes is proposed. However, the development of a detailed design and adaption scheme for this structure remain challenging.
(3) Digital twin: The digital twin represents a high-fidelity digital mirror of the physical entity, and the live replica of a process or whole 6G network has the potential to predict the impact of decisions and optimize the organization modes [36]. By penetrating AI into the edge, the 6G network will better support vertical use cases and predict outbreak events accordingly.
(4) Metaverse: Beyond simply creating a physical, virtual space, the metaverse will be able to provide an immersive experience with a story through user interaction [37]. However, in addition to hardware and software limitations, another challenge is determining how to protect the massive amount of private information.

3. Intelligent RAN Architecture for the Space–Air–Ground–Sea Integrated 6G Network

Despite the attractive characteristics of the current RAN in 5G networks, e.g., the centralized design and hierarchical collaboration, it is still not capable of supporting the emerging vertical services in the 6G era [38]. Specifically, the limitation on the fronthaul capacity severely weakens the cloud's ability to optimize resource allocation, organize cooperation, and fulfill users' QoS requirements [7,11]. On the other hand, motivated by the more complex communication scenarios and customized services in the 6G network, the prospects of service-awareness access are becoming more attractive, which requires the service providers to fully consider the users' personal demands and sense the resource status [39]. Moreover, communication in various environments, i.e., space, air, ground, and sea, needs to be organically unified and seamless coverage needs to be achieved.

Next, we show that the alternative way to alleviate the above problems is through the employment of an intelligent cloud architecture, as illustrated in Figure 3, consisting of three layers, namely *the cloud*, *edge servers*, and *integrated air interfaces*. The first layer is the cloud layer, which plays the core role of the RAN, taking the majority of the computing tasks for signal processing, resource allocation, and other compute-intensive tasks. The cloud pool is not only linked to the edge server layer but is also connected to many air interfaces directly through high-bandwidth fiber or electromagnetic (EM) waves, depending on the actual physical properties of the links. The majority of resources in the network are controlled by the cloud and allocated to edge servers or users. Fulfillment of the functionality of this layer relies on infrastructure built by the major service providers as well as the government.

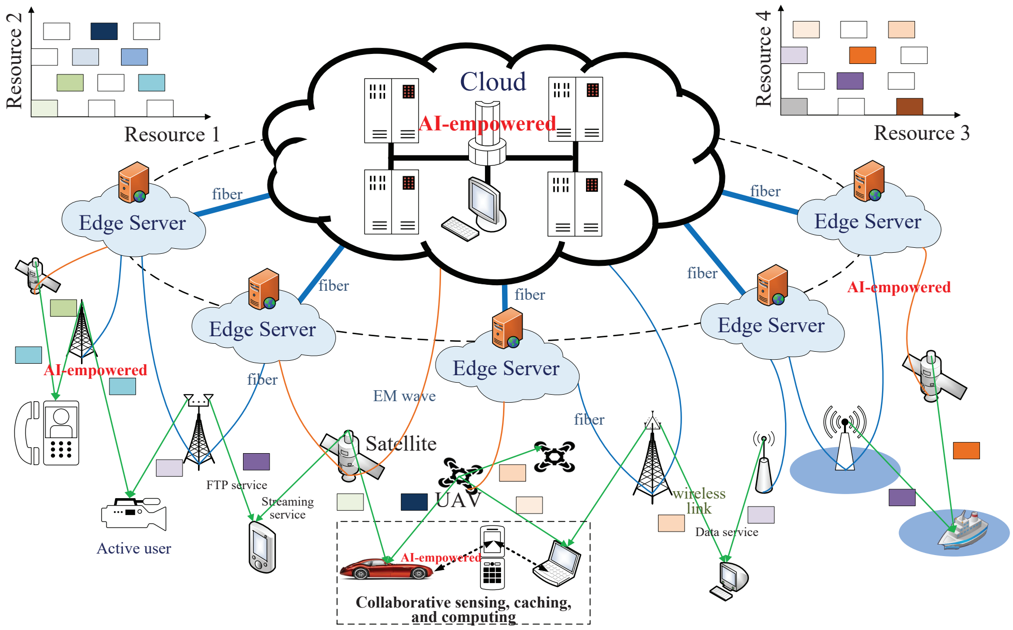

Figure 3. Intelligent cloud architecture for the 6G RAN.

The second layer, i.e., the edge server layer, is composed of the computing units, namely the edge servers that perform similar yet independent functions to the cloud with less computing capability. Its major task is to offload computing tasks at the cloud to itself, which can effectively relieve the burden of the fronthaul connecting the cloud pool. Necessary and important information, such as resource allocation results or mobile users' QoS satisfaction, are fed back to the cloud layer, aiding in efficient computation. The edge servers often deal with delay-sensitive services, and the service delay can be shortened by avoiding excessive message exchange overhead.

The third layer is the integrated air interface layer, which includes various air interfaces scattered in the space–air–terrestrial–sea environment, e.g., UAV, airplane, satellite, sea-level base station, and typical RRH, supporting 2G–5G, WiFi, and other RATs. In this layer, all air interfaces receive signals to be emitted from the cloud or edge servers. If the computing load for certain services is low and/or the delay requirement is very stringent, the cloud typically designates the edge server near the air interface/users to cope with the corresponding processing tasks. Furthermore, the cloud assigns resources along with constraints, such as the power budget and spectrum mask, to the edge server for flexible configuration. In contrast, if some services, such as complicated coding techniques with large data blocks, cooperation with a large user population, or network-wide optimization, cause a very large computing load that is beyond the edge server's capability, the cloud handles the computing task directly and then forwards baseband-ready signals to the corresponding air interfaces. This setup evidently exhibits the service-aware feature. The association between the cloud/edge server and air interface obeys the following rules: (1) an air interface can receive signals from the cloud and the edge server simultaneously but cannot communicate with two edge servers concurrently, thereby avoiding potential conflicts in resource allocation due to independent edge server operations; (2) the association between the air interface and the edge server may vary over time, allowing a flexible configuration.

The distributed artificial intelligence (AI) provides integrated coordination for both cross-layers and internal function units. The purpose of this is not only to assist in the prediction of the networks' behaviors prospectively and conduct optimization to achieve faster service responses, it also attends to sensing heterogeneous resources and coordinating various users' requirements. We illustrate the logical structure of the distributed AI in Figure 4 across two dimensions. One is AI's vertical dimension in each function layer of a communication system, and the other is AI's horizontal dimension between multiple users and between users and the cloud.

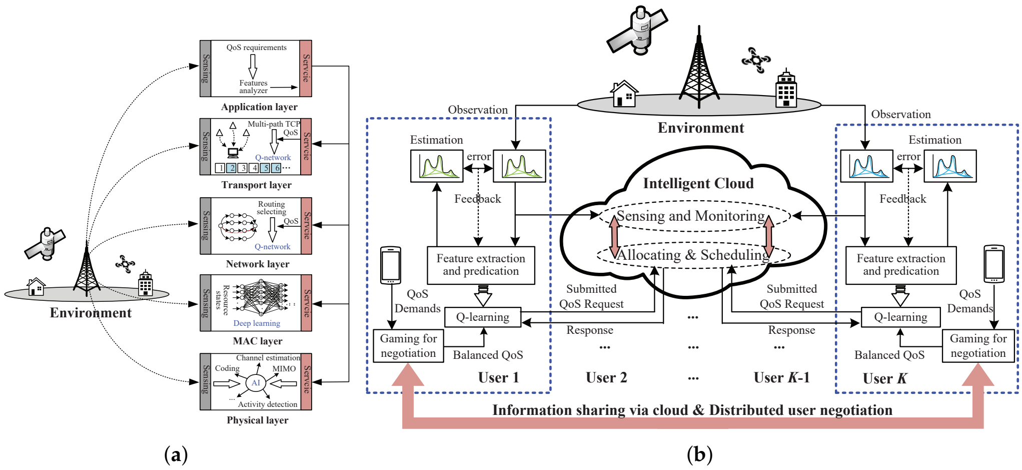

Figure 4. The model of the distributed AI: (**a**) the vertical dimension; (**b**) the horizontal dimension.

As shown in Figure 4a, we suggest that the adoption of the AI structure should be driven by data and model simultaneously. Each layer has a sensor node that is used to sense and collect available information from the space–air–ground–sea environment. The top application layer can analyze and extract features from the service's QoS and then deliver these to other layers based on different models' requirements. Data from both sensing and services are used to train and update the AI models adaptively, i.e., model self-learning. Typical schemes employed to fulfill the model's self-learning include GAN, deep learning (DL), and LSTM. GAN has the ability to dynamically monitor and simulate the internal features of observations and then derive an abstract model to depict statistical rules. DL can seize the deep connection of input data and iterate towards expected outputs. LSTM is suitable for models with correlations in the time domain; thus, it can predict the model's future trends based on its historical performances. Specifically, the traditional models in the physical layer, which were mainly derived according to the characteristics of transmitters and common protocols, can be generated and updated adaptively with the assistance of AI techniques in 6G networks. In this regard, GAN has shown the potential to model the channels [40]. For the MAC layer, we consider employing DL to extract implicit information from resource states. In the transport layer and network layer, on the one hand, we can use the Q-learning network to solve multi-scale computing problems (e.g., routing selections and multi-path transmissions), and on the other hand, these models can learn from external data driven by GAN, DL, LSTM, etc.

In Figure 4b, we illustrate the process of information sharing and negotiation across users, which is driven by the distributed AI framework. First, each user obtains observations from the network environments, including information of diverse resources (e.g., spectrum, caching, and computing resources) and transmission conditions. These observations are injected into an extraction/predication module to get key features of the transmission and network environments. Furthermore, future environments can be estimated via the predication function, which is used to guide user's transmission strategy. The estimations are further compared with the observations, and the obtained error is also fed back to the extraction/prediction module, which will be used to optimize the local extraction and predication algorithms. In the meantime, the observations by users are also uploaded to the cloud, so that the cloud can sense and monitor environment information over the entire network. Second, the cloud will help distribute user's QoS demands to each other, together with the resource and network environment information. With these information, each user will conduct negotiation in a distributive manner, through which each user will adjust its own QoS requirements to a more balanced level, i.e., acceptable for itself yet potentially feasible for the network. For QoS negotiation, users will adopt game-theory based approaches, as the game theory has been proven to be capable of effectively solving complicated problems such as resource competition, coordination, allocation, etc., over wireless networks and the game theory itself is persistently evolving with more intelligence [41–45]. The negotiated QoS and extracted features from observations and estimation are used to finalize its QoS requirements by the reinforcement learning approach. Then, by analyzing users' QoS requests and the available resources over the network, the intelligent cloud can effectively optimize resource allocation and scheduling improving the network performances while meeting users' QoS requirements. Moreover, the optimized scheduling and resource allocation strategies will be sent back to users such that they can better train their learning module.

The intelligent cloud architecture has the following essential features. First, the distributed AI can assist the architecture to adaptively coordinate the limited resources with the users'demands. Through estimating the possible access performance with multiple available RATs, each user will select the best solution for itself. This process will also be coordinated with the cloud and/or edge servers, depending on the serving load and available resources from both a local and global perspective. Second, the cloud layer and the edge server layer can balance the computing load. Cloud and edge servers have similar serving functions for end users, and thus this structure can be transparent to end users with the integrated air interface layer. In this sense, the hierarchical architecture does not complicate the network. Third, this architecture fully embraces the RAT diversity in the 6G network. In particular, each end user can receive multiple services simultaneously from multiple air interfaces with different RATs. Furthermore, to accommodate more stringent QoS requirements, it is also possible for a single service to deliver data over multiple air interfaces with different RATs. Last but not least, by efficiently distributing and balancing computing loads and AI-driven resource allocation and coordination by the cloud, the hierarchical architecture is potential to considerably improve the overall network performance while meeting users' QoS requirements. It is worth noting that, even with advanced machine learning (ML) algorithms and fast-developing computing abilities, completely centralized optimization over the large-size network is still impractical. Hence, the employment of a hierarchical structure with distributed AI can no doubt lower the complexity and improve the scalability of the 6G network. To sum up, the intelligent cloud structure is desirable as it can flexibly distribute the computing load and scales well with the network size. However, under this architecture, how to deal with the essential heterogeneousness over the varying communication environments of the 6G network, as well as providing service-aware resource configuration to fulfill users' diverse requirements still faces many challenges, which will be discussed in the next section.

4. Challenges and Potential Solutions for Achieving a Service-Aware Heterogeneous Resource Configuration

The intelligent cloud-based RAN mainly concentrates on the way computing abilities are provided and organized. Determining how to efficiently configure the highly heterogeneous resources of the 6G network, which are expected to be aware of users' QoS/QoE requirements, requires thorough study on diverse aspects. We next discuss four major issues in service-aware heterogeneous resource configuration in the RAN of 6G networks: (1) From a network protocol perspective, does the current protocol stack fit well with the highly varied heterogeneous networks? (2) From the resource-control perspective, how is resource sensing and monitoring organized in the intelligent cloud? (3) From the air interface perspective, how can we support users' services via multiple heterogeneous RATs? (4) From the user socialization perspective, how can we achieve user cooperation via D2D communications?

4.1. Tightly Coupled Cross-Layer Design with Service Awareness

The RAN architecture and diversified heterogeneous resources of the 6G network exhibit significant differences from conventional networks. However, the network protocol stack in each network node is almost unchanged. Thus, the question of "does the current protocol stack fit well with the network environments nowadays?" arises. Two attractive new features for the 6G network are service-aware access with seamless coverage and the presence of various heterogeneous RATs. Correspondingly, the moving users and varying mobile channels, even those moving at a very slow speed, result in highly varying network statuses [46]. Figure 5 shows how the conventional protocol stack reacts to fast changes in the network. The physical (PHY) and link layer can adapt their parameters on the millisecond time scale (e.g., 5G NR divides the 1 ms subframe into several slots according to subcarrier spacing, and one of the typical slot lengths is 0.125 ms with 120 Khz subcarrier spacing). The higher the protocol layer, the longer the response time, and this may reach some tens of seconds. Therefore, with a fast and highly varying network status, the upper protocol layers cannot match the pace of lower protocol layers, resulting in severe inconsistent strategies, as depicted in Figure 5. This significantly lowers the efficiency of mobile networks. Note that the inconsistency problem is not an issue in wired or traditional mobile networks, because the networks are typically rather stable. Moreover, the distributed AI of the aforementioned intelligent RAN architecture may separately and automatically conduct local optimization and make decisions, which should be supported and coordinated by the cloud.

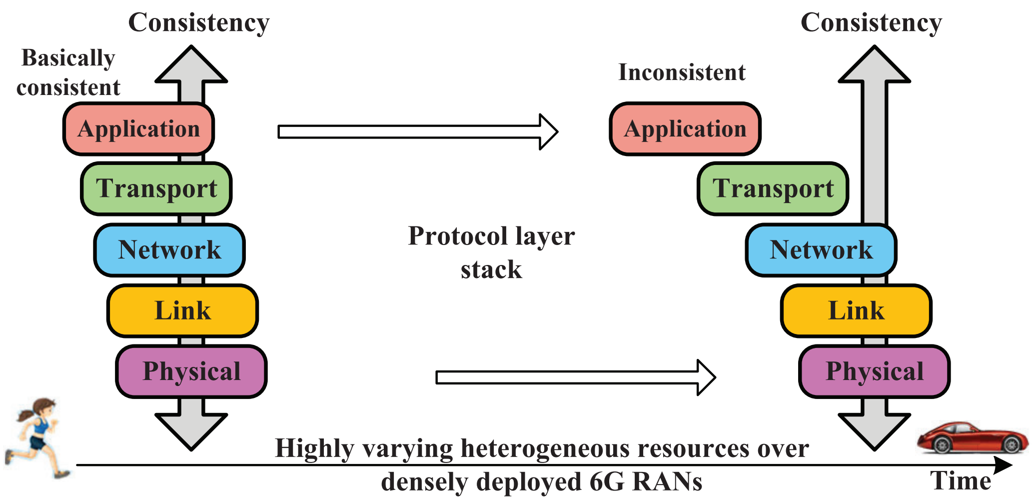

Figure 5. Inconsistent responses to the highly varying heterogeneous network status caused by the loosely coupled protocol stack.

To address this problem, we propose a bidirectional iterative protocol stack framework based on a cross-layer design, as depicted in Figure 6. The designed framework consists of two main parts: the protocol stack entity and the bidirectionally iterative strategy optimizer (strategy optimizer in short). The former carries out the networking functionality following the traditional layered structure, whereas the latter determines the transmission strategy, through which different protocol layers are expected to be tightly coupled to avoid inconsistency. In Figure 6, we use the typical stack structure of the 3GPP as an example, where the link layer is replaced by the PDCP (Packet Data Convergence Protocol) layer, RLC (Radio Link Control) layer, and MAC layer.

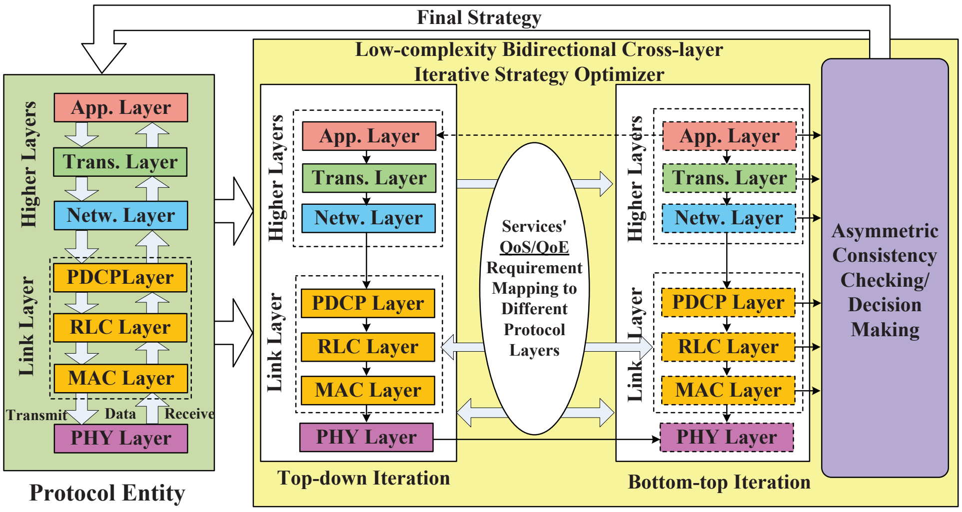

Figure 6. Tightly coupled cross-layer framework with service-awareness achieved via bidirectional iterative strategy optimization (using the typical architecture of a protocol stack in 3GPP as an example).

The protocol entity inherits the classical layer structure, as it can be easily transplanted for diverse wireless networks with distributive regulation. The strategy optimizer bidirectionally optimizes the transmission strategy across all layers, representing the core innovative concept. It includes four sub-function blocks: the QoS-/QoE-mapper, top-down sub-optimizer, bottom-up sub-optimizer, and consistency check module. The QoS-/QoE-mapper collects users' QoS/QoE requirements, which are used to orient parameter optimization in different protocol layers, integrating service awareness. Bidirectional strategy optimization, carried out by the top-down and bottom-up sub-optimizers, is the key element required to overcome mismatching decisions by different layers. Optimization starts with the bottom-up sub-optimizer, which initiates the transmission strategy of each protocol in a consecutive order from the PHY layer to the application layer. Different protocol layers optimize their respective objectives based on the channel quality and available resources (information from lower protocol layers) and are subject to the QoS/QoE constraints (information from the QoS/QoE mapper). The bottom-up sub-optimizer operates in the opposite direction, from the application to the physical layer, where adjustments at each layer are based on updated strategy information from the upper layers. Bottom-up and top-down optimization are performed in an iterative fashion to sufficiently exchange information across protocol layers. On one hand, this iteration mechanism achieves a low level of complexity as each layer's optimization is conducted separately, while on the other hand, the information is fully exchanged such that the strategies at different protocol layers are tightly coupled and are as consistent as possible.

We further integrate a consistency-check module, which examines whether the strategies of different protocol layers are consistent. If they are, this module terminates the iteration and passes the results to the protocol entity. If the iteration does not converge until timeout, the strategy rules in favor of the lower layer while keeping the higher protocol layers' strategies before the iteration unchanged. The consistency check is asymmetric, because the lower layers are close to the propagation environments.

This design presents a framework that has the potential to overcome the loosely coupling issue for the conventional protocol stack. Moreover, the proposed cross-layer framework can be designed within part of the protocol layers, allowing flexibility. It will powerfully support the new services in the 6G network beyond mMTC, uRLLC, and eMBB, significantly reducing the latency. Considering some real-time spatial and temporal changes, e.g., the device density, traffic patterns, and spectrum availability, the Event Defined uRLLC (EDuRLLC) is defined as the evolved version of the uRLLC [17]. It has higher requirements in terms of the response speed and coupling of cross-layers. With the frequent switching of air interfaces and RATs in high-speed mobile services, the independent design of QoS optimization at each protocol layer can support more types of latency-sensitive services [47]. However, determining how to integrate diverse QoS/QoE requirements when designing the specifc protocol stack is an attractive yet widely open problem.

4.2. Service-Aware Heterogeneous Resource Sensing and Allocation

In practical systems, resource allocation solutions are computed based on specific protocols and access mechanisms [48–51]. After these factors have been standardized, despite the powerful computing ability provided by the cloud, there are still some unique challenges that require further attention. One challenge is determining how to optimize service-aware resource allocation for the edge server; the other is determining how to discover and aggregate resource fragments for service provisioning requested end users.

As discussed previously, the edge server also has a certain computational ability to deal with base-band signal processing as well as resource allocation for users associated with it. However, the numbers of resources allocated to the air interfaces and edge servers need to be determined at the upper-layers. One possible solution is that the air interfaces each upload the occupation states of the resources to the corresponding edge server within the specified time scale, and then the edge server is able to estimate the current network loads according to the local mathematical model [51]. The other method is to extract the model from the history data of the edge server and then conduct estimations and predictions. This method is especially suitable for machine communication with periodic traffic patterns. Based on the estimated real-time loads and users' QoS requirements, the edge server will allocate a reasonable number of resources to each air interface. The edge server will also report partial loads and solutions to the cloud layer. For some services with strict QoS/QoE requirements or dramatic densities, the cloud will dynamically release more resources for the edge servers according to demands.

Delay-sensitive and real-time services will occupy a large proportion of the 6G network. To meet the QoS/QoE requirements for these services, we expect that continuous (in terms of time, frequency [52], space, power, computing, caching [53], or other resource domains) yet stable resource blocks will be allocated to them, as shown in Figure 7. Then, an evident consequence will be the existence of resource fragments, which exhibit strong heterogeneousness. The identification of these heterogeneous fragments includes two folds: multi-scale resource sensing and non-uniform resource sampling. The multi-scale characteristic emphasizes the diversities of methods (e.g., satellites, UAVs, robots, smartphones) and content (different resource domains) as well as accuracy (in meters, even decimeters) and timeliness (real-time or intermittent) in resource sensing. The non-uniform characteristic leads to higher requirements for heterogeneous resource analysis and matching. These fragments, if well organized, can effectively support many services that do not have stringent QoS/QoE requirements. Based on the above discussion, a potential strategy is to build a database for users' resource-usage patterns that reflects the users' RAT types and major service categories and maintain the resource-fragment map along with multiple resource domains. Then, a quick matching algorithm between the users and the resource-fragment map with relatively low complexity can be designed. Another topic is determining how to efficiently impose monitoring on the resource status, including the correlation of available resources, network-load estimation, as mentioned above, and even the anomaly detection [54,55]. Overall, we predict that the design of fast and efficient algorithms for the intelligent architecture will become a high-priority task to deal with the explosively increasing computation loads in the densely deployed heterogeneous 6G network.

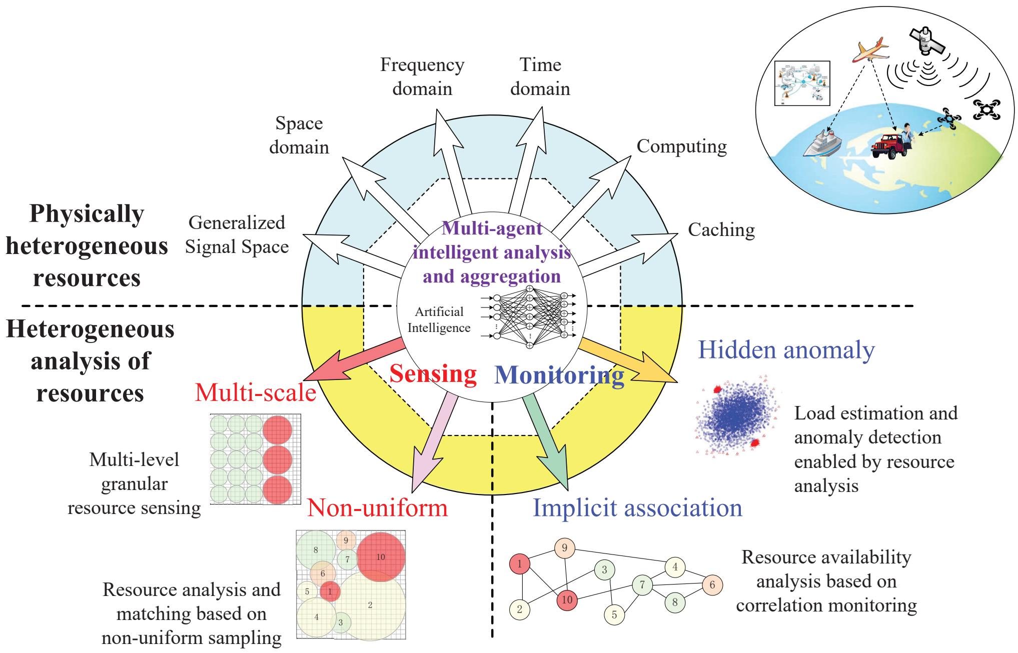

Figure 7. Heterogeneous resource analysis and aggregation via intelligent sensing and monitoring.

4.3. Service-Aware Transmission over Heterogeneous RATs

In the 5G network, users can switch connections between the NR and LTE base station depending on which one can provide better QoS. As mentioned previously, in the heterogeneous 6G network, a mobile user can connect via multiple air interfaces (different RATs) to receive single or multiple services from the cloud, and this will be more complicated than the switch in 5G. We have discussed the tightly coupled protocol design in the above, and it has the potential to support more frequent and diverse switching. If the end user is requesting multiple services, each service can be simply allocated to one RAT that is separated from the others. Clearly, the challenge lies in the attempt to obtain a single service from multiple RATs. This goal has been persistently tackled by the research community. The major problem is packet reordering, as illustrated in Figure 8a. Data packets belonging to a single service might come from different RRHs due to the diversified RATs present in the 6G network. However, the packet sequence is not consecutive in each path, and thus, packet reordering in the mobile user is needed for data recovery.

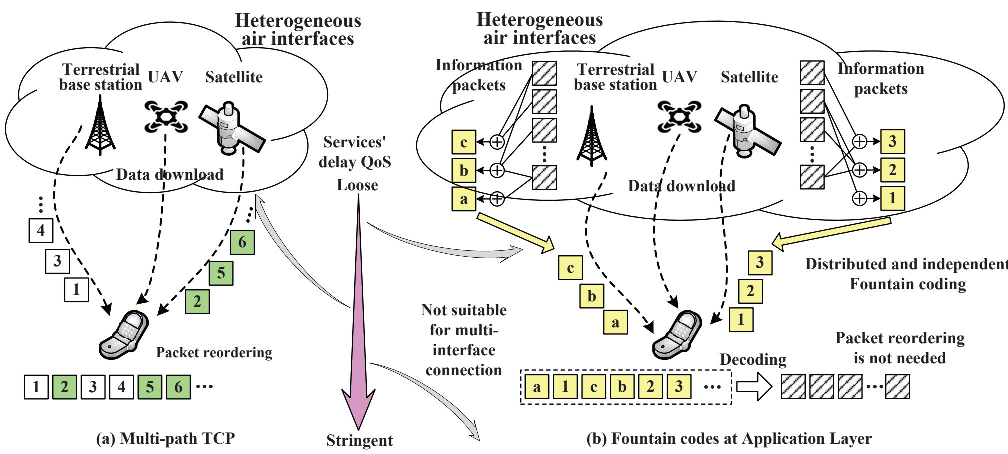

Figure 8. Service-aware access of heterogeneous RATs: the multi-path TCP approach and fountain-code-based approach.

One of the most popular approaches to support multi-path transmission is the multi-path TCP [56], as depicted in Figure 8a. The essential issue in multi-path TCP is the disharmony between the TCP's error control mechanism and the independent channel-quality/path-bandwidth of different paths. The TCP typically maintains a sliding window for the transmitter and receiver. At the receiver end, if a packet is lost or delayed, causing timeout, retransmission of all packets with proceeding sequence numbers, even some of which have correctly arrived, will be requested. This problem may occur often in the 6G network because the heterogeneous RATs have independent channel fading and congestion statuses. Moreover, the imbalanced bandwidths across different RATs might severely aggravate this problem. Aside from the multi-path TCP, the P2P approach via a consistent hash table and resource discovery and cooperation, might be an alternative approach. However, it is not suitable for the 6G RAN, as it requires cooperation across multiple paths, yielding a considerable overhead and thus not making itself a good candidate.

We next show that the fountain code has the potential to solve this multi-path data aggregation problem. The fountain code was originally developed for multi-cast common data from distributed servers on the Internet [57]. It is able to generate an infinite encoded sequence from given source symbols. Under ideal conditions, the source symbols can be recovered by simply obtaining an arbitrary subset of encoded symbols whose size is not less than that of the source symbols. The distributed servers can independently generate fountain-coded packets, which are random linear XOR combinations of a block of equal-length information packets. For usage in decoding, the random combination pattern is put in the header of each coded packet. The decoding takes advantage of the low-complexity iterative algorithm, and more importantly, the order of coded packets from distributed servers is not concerned by receivers. As long as a receiver can sufficiently accumulate many coded packets, it can recover the packet block successfully. It is notable that for the large block size (i.e., the number of information packets in a block), the optimized design of the distribution of random combination patterns will introduce only a very small amount of extra redundancy compared with the size of the information packets.

These benefits make the fountain code a strong candidate to increase the speed of delivery for single service over multiple air interfaces with different RATs. Multiple paths are analogous to the distributed servers mentioned above. As shown in Figure 8b, the coded packets from different RATs are not restricted by their arrival order. Furthermore, unlike the multi-path TCP, the decoding of the fountain code concentrates on the accumulation of coded packets and is not affected by the bandwidth imbalance across multiple RATs.

A drawback of the fountain-code-based approach is the need for a large block size to achieve a high efficiency. Thus, it is mainly suitable for delay-tolerant services, such as software updating and file downloading, which further motivates the achievement of service-aware access for heterogeneous RATs. Particularly, determining whether to enable connections via multiple RATs or not is driven by the delay quality-of-service (QoS) required by the user. As illustrated in Figure 8, when the delay-QoS requirement is very loose, the fountain-code-based approach is enabled. In contrast, when the delay-QoS requirement gets more stringent, the multi-path TCP approach will be activated, such that the timeout setup for the sliding window can assist in the delay assurance to some extent. In addition, if the service is very delay-sensitive, the RAN has to drop multi-RAT connections to avoid a long delay. Note that after determining the specific access mode for the RATs, the number of allocated resources depends on computation by the cloud or edge servers.

However, multi-path transmission will also introduce or aggravate the problem of co-site interference, in which a single device served by multiple RATs may receive different types of signals with diverse time, frequency, or space domain characteristics in a certain communication period, and this will eventually severely degrade the QoS. As interference introduced by accessing the LTE and 5G air interfaces in 5G networks simultaneously, co-site interference becomes more serious and significant because of the integrated communication scenarios and heterogeneous air interfaces in 6G networks. This issue has gained attention from many researchers in terms of determining how to effectively identify interference and extract useful signals [58,59]. Specifically, the digital-extraction cancellation scheme can overcome the weakness of the analog scheme caused by the limited accuracy of multiple RF components, typically by building a signal model to reconstruct the interference and then offsetting the interference. Based on the above discussion, determining how to guarantee the availability of the generated signal model and channel model while considering the protean parameters and coupled signals is a key problem. One possible solution is to insert some additional information into the pilot sequences, and when the signals reach the receiver, the known sequence fragments can assist to judge the distortion caused by interference and noise in the transmission channel. More discussion on the design of pilot sequences considering different frame structures and heterogeneous RATs is needed.

4.4. Enabling User Socialization for Cloud Extension

Regardless of how robust and delicate the design architecture is, the provision of seamless high-quality services for users in various environments and locations is still extremely difficult. One promising solution is to enable user-centric networking in the 6G network, where the user plays the role of not only the consumer but also the provider, such that the cloud-based architecture of RAN can be extended to the user level [6,30,60–63]. Device-to-device (D2D) communications provide strong support to achieve this goal, as they allow a device to relay the data to other devices in proximity, effectively benefiting the dead-corner coverage, data sharing, and even the handover in the RAN of 6G networks [64].

One typical example is atomized caching in mobile edge computing (MEC). As shown in Figure 9, the encoded packages in a file are dispersed at the edge of the network like fog, and the content requester is able to obtain the required file through D2D communication with the surrounding caching servers to carry out content sharing. Enabled by fountain codes, as mentioned in Section 4.3, the content requester does not need to figure out the associations between the encoded packets and received packets and only needs to concentrate on the number of received packets. Thus, with the assistance of multiple caching servers, a single requester will have the ability to download at higher speeds, while on the other hand, it can also act as a caching provider for other users.

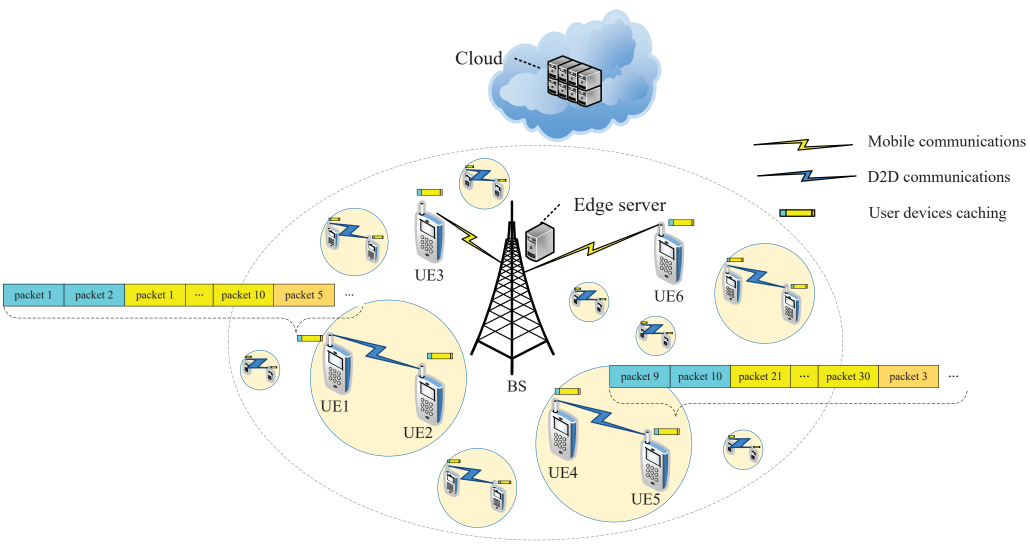

Figure 9. Atomized caching enabled by fountain codes with support by D2D communications and user socialization.

Unfortunately, recently, the standardization of D2D has progressed slowly due to a lack of economic incentives for both service providers and mobile users. To solve this problem, data relay among users requires some incentive [65]. In contrast, the beneficiaries need to pay the helper in some way, enabling socialization across end users [31]. This can also be easily integrated into the cloud architecture from the provider side. However, there are still many open problems to be researched and studied, including how to define and manage the payment, whether part of the payment should be forwarded to the provider, how the payment is associated with service types, etc.

5. Future Research Areas for the Service-Aware Intelligent RAN

5.1. Model Selection and Training for Multi-Level AI

With the explosive growth in the number of users and the types of users in the 6G era, the computing resources in each layer are always limited. This is especially obvious in some widely distributed terminals, e.g., IoT sensor nodes [66,67]. Moreover, no matter whether serving for a so-called latency-sensitive application or not, the communication algorithms are always strictly required on the delay performances. Typically, users do not want their requests to be answered when they no longer need them. This makes the time complexity of the algorithm a critical issue in model selection. Most of the proposed algorithms rely on databases and pre-training. Thus, in the future, even in the segmented communication field, we hope that the models will achieve a considerable degree of universality and adaptability.

5.2. Service-Resource Matching via Big Data Flying in the Air

With the fast-growing network size, the 6G network imposes overwhelming computational burden in the cloud. Rather than brutally computing for optimized resource allocation, it is desirable for the cloud to gain knowledge from the big data generated from networking [31,68], such that the quick service-resource matching pattern can be identified. However, it is worth noting that unlike the wired networks, the big data for mobile networks are often dispersed and flying over the air. Therefore, how to take advantage of this feature for deep learning will intrigue a very interesting research topic.

5.3. Service-Aware Security Assurance for User-Assisted Cloud

In traditional model updates, user data are necessary to optimize user experiences. However, people are reluctant to provide important personal information to the cloud, even if this can help to achieve more refined services through big data analysis. Federated learning is a new security strategy that enables local updates of models without uploading data to the cloud [69]. This will introduce a significant communication overhead during training; thus, some devices with weak computing and storage capabilities may not support federated learning well. Another potential PHY layer strategy is to generate a secret key by exploiting the noise channel and use the key as a one-time pad or several-time pad to ensure secure transmission [70]. Therefore, we should customize special security policies for different services, fully considering the diversities of hardware, resources, and even the external environment [71].

5.4. Deeper Virtualization down to the Physical Layer

Though the definition of network slicing has been proposed in 5G networks, it unfortunately still mainly concentrates on the resources of the upper layers, i.e., data and service slicing [72]. The slicing technique in the 6G network, which is supported by cloud computing, will have the ability to achieve deeper and more refined slicing for wireless links and air interfaces in the PHY and MAC layers, i.e., RAN slicing [73]. As the PHY abilities of base stations and terminals are further abstracted, the trend of network function virtualization (NFV) will become more obvious and finally, “PHY function as a service” will be achieved.

5.5. Concise Configuration in Intent-Driven Networks

In traditional networks, on one hand, a single network function can usually serve many core applications, but in the 6G network, the vertical trend of services calls for the ability to carry out customized networking. On the other hand, configuring networks manually inevitably leads to low efficiency and late responses, consequently increasing the management costs and expenditure for operators. To fully embrace the diversity of 6G services, intent-driven networks (IDNs) have shown huge potential to intelligently sense and respond to not only users' and operators' requirements but also their intent, including optimizing, monitoring, configuring, verifying, etc. [74,75]. Though IDN are hopeful candidates to truly fulfill the vision of intelligent networking, which have received wide attention from researchers, the concrete architecture and technique schemes of IDNs still require further development.

6. Conclusions

In this article, we first illustrated the concept of service-awareness and then briefly previewed the evolution of cloud-based RAN architectures from 5G to 6G. After analyzing the advantages and issues of the existing architectures, we presented an intelligent RAN architecture for service-aware access in 6G networks. We also shared our opinions on the issues and potential solutions for the service-aware resource configuration from network-protocol, resource-control, heterogeneous-RAT, and user socialization perspectives. Future research efforts will involve the selection and training of ML models, deep learning via big data, service-aware security assurance, deeper virtualization, and intent-driven networks for the cloud-based RAN of 6G networks. This article provides a roadmap for researchers working on hierarchical cloud and service-aware heterogeneous resource configuration for 6G radio access networks.

Author Contributions: Conceptualization, Z.Z. and Q.D.; Formal analysis, Q.D.; Investigation, H.S.; Writing—original draft, Z.Z., Q.D., D.W. and X.T.; Writing—review and editing, Q.D., D.W., X.T. and H.S. All authors have read and agreed to the published version of the manuscript.

Funding: The reported research work in this paper is supported by National Key Research and Development Program of China under Grant No. 2020YFB1807700.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Granelli, F.; Gebremariam, A.A.; Usman, M.; Cugini, F.; Stamati, V.; Alitska, M.; Chatzimisios, P. Software Defined and Virtualized Wireless Access in Future Wireless Networks: Scenarios and Standards. *IEEE Commun. Mag.* **2015**, *53*, 26–34. [CrossRef]
2. Soldani, D.; Manzalini, A. Horizon 2020 and Beyond: On the 5G Operating System for a True Digital Society. *IEEE Veh. Technol. Mag.* **2015**, *10*, 32–42. [CrossRef]
3. Marsch, P.; Raaf, B.; Szufarska, A.; Mogensen, P.; Guan, H.; Färber, M.; Redana, S.; Pedersen, K.; Kolding, T. Future Mobile Communication Networks: Challenges in the Design and Operation. *IEEE Veh. Technol. Mag.* **2012**, *7*, 16–23. [CrossRef]
4. Demestichas, P.; Georgakopoulos, A.; Karvounas, D.; Tsagkaris, K.; Stavroulaki, V.; Lu, J.; Xiang, C.; Yao, J. 5G on the Horizon: Key Challenges for the Radio-Access Network. *IEEE Veh. Technol. Mag.* **2013**, *8*, 47–53. [CrossRef]
5. Dahrouj, H.; Douik, A.; Dhifallah, O.; Al-Naffouri, T.Y. Resource Allocation in Heterogeneous Cloud Radio Access Networks: Advances and Challenges. *IEEE Wirel. Commun.* **2015**, *22*, 66–73. [CrossRef]
6. Jo, M.; Maksymyuk, T.; Strykhalyuk, B.; Cho, C.-H. Device-to-device-based Heterogeneous Radio Access Network Architecture for Mobile Cloud Computing. *IEEE Wirel. Commun.* **2015**, *22*, 50–58. [CrossRef]
7. Ran, C.; Wang, S.; Wang, C. Balancing Backhaul Load in Heterogeneous Cloud Radio Access Networks. *IEEE Commun. Mag.* **2015**, *22*, 42–48. [CrossRef]
8. China Mobile Research Institute. *C-RAN the Road towards Green Ran*; Technical Report; China Mobile Research Institute: Beijing, China, 2011.
9. Checko, A.; Christiansen, H.L.; Yan, Y.; Scolari, L.; Kardaras, G.; Berger, M.S.; Dittmann, L. Cloud RAN for Mobile Networks-A Technology Overview. *IEEE Commun. Surv. Tutor.* **2014**, *17*, 405–426. [CrossRef]
10. Zhang, N.; Cheng, N.; Gamage, A.T.; Zhang, K.; Mark, J.W.; Shen, X. Cloud Assisted Hetnets toward 5G Wireless Networks. *IEEE Commun. Mag.* **2015**, *53*, 59–65. [CrossRef]
11. Peng, M.; Li, Y.; Jiang, J.; Li, J.; Wang, C. Heterogeneous Cloud Radio Access Networks: A New Perspective for Enhancing Spectral and Energy Efficiencies. *IEEE Wirel. Commun.* **2014**, *21*, 126–135. [CrossRef]
12. Chen, S.; Zhao, J. The Requirements, Challenges, and Technologies for 5G of Terrestrial Mobile Telecommunication. *IEEE Commun. Mag.* **2014**, *52*, 36–43. [CrossRef]
13. Chen, S.; Hu, J.; Shi, Y.; Zhao, L.; Li, W. A Vision of C-V2X: Technologies, Field Testing, and Challenges with Chinese Development. *IEEE Internet Things J.* **2020**, *7*, 3872–3881. [CrossRef]

14. Zhou, Z.; Abawajy, J.; Shojafar, M.; Chowdhury, M. DEHM: An Improved Differential Evolution Algorithm using Hierarchical Multi-strategy in a Cybertwin 6G Network. *IEEE Trans. Ind. Inform.* **2022**, *99*, 1. [CrossRef]
15. Chen, S.; Liang, Y.-C.; Sun, S.; Kang, S.; Cheng, W.; Peng, M. Vision, Requirements, and Technology Trend of 6G: How to Tackle the Challenges of System Coverage, Capacity, User Data-Rate and Movement Speed. *IEEE Wirel. Commun.* **2020**, *27*, 218–228. [CrossRef]
16. Zong, B.; Fan, C.; Wang, X.; Duan, X.; Wang, B.; Wang, J. 6G Technologies: Key Drivers, Core Requirements, System Architectures, and Enabling Technologies. *IEEE Veh. Technol. Mag.* **2019**, *14*, 18–27. [CrossRef]
17. Letaief, K.B.; Chen, W.; Shi, Y.; Zhang, J.; Zhang, Y.-J.A. The Roadmap to 6G: AI Empowered Wireless Networks. *IEEE Commun. Mag.* **2019**, *57*, 84–90. [CrossRef]
18. Chen, F.; Jiang, Z.; Ji, F.; Yu, H.; Guan, Q.; Liu, F. Radio-acoustic Integrated Network for Ocean Information Transmission: Framework and Enabling Technologies. *China Commun.* **2021**, *18*, 62–70. [CrossRef]
19. Cheng, N.; Quan, W.; Shi, W.; Wu, H.; Ye, Q.; Zhou, H.; Zhuang, W.; Shen, X.; Bai, B. Space-air-ground Network Integration: A Comprehensive Simulation Platform. *IEEE Wirel. Commun.* **2020**, *27*, 178–185. [CrossRef]
20. Hui, Y.; Cheng, N.; Huang, Y.; Chen, R.; Xiao, X.; Li, C.; Mao, G. Personalized Vehicular Edge Computing in 6G. *IEEE Netw.* **2021**, *35*, 278–284. [CrossRef]
21. You, X.; Wang, C.; Huang, J.; Gao, X.; Zhang, Z.; Wang, M.; Huang, Y.; Zhang, C.; Jiang, Y.; Wang, J.; et al. Towards 6G Wireless Communication Networks: Vision, Enabling Technologies, and New Paradigm Shifts. *Sci. China Inf. Sci.* **2021**, *64*, 1–74. [CrossRef]
22. Zhang, P.; Xu, X.; Qin, X.; Liu, Y.; Ma, N.; Han, S. Evolution Toward Artificial Intelligence of Things Under 6G Ubiquitous-X. *J. Harbin Inst. Technol. (New Ser.)* **2020**, *27*, 116–135.
23. Yang, P.; Xiao, Y.; Xiao, M.; Li, S. 6G Wireless Communications: Vision and Potential Techniques. *IEEE Netw.* **2019**, *33*, 70–75. [CrossRef]
24. Li, B.; Fei, Z.; Chu, Z.; Zhou, F.; Wong, K.-K.; Xiao, P. Robust Chance-Constrained Secure Transmission for Cognitive Satellite–Terrestrial Networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 4208–4219. [CrossRef]
25. Zhang, R.; Zhou, Y.; Lu, X.; Cao, C.; Guo, Q. Antenna De-embedding for mmWave Propagation Modeling and Field Measurement Validation at 73 GHz. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 3648–3659. [CrossRef]
26. Liu, L.; Zhou, Y.; Yuan, J.; Zhuang, W.; Wang, Y. Economically Optimal MS Association for Multimedia Content Delivery in Cache-Enabled Heterogeneous Cloud Radio Access Networks. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1584–1593. [CrossRef]
27. Wang, B.; Gao, F.; Jin, S.; Lin, H.; Li, G.Y. Spatial- and Frequency-Wideband Effects in Millimeter-Wave Massive MIMO Systems. *IEEE Trans. Signal Process.* **2018**, *66*, 3393–3406. [CrossRef]
28. Sun, L.; Du, Q. A Review of Physical Layer Security Techniques for Internet of Things: Challenges and Solutions. *Entropy* **2018**, *20*, 730. [CrossRef]
29. Yang, X.; Shu, L.; Liu, Y.; Hancke, G.P.; Ferrag, M.A.; Huang, K. Physical Security and Safety of IoT Equipment: A Survey of Recent Advances and Opportunities. *IEEE Trans. Ind. Inform.* **2022**. [CrossRef]
30. Wang, Y.; Lin, X. User-Provided Networking for QoE Provisioning in Mobile Networks. *IEEE Wirel. Commun.* **2015**, *22*, 26–33. [CrossRef]
31. Su, Z.; Xu, Q.; Qi, Q. Big Data in Mobile Social Networks: A QoE-oriented Framework. *IEEE Netw.* **2016**, *30*, 52–57. [CrossRef]
32. Cheng, W.; Zhang, W.; Jing, H.; Gao, S.; Zhang, H. Orbital Angular Momentum for Wireless Communications. *IEEE Wirel. Commun.* **2019**, *26*, 100–107. [CrossRef]
33. Patterson, R.E.; Eggleston, R.G. Human–Machine Synergism in High-Level Cognitive Functioning: The Human Component. *IEEE Trans. Emerg. Top. Comput. Intell.* **2018**, *2*, 249–257. [CrossRef]
34. Wiltshire, T.J.; Fiore, S.M. Social Cognitive and Affective Neuroscience in Human–Machine Systems: A Roadmap for Improving Training, Human–Robot Interaction, and Team Performance. *IEEE Trans.-Hum.-Mach. Syst.* **2014**, *44*, 779–787. [CrossRef]
35. Braten, A.E.; Kraemer, F.A.; Palma, D. Autonomous IoT Device Management Systems: Structured Review and Generalized Cognitive Model. *IEEE Internet Things J.* **2021**, *8*, 4275–4290. [CrossRef]
36. Nguyen, H.X.; Trestian, R.; To, D.; Tatipamula, M. Digital Twin for 5G and Beyond. *IEEE Commun. Mag.* **2021**, *59*, 10–15. [CrossRef]
37. Park, S.-M.; Kim, Y.-G. A Metaverse: Taxonomy, Components, Applications, and Open Challenges. *IEEE Access* **2022**, *10*, 4209–4251. [CrossRef]
38. Mei, J.; Wang, X.; Zheng, K. An Intelligent Self-Sustained RAN Slicing Framework for Diverse Service Provisioning in 5G-Beyond and 6G Networks. *Intell. Converg. Netw.* **2020**, *1*, 281–294. [CrossRef]
39. Al-Rahamneh, A.; Astrain, J.J.; Villadangos, J.; Klaina, H.; Guembe, I.P.; Lopez-Iturri, P.; Falcone, F. Enabling Customizable Services for Multimodal Smart Mobility with City-Platforms. *IEEE Access* **2021**, *9*, 41628–41646. [CrossRef]
40. Sun, L.; Wang, Y.; Swindlehurst, A.L.; Tang, X. Generative-Adversarial-Network Enabled Signal Detection for Communication Systems with Unknown Channel Models. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 47–60. [CrossRef]
41. Tang, X.; Wang, D.; Zhang, R.; Chu, Z.; Han, Z. Jamming Mitigation via Aerial Reconfigurable Intelligent Surface: Passive Beamforming and Deployment Optimization. *IEEE Trans. Veh. Technol.* **2021**, *70*, 6232–6237. [CrossRef]
42. Tang, X.; Ren, P.; Gao, F.; Du, Q. Interference-Aware Resource Competition Toward Power-Efficient Ultra-Dense Networks. *IEEE Trans. Commun.* **2017**, *65*, 5415–5428. [CrossRef]

43. Shen, J.; Yang, C.; Li, T.; Wang, X.; Song, Y.; Guizani, M. Interactive Artificial Intelligence Meets Game Theory in Next-Generation Communication Networks. *IEEE Wirel. Commun.* **2021**, *28*, 128–135. [CrossRef]
44. You, T.; Zhang, H.; Zhang, Y.; Li, Q.; Zhang, P.; Yang, M. The Influence of Experienced Guider on Cooperative Behavior in the Prisoner's Dilemma Game. *Appl. Math. Comput.* **2022**, *426*, 127093. [CrossRef]
45. Han, Z.; Niyato, D.; Saad, W.; Başar, T.; Hjørungnes, A. *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*; Cambridge University Press: Cambridge, UK, 2012. [CrossRef]
46. Saad, M.M.; Khan, M.T.R.; Shah, S.H.A.; Kim, D. Advancements in Vehicular Communication Technologies: C-V2X and NR-V2X Comparison. *IEEE Commun. Mag.* **2021**, *59*, 107–113. [CrossRef]
47. Zhou, L.; Wu, D.; Chen, J.; Wei, X. Cross-Modal Collaborative Communications. *IEEE Wirel. Commun.* **2020**, *27*, 112–117.
48. Yao, Z.; Cheng, W.; Zhang, W.; Zhang, H. Resource Allocation for 5G-UAV-Based Emergency Wireless Communications. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3395–3410. [CrossRef]
49. Liu, L.; Zhou, Y.; Garcia, V.; Tian, L.; Shi, J. Load Aware Joint CoMP Clustering and Inter-Cell Resource Scheduling in Heterogeneous Ultra Dense Cellular Networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 2741–2755. [CrossRef]
50. Li, B.; Fei, Z.; Xu, X.; Chu, Z. Resource Allocations for Secure Cognitive Satellite-Terrestrial Networks. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 78–81. [CrossRef]
51. Zhao, Z.; Du, Q.; Sun, L. Network-Load Estimation for K-Repetition Grant-Free Access Enabling Adaptive Resource Allocation Towards QoS Enhancement. In Proceedings of the 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, 13–16 September 2021. [CrossRef]
52. Zhao, H.; Ding, K.; Sarkar, N.I.; Wei, J.; Xiong, J. A Simple Distributed Channel Allocation Algorithm for D2D Communication Pairs. *IEEE Trans. Veh. Technol.* **2018**, *67*, 10960–10969. [CrossRef]
53. Li, L.; Wang, M.; Xue, K.; Cheng, Q.; Wang, D.; Chen, W.; Pan, M.; Han, Z. Delay Optimization in Multi-UAV Edge Caching Networks: A Robust Mean Field Game. *IEEE Trans. Veh. Technol.* **2021**, *70*, 808–819.
54. Hussain, B.; Du, Q.; Sun, B.; Han, Z. Deep Learning-Based DDoS-Attack Detection for Cyber–Physical System over 5G Network. *IEEE Trans. Ind. Inform.* **2021**, *17*, 860–870. [CrossRef]
55. Hussain, B.; Du, Q.; Imran, A.; Imran, M.A. Artificial Intelligence-Powered Mobile Edge Computing-Based Anomaly Detection in Cellular Networks. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4986–4996. [CrossRef]
56. Leung, K.-C.; Li, V.O.K.; Yang, D. An Overview of Packet Reordering in Transmission Control Protocol (TCP): Problems, Solutions, and Challenges. *IEEE Trans. Parallel Distrib. Syst.* **2007**, *18*, 522–535. [CrossRef]
57. Byers, J.W.; Luby, M.; Mitzenmacher, M. A Digital Fountain Approach to Asynchronous Reliable Multicast. *IEEE J. Sel. Areas Commun.* **2002**, *20*, 1528–1540. [CrossRef]
58. Ge, S.; Meng, J.; Xing, J.; Liu, Y.; Gou, C. A Digital-Domain Controlled Nonlinear RF Interference Cancellation Scheme for Co-Site Wideband Radios. *IEEE Trans. Electromagn. Compat.* **2019**, *61*, 1647–1654. [CrossRef]
59. Ge, S.; Xing, J.; Liu, Y.; Liu, H.; Li, Y.; Meng, J. Dual-Stage Co-Site RF Interference Canceller for Wideband Direct-Conversion Receivers Using Reduced Observation Chain. *IEEE Trans. Electromagn. Compat.* **2020**, *62*, 923–932. [CrossRef]
60. Garcia, V.; Zhou, Y.; Shi, J.L. Coordinated Multipoint Transmission in Dense Cellular Networks with User-Centric Adaptive Clustering. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 4297–4308. [CrossRef]
61. Cheng, W.; Zhang, X.; Zhang, H. Optimal Power Allocation with Statistical QoS Provisioning for D2D and Cellular Communications over Underlaying Wireless Networks. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 151–162. [CrossRef]
62. Du, Q.; Song, H.; Zhu, X. Social-Feature Enabled Communications among Devices toward the Smart IoT Community. *IEEE Commun. Mag.* **2019**, *57*, 130–137. [CrossRef]
63. Zhou, Y.; Liu, H.; Pan, Z.; Tian, L.; Shi, J.L. Energy Efficient Two-stage Cooperative Multicast Based on Device to Device Transmissions: Effect of User Density. *IEEE Trans. Veh. Technol.* **2016**, *65*, 7297–7307. [CrossRef]
64. Zhou, L.; Wu, D.; Chen, J.; Dong, Z. Greening the Smart Cities: Energy-Efficient Massive Content Delivery via D2D Communications. *IEEE Trans. Ind. Inform.* **2018**, *14*, 1626–1634. [CrossRef]
65. Gao, Y.; Xiao, Y.; Wu, M.; Xiao, M.; Shao, J. Dynamic Social-Aware Peer Selection for Cooperative Relay Management with D2D Communications. *IEEE Trans. Commun.* **2019**, *67*, 3124–3139. [CrossRef]
66. Hassan, M.R.; Hassan, M.M.; Altaf, M.; Yeasar, M.S.; Hossain, M.I.; Fatema, K.; Shaharin, R.; Ahmed, A.F. B5G-Enabled Distributed Artificial Intelligence on Edges for COVID-19 Pandemic Outbreak Prediction. *IEEE Netw.* **2021**, *35*, 48–55. [CrossRef]
67. Wang, Z.; Ogbodo, M.; Huang, H.; Qiu, C.; Hisada, M.; Abdallah, A.B. AEBIS: AI-Enabled Blockchain-Based Electric Vehicle Integration System for Power Management in Smart Grid Platform. *IEEE Access* **2020**, *8*, 226409–226421. [CrossRef]
68. Xia, Q.; Xu, Z.; Liang, W.; Yu, S.; Guo, S.; Zomaya, A.Y. Efficient Data Placement and Replication for QoS-Aware Approximate Query Evaluation of Big Data Analytics. *IEEE Trans. Parallel Distrib. Syst.* **2019**, *30*, 2677–2691. [CrossRef]
69. Sattler, F.; Wiedemann, S.; Müller, K.-R.; Samek, W. Robust and Communication-Efficient Federated Learning from Non-i.i.d. Data. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 3400–3413. [CrossRef]
70. Sun, L.; Du, Q. Physical Layer Security with Its Applications in 5G Networks: A Review. *China Commun.* **2017**, *14*, 1–14. [CrossRef]
71. Sookhak, M.; Tang, H.; He, Y.; Yu, F.R. Security and Privacy of Smart Cities: A Survey, Research Issues and Challenges. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 1718–1743. [CrossRef]
72. Lv, Z.; Xiu, W. Interaction of Edge-Cloud Computing Based on SDN and NFV for Next Generation IoT. *IEEE Internet Things J.* **2020**, *7*, 5706–5712. [CrossRef]

73. Jiang, Z.; Chen, X.; Ma, J.; Zhang, Y.; Gu, J. Traffic Dynamics Evaluation for the Future NFV Deployment. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 2214–2218. [CrossRef]
74. Ouyang, Y.; Yang, C.; Song, Y.; Mi, X.; Guizani, M. A Brief Survey and Implementation on Refinement for Intent-Driven Networking. *IEEE Netw.* **2021**, *35*, 75–83. [CrossRef]
75. Pang, L.; Yang, C.; Chen, D.; Song, Y.; Guizani, M. A Survey on Intent-Driven Networks. *IEEE Access* **2020**, *8*, 22862–22873. [CrossRef]