Author(s): Hu, Yun; Chang, Zheng; Chen, Yanhui; Han, Zhu

Title: Service-Oriented Wireless Virtualized Networks : An Intelligent Resource Management Approach

Year: 2022

Version: Accepted version (Final draft)

Please cite the original version:

Hu, Y., Chang, Z., Chen, Y., & Han, Z. (2022). Service-Oriented Wireless Virtualized Networks : An Intelligent Resource Management Approach. IEEE Vehicular Technology Magazine, 17(1), 57-65. https://doi.org/10.1109/MVT.2021.3114651

Service-Oriented Wireless Virtualized Networks : An Intelligent Resource Management Approach

Yun Hu, *Member, IEEE,* Zheng Chang, *Senior Member, IEEE,* Yanhui Chen, *Member, IEEE,* and Zhu Han, *Fellow, IEEE*

***Abstract*—It can be anticipated that the future wireless network will be composed of highly heterogeneous resources and experience explosively growing demands for mobile traffic services. Many challenges arise on how to achieve an efficient operation from the network management point-of-view to utilize heterogeneous resources for a large variety of services over a significant number of end-users. By the idea of virtualizing wireless access and resources, wireless network virtualization is proposed as the main evolution, which attracts increasing attention. In this work, the concepts and the challenges for realizing dynamic and intelligent resource management in wireless virtualized networks (WVNs) are discussed. A deep reinforcement learning framework is proposed as an example to study how to utilize resources and provide an efficient management scheme in WVNs. Future research directions and opportunities are also highlighted to shed the light towards an efficient WVN paradigm.**

I. Introduction

With an ever-increasing demand for broadband communications, new challenges for enhancing quality of experience (QoE) of the end-user keep rising, such as intelligent and ubiquitous high-speed connectivity, efficient and flexible allocation of resources, etc. It can be anticipated that the next generation wireless networks will boost the development of many industries, ranging from mobiles networks, to entertainment, manufacturing, vehicles, healthcare, agriculture and will provide an enhanced QoE to its subscribers. To realize the vision of essentially unlimited access to data services for anyone and anything, the recent emerging network platforms, such as Software Defined Network (SDN) and Network Function Virtualization (NFV), urges us to redesign the cellular-based network infrastructure [1]. Incorporating with the wireless network, the advanced SDN/NFV architectures are applied to Radio Access Networks (RANs), which creates the Wireless Virtualized Networks (WVNs) framework. In the WVN, RAN functions are performance on the commoditized network platforms owned by multiple Infrastructure Providers (InPs), instead of dedicated telecom hardware owned by single operator. In particular, the concept of virtualization allows the customized WVNs for particular applications on top of a physical network. The WVN makes physical infrastructure and radio resources being abstracted, sliced and shared, which makes them well suited to address the diverse requirements of future wireless network and ease the network management. After virtualization, the virtual slices containing radio resources are offered to the service provider based on their demands. The mobile operator and service provider can rent the virtual resources, instead of owning them, to provide services to the users. Consequently, the overall expenses of network deployment and operation can be significantly decreased [2].

While the WVN may lower the standard for becoming a network operator, opportunities come with challenges, including isolation, control signalling, resource discovery and allocation, mobility management, network management and operation, and some non-technical issues etc [3]. Due to the fact that the future network is expected to be highly heterogeneous with ultra-density, it is natural to consider how the resource efficiently utilized in an optimized manner. These challenges need to be tackled broadly and carefully by extensive and dedicated research efforts. Apparently, it is of profoundly significance to explore the interactions and relations among all the involved parties so as to properly manage the limited resources with the different services. As the InPs are heterogeneous with different preference towards joining in and providing services, in terms of operation cost, energy consumption, available spectrum, etc., there is a need to design a sophisticated mechanism in which Mobile Virtual Network Operators (MVNOs) will be stimulated to rent and manage the resources of the InPs. To this end, we turn to utilize the intelligent approaches to design the interaction between the InPs and MVNO, and provide the efficient resource allocation scheme.

Recently, the design and development of WVN has received many research interests [4]. In particular, it can be found that resource allocation problems in the WVN attracts considerable interests. In order to minimize the operation cost of running a physical network of an InP, in [5], the authors propose a new formulation for the bandwidth allocation and routing problem for multiple WVNs that operate on top of the physical substrate network. In [6],

Y. Hu and Y. Chen are with National Demonstration Center for Experimental Electronic Information & Communications Engineering Education, Xidian University, Xi'an 710071, China. Z. Chang is with School of Computer Science and Engineering, University of Electric Science and Technology China, Chengdu, China and also with Faculty of Information Technology, University of Jyväskylä, P.O.Box 35, FIN-40014 Jyväskylä, Finland. Z. Han is with the Department of Electrical and Computer Engineering in the University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea, 446-701. This work is partly supported by NSFC (No. 62071105), NSF EARS-1839818, CNS1717454, CNS-1731424, and CNS-1702850. The corresponding author is Zheng Chang. Email: huyun@xidian.edu.cn, zheng.chang@uestc.edu.cn, yhchen@mail.xidian.edu.cn, zhan2@uh.edu

the authors investigate the WVN with device-to- device (D2D) communications and present resource allocation mechanism. The authors of [7] propose an efficient radio resource allocation algorithm for an Orthogonal Frequency Division Multiplexing (OFDM)-based WVN. In order to maximize the uplink energy efficiency, the authors propose to jointly optimize the usage of transmit power, subcarrier, and antenna, so as to determine the resource allocation in each virtual slice in WVN. In [8], the authors present a joint resource allocation and user association algorithm to optimize the system throughput in the downlink of WVN. Meanwhile, the minimum required rate of each Network Service Provider (NSP) is explicitly considered. In [9], an information-centric WVN is considered and the authors first define the system utility related to earned revenue from the users and cost paid to the InPs. Then, the authors propose a resource allocation scheme to maximize the system utility. Recently, along with the fast development of machine learning and artificial intelligence (AI), there are growing interests on applying AI for designing wireless networks [10]-[14]. The authors of [10]-[13] have utilized the deep reinforcement learning (DRL) scheme for addressing the resource allocation problems in different network scenarios. By using multi-agent DRL, the author of [14] present a virtual resource allocation schemes for virtualized IoT network. In [15], the authors tackle a general DRL-based resource allocation problem for network slicing which considers a mixed action space including both discrete channel allocation and continuous energy harvesting time division. A constrained discrete-continuous soft actor-critic (CDC-SAC) is used to redesign the network architecture and policy learning process.

In this article, we first briefly introduce the WVN concept. Then we discuss the WVN taxonomy and also present an overview of key enabling technologies and encountered challenges towards an mature and intelligent resource management in a service-oriented WVN framework. Then, we utilize the DRL framework and evaluate its performance and capability on designing resource allocation schemes for the WVN. Future directions and opportunities are discussed. We hope this article can attract interests on this emerging field and yield innovative solutions to the development of the intelligent resource allocation for WVN infrastructure.

II. Wireless Virtualized Network: Concept, Architecture and Advantages

Virtualization has mainly been studied for computing server to deal with the computing resources. While talking about the virtualization in the wireless networks, radio resources and physical infrastructure of the wireless network are abstracted and allocated into virtual slices with certain functionalities. The virtual slices are then ready to be shared by different parties after resource isolation [2]. Therefore, after the virtualization process, the resources will be provided to different Network Service Providers (NSPs). Then, the virtual resources are used for service provisioning.

Similar to the concept in computer virtualization, in the WVNs, physical wireless network infrastructure and radio resources, owned by different InPs, can be abstracted and sliced into virtual wireless network resources, and shared by multiple parties after a certain level of isolation [2]. After resource virtualization, the virtual resources can be offered to the end-users via the MVNO or dedicated NSPs. In Fig. 1, a brief illustration of WVN is presented from both logical and physical domains. The radio resources owned by the InPs and the MVNO acts as central controller coordinating the virtualized resources. When the NSPs receive the demand of service from the users, they can request the radio resources from InPs. Then, the radio resources and physical infrastructures owned by different InPs will be processed by the network controller. After being isolation, abstract and virtualization, the virtual slices containing radio resources are offered to the NSPs based on their demands. The MVNO is able to virtualize the physical resources according to the demand of NSPs. The users is able to logically connect to the virtual network through service subscription and communicates with the cellular network physically.

Moreover, wireless virtualization also has a wide range of potential benefits in both commercial and research contexts by enabling a flexible reuse of existing infrastructure and many other resources. In commercial markets, Capital Expenditure (CapEx) and Operating Expenditure (OpEx) can be reduced significantly due to the sharing enabled by WVN [2]. In addition, it is expected that the WVN is able to help ease and lower the doorsill of the network operation, which consequently attracts many parties to step into this wireless market to become the NSPs and/or InPs. The bloom of wireless market can further present diverse and flexible services to the end-users of all kinds, and provide more satisfied QoS. Nevertheless, the distinctive properties of the wireless environment make the resource management problem in WVN complicated. We will discuss how to address these challenges in the following section.

III. Research Challenges for Intelligent Resource Management in WVN

As mentioned, for the wireless virtualized network, realizing efficient, sophisticated and intelligent resource management scheme is still under-investigation confronts many challenges from the interaction of different network layers, network management, resource virtualization and allocation, and involved signalling issues. In the following, these challenges are elaborated together with their impacts.

A. *Virtual Resource Allocation*

Intuitively, virtual resource allocation is one of the most significant features of WVNs. Based on the service requests, virtual resource allocation should be able to abstract, virtualize and isolate the heterogeneous resources from various physical network sources. The steps of the

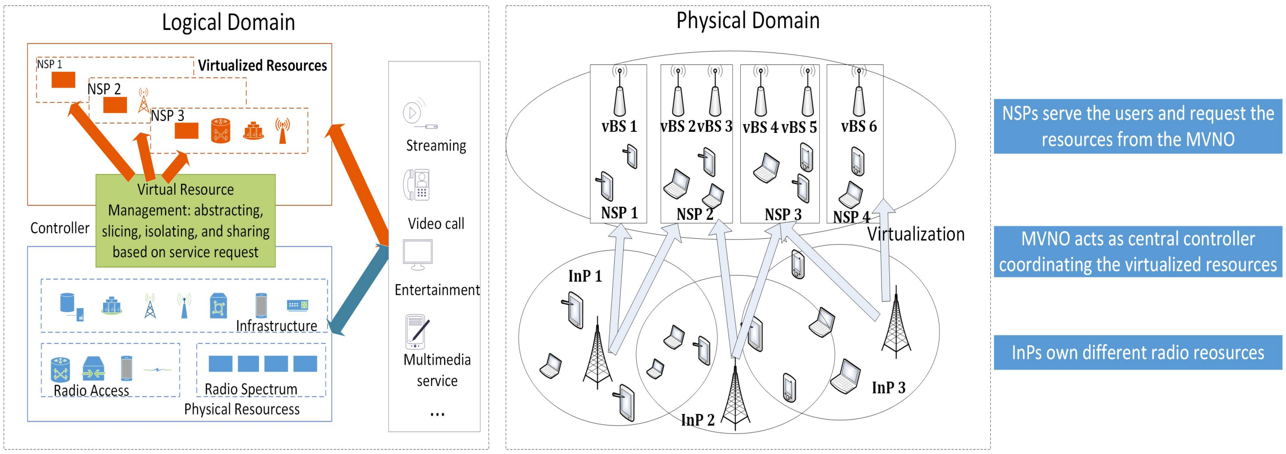

Fig. 1. An example of wireless network virtualization

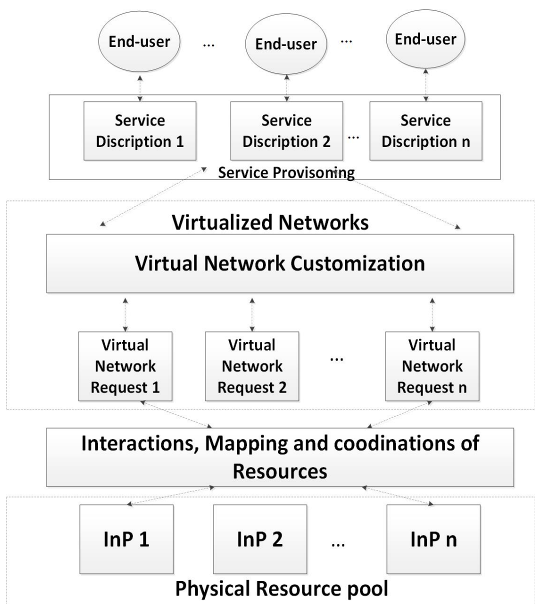

Fig. 2. Virtual resource allocation.

resource allocation are summarized in Fig. 2. The first step of resource allocation is to customize the WVN to become a service-oriented network. Accordingly, optimal network protocols and resource requirement should be decided in the customization to satisfy the end-to-end service quality.

Moreover, the service requirement, virtual resources and physical resources should be mapped, which is usually a NP-hard problem. At the InP side, the physical radio resources that are used for virtualization need to be decided. At the NSP side, what kind of resources to be chosen and what kind of cost it may occur should be determined depending on their subscriber's demands. The jobs of the MVNOs are to be aware of the available active and passive resources in the underlying physical cellular networks and ask for resources from the InPs. The network slicing will then be performed to realize the virtualization process and the virtualized resources will be provided to the NSPs. As we can see, to realize the efficient resource allocation in WVN, efficient coordination mechanism and interaction model are essential among the InP, MVNO and NSP, which calls for a contiguous efforts. In this context, all these three parties attempt to maximize their own benefits, so the dynamic adaptation and intelligent management of resources across different providers and operators should be investigated.

B. Network Management

The network management and deployment poses another significant challenge to the WVN to guarantee the efficient operations of the physical and virtual network and the quality of supported services. Since essentially the underlying physical networks are heterogeneous in nature and each of them may have unique properties, solutions and mechanisms should be dedicated to address these differences for the efficient maintenance, operation and service provisioning.

Specifically, end-users in the WVN should be able to switch to their contracted NSP smoothly whenever they want. The end-users in the perfect situation are able to access any NSP that can provide the best QoS wherever they are located. Thus, the network architecture and resource management need to facilitate the mobility management through resource sharing and protocols development among different involved parties to ensure that end-users can successfully access to the selected NSP.

Moreover, from the system perspective, the WVN can centrally control the physical resources of the InP, which brings an opportunity for efficient network operation. For example, when the coverage of several InPs are overlapped, or the demand of a certain area is low, effective network management of WVN allows to dynamically allocate radio resources, e.g., putting some base stations (BSs) of the InPs into the sleep mode, to save operation cost while maintaining the QoS. However, due to the heterogeneity and dynamics of user or service demand, the network management should be carefully designed. From perspective of the InP, how to optimally deploy their infrastructures and offer radio resources to the MVNO should be investigated. While for the MVNO, how to manage the physical resources and perform network slices to achieve efficient network operation and enhanced service quality calls for proposals from both algorithmic and implementation levels.

C. Control and Signalling

Obviously, the virtualization process and operation may generate a considerable amount of signalling overheads comparing with the traditional cellular networks. Specifically, the interactions among all the involved parties in intelligent management require additional effort on the design of exchanging signalling and control information. For example, the connections between the MVNO and InPs should be established before the virtualization starts. With this connectivity, both parties can express their preference and required different types of resources. In addition, since resource sharing may happen among InPs, a standard protocol for information sharing among the InPs is necessary. Thus, the current control and signalling protocols cannot satisfy the requirement during the virtualization process and need to be refined.

In addition, the control functions of WVNs should be enriched. To realize the automated operation of virtualization and resource allocation, the interplay of big data management and WVN controller needs to be investigated. Due to the large scale of WVNs, the abstraction, isolation, and virtualization procedures also require a frequent update to customize and adaptive to the stochastic fluctuation, variation of the wireless environment and end-user statues in a fast and dynamic way. However, due to the distinct nature of network entities and service requirement, the update of WVN procedures may vary over time and location, which results in inconsistency of the networks and poses significant challenges to the controller. Therefore, an efficient control plane with low latency is of profound importance to the resource management in WVN.

D. Big Data Management

There is no doubt that the future network will cope with massive data. The big amount of data is not limited to transmission data, but also contains the data from various sources to reflect the daily behaviors of network users and subscribers. Thanks to fast development of machine learning and data mining techniques, network-generated big data, such as user and service history log, user preference and traffic history, no longer only cause network resource consumption, but rather, can be viewed as an valuable sources to assist network operators implement and enhance the network services to massive devices with heterogeneous requirements.

As the WVN is based on the increased computational capabilities of network platform, managing and utilizing big data in WVN is full of possibility. However, there are still some challenges ahead. From the transmission level, all the wireless and wired links have the limitation of data rate, which is an inevitable bottleneck for big data transmission over the network. From the data processing point-of-view, many data in the wireless network are sparse, e.g., sparse channel and spare user activity, etc.. Therefore, how to address the sparsity in the data process is of significance. Moreover, what data, how to abstract data and how to process the data are essential for the NSP to deeply understand and predict the requirements of end-users and provide reliable services. Correspondingly, what to acquire and how to process and abstract the data to become meaningful information are considered as the main challenges for big data management and is important for effective network operation in WVNs.

IV. DRL-based Resource Management for WVN

As presented, there are several confront challenges on realizing the intelligent resource management for the WVN. In this section, we will adopt the Artificial Intelligence (AI) framework into the development of virtual resource management in WVN and explore its potential for virtual resource allocation and network management. The proposed AI-based virtual resource management scheme is to design a mechanism to maximize the utility of the MVNO and meanwhile enhance the satisfactions of the InPs.

In the considered system, there is one MVNO that needs to operate the rented resources from InPs in a virtualized way to provide services to multiple end-users. Over a certain amount of time slots, effective resource management should have the complete knowledge about the future time slots to reach the optimal solution of the next time slot. Therefore, absence of prior information about channel state and energy arrival may lead to a degraded system performance. Correspondingly, we will use a RL framework to address such a problem without prior knowledge. It is an early attempt to investigate the relations of the MVNO and InPs and corresponding resource management problems from machine learning perspectives. In the following, the basics of RL are first presented, including the defined state, action and reward strategies. To avoid high dimensionality problem, DNN is applied and a DQN scheme is used to perform Q-learning action-value function estimation.

A. RL Framework Formulation

In the RL framework, the agent can chose actions to interact with the environment. In general, there are 3 basic elements: state, action and reward. In the presented WVN, the network controller is the agent and all the other entities can be considered as the environment. Within the action space, the network controller chooses an action in each time slot from the action space, which decides the resource allocation, and then emerges to next state. After action execution, a reward or punishment can be obtained from the environment. Such a framework aims to maximize the cumulative received rewards of the system during the interactions with the environment.

1) State: In the WVN, the network controller is able to have the necessary information of Base Stations (BSs) from the InP, e.g., battery level, maximum transmit power, available spectrum resources, and channel information. Thus, the state at each time slot should contain all these information.

2) Action: In the presented WVN, the objective is to maximize the utility of the MVNO, which may concern with the BS-user association and resource allocation factors and possible cost. Thus, action space can contain different strategies, e.g., the spectrum allocation, power allocation and/or user association. Then, the action space comprises of all the possible strategies.

3) Reward: The network controller can obtain a reward after executing action. The definition of reward is crucial as it can enforce the network controller to take proper action. As introduced, the main target of the considered scenario is to optimize the utility of all the MVNOs while satisfying each end-user's QoS. In order to link the reward to the objective function, following points are explicitly considered.

- Since the goal of RL framework is maximizing the reward, there should be a positive relations between the utility and the defined reward;
- In order to meet the the users' QoS requirements, the reward will be decreased if there is any loss of the QoS.

Therefore, we can define the immediate reward as the weighted combination of utility and end-user's requirement.

4) Q-Learning Method: At time slot l, the network controller first watches the state s_l of all the BSs, and then chooses an action a_l according to a stochastic policy π. After selecting an action, the network controller can transmit the action information to the InPs or BSs via control signalling and a reward r_l can be achieved. Then the network will take a transition to next state at time slot $l+1$.

Therefore, each pair of state-action has a Q-value Q_l for time slot l, which is the expected cumulative future discounted reward at state s_l and action a_l. The network controller computes Q_l, the value of which is stored in a Q-table for each time slot. When the optimal policy that can maximize Q_l is satisfied, the optimal function Q_l^* for action a_l is obtained and it should follow a Bellman optimality equation related to r_l and Q_{l+1}^* [11]-[13], such as

$$Q_l^* = r_l + \xi \max_{a_{l+1}} Q_{l+1}^*, \tag{1}$$

where ξ is a discount factor.

B. DRL-based Solution

In the Q-learning, there should be a Q-table which consists of all possible states as its rows and actions as its columns for each BS. This Q-table will be the reference for the network controller to select the proper action according to the Q-value. Although using such a table relieves the dependence on full network statistics information, but Q-learning still needs to confront the problem of a huge state space.

In the considered WVN, as many different entities are involved, the possibility that a very large amount of states and actions co-exist will be very high. Then, the dimension of Q-table will be very high if all the state and actions are stored. Consequently, the algorithm may not be working properly as it is difficult to get enough samples to traverse each state. Therefore, we can utilize the Q-learning with neural network (NN) to estimate instead of calculating each pair's Q-value, which leads to the concept of Deep Q-Network (DQN).

When incorporating with the NN, the performance of the Q-learning on flexibility is able to be improved. The training procedure of the DQN is the same as the one in [11], [12], which uses experience replay to reduce the correlation between training samples. When it comes to the DQN, DNN is used instead of NN in Q-network and it has been proved as a robust learning approach with better performance. Comparing with the Q-network, there are three major improvements in the DQN.

The first one is that DNN can replace the ordinary NN with a multiple layer structure. In the DNN, the multiple layers of convolution filters are used to explore the local spatial correlations. Therefore, DQN is able to extract the high-level features of input raw data. In addition, the experience replay in DQN is able to save the experience tuple into a replay memory. Then from the memory, a randomly sample batch can be used to train the DNN. In this way, DQN can learn from past experience instead of only from the current one. Moreover, a second network is adopted, by which we keep a copy of the NN and use it for the $Q(s', a')$ value in the Bellman equation. The computed target Q-values can be used to compare with the estimated Q-values to obtain the loss of each action. The idea is that using the second network's Q-values to train the main Q-network will improve the stability of the training. Using one network for the target Q-values and estimated ones can fall into feedback loops between the target and estimated values. The framework of the presented DQN can be found in Fig. 3.

C. Performance Evaluation

In this section, we evaluate the proposed DRL-based scheme with the WVN scenario consisting of 2 MVNO, 2 InPs, 4 Small cell BSs (SBSs), and 15 users. In this scenario, each InP owns physical resources, one Macro BS (MBS) and 2 SBSs which are with energy harvesting capabilities and can be powered by renewable energy supply. In this setting, the MVNOs need to rent radio resources, virtualize and then operate them to provide services to a number of users. The aim is to maximize the utility of MVNOs, which concerns both the revenue earned from the users and the cost paid to the InPs. To achieve such a goal, we consider to jointly optimize power and spectrum allocation, and user association in such a network.

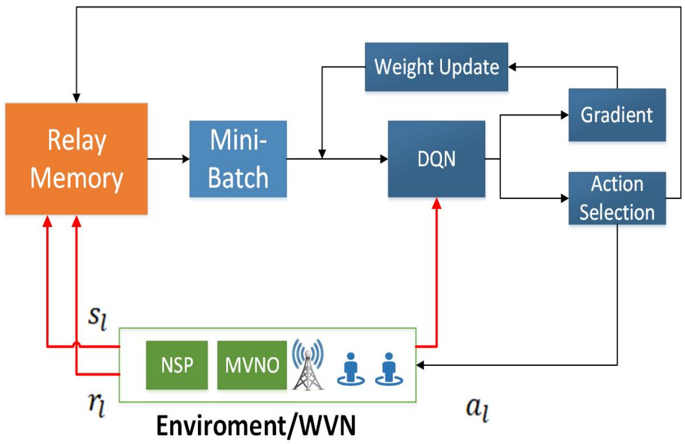

Fig. 3. Flow of the DRL-based scheme

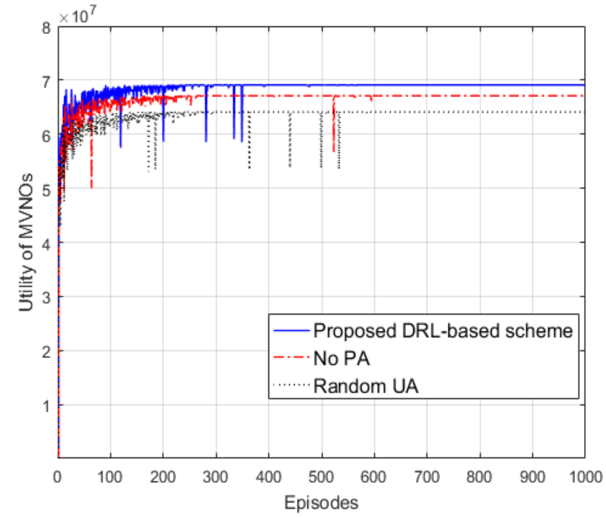

Fig. 4. Convergence performance of different DRL-based schemes.

First, we examine the impact of various factors on the system utility to see their impact. Then we evaluate the proposed schemes and show their performance. We present the convergence performance of DRL-based scheme in Fig. 4. In this figure, we also show the effectiveness of power allocation and user association, by comparing the proposed DRL-based scheme with the one without power allocation ('No PA') and the one with random user association (random UA). One can observe that the proposed DRL-based scheme has a good convergence rate and outperforms the others. From this figure, it can also be found that utility of MVNOs of all cases are low at the beginning. As the time goes on, the utility all three cases become larger until reaching a relatively stable value. In addition, the proposed DRL-based scheme outperforms the other two, which shows the necessity of investigating the power allocation and user association.

To evaluate the effects of the number of end-users, Fig. 5 varies the number of end-users and plots the utility performance. Moreover, we compare the proposed DRL-based scheme and the a distributed optimization scheme, i.e. alternating direction method of multipliers (ADMM)-based scheme. In addition, we also plot a ADMM-based scheme [5] with random user association (Random UA). We can see that the DRL-based scheme outperforms the others. This is mainly because the proposed ADMM-based scheme may fail to perform accurate resource allocation in a dynamic scenario and then results in performance loss.

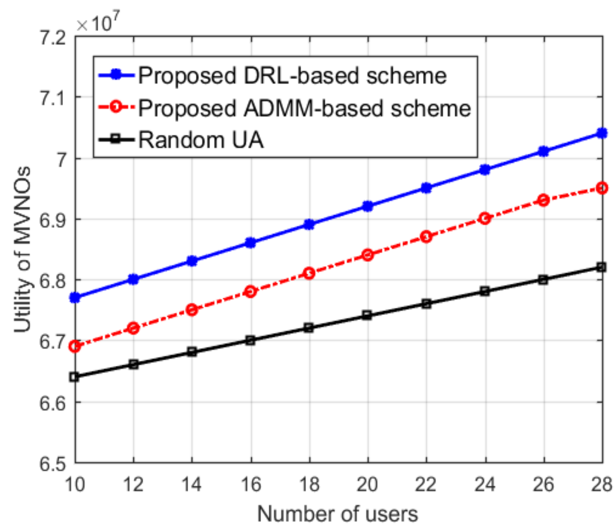

Fig. 5. Utility of MVNO vs. number of users.

V. Future Research Direction

While enabling intelligent resource management in WVN has a great potential to provide flexible services, there is still a long way to achieve it. In the following, some research directions are presented. These pointed research areas, we hope, may shed the light towards design of intelligent resource management in the WVN.

A. Integration of Network and Cloud

To realize and implement the wireless network virtualization and effective resource management schemes, the successful integration of wireless networks and cloud computing platform is vital. From the network point-of-view, the involvement of powerful computing units allows to execute more complex resource virtualization, management and optimization mechanism for operating the network in an efficient manner.

Nevertheless, most of the previous work about wireless network only concentrate on the radio resource allocation, such as frequency spectrum and power allocation. While the research of computing system focus on the computing resources and task allocation and assignment. Such limitations will motivate the resource on the joint optimization of radio and computing resources. In this context, the challenges come from how to provide an unified measurement of the radio and computing resources and develop effective optimization methods. In addition, mobility poses another challenges to the convergence and integration. Considering the ultra dense network and a fast mobility environment, resource allocation of both radio and computing resources and service provisioning are hard to address.

B. Edge Intelligence-enabled Resource Management

On one hand, the emerging multimedia-rich applications such as augmented reality (AR), video streaming, and on-line gaming, require advanced processing and storage operations of large volume data at the network edge. On the other hand, the huge potential value of the massive amounts of data collected from every aspect of transportation, Information and Communications Technology (ICT), energy, and social sectors have not been comprehensively investigated. Moreover, implementing AI at the network edge is able to make more accurate resource management decisions with reduced latency. Therefore, how to utilize the large amount of collected data and advanced AI techniques at the network edge to provide reliable and efficient resource management mechanism for wireless virtualized network call for dedicated works.

In particular, considering the large volume, great variety, high velocity, and temporal/spatial variation features of big data, it is not obvious to find effective data modeling and analysis approaches, which need sophisticated data processing schemes. The advanced machine learning schemes should be explored at the network edge to study the classifications, features, values, and potential application scenarios of big data, of which the learning outcome should further complement the development of resource allocation in WVN.

C. Software-defined Architecture for Network Integration

Meanwhile, it is also worth noticing that the WVN consists of various different kind of networks, especially considering the fact that the cells will be ultra-densely deployed. Therefore, it is envisioned that intelligent and effective architectures and approaches are urgently required to dynamically match supply and demand across time, space, and spectrum over large number of different cells in WVNs. In this context, a software-defined architecture is a promising solution for providing coordination and resource management over space-air-ground integrated cells which consist of radio access infrastructures mounted on satellites, unmanned aerial vehicles (UAVs), vehicles and regular BSs. The considered architecture should be able to provide a unified platform for centralized network management and intelligent resource allocation via the separation of control and data functions. Fast network reconfigurability, seamless interoperability, adaptive resource allocation and differentiated QoS provisioning can be provided via successful proposal.

VI. Conclusion

Wireless virtualized network (WVN) enables virtualization and sharing of physical resources, reduced expenses of network deployment and operation, easier migration to newer services, and flexible management. In this article, we highlighted the concepts and importance of efficient virtual resource management in the WVNs and discussed the related challenges and possible solutions. As one example, A deep reinforcement learning framework is proposed along with proper illustration as an example to study how to utilize resources and provide an efficient management scheme in WVNs. Future directions were presented to highlight the research potentials.

References

[1] P. Rost, I. Berberana, A. Dekorsy, G. Fettweis, A. Maeder, H.Paul, V. Suryaprakash, M. Valenti, and D. Wübben, "Benefits and challenges of virtualization in 5G radio access networks," *IEEE Communications Magazine*, vol. 53, no. 12, pp. 75-82, Dec. 2015.

[2] C. Liang and F. R. Yu, "Wireless virtualization for next generation mobile cellular networks," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 61-69, Feb. 2015.

[3] Z. Chang, Z. Zhou, S. Zhou, T. Ristaniemi and T. Chen, "Towards service-oriented 5G: Virtualizing the networks for everything-as-a-service," *IEEE Access*, vol. 6, pp. 1480-1489, 2018.

[4] Z. Feng, C. Qiu, Z. Feng, Z. Wei, W. Li and P. Zhang, "An effective approach to 5G: Wireless network virtualization," *IEEE Communications Magazine*, vol. 53, no. 12, pp. 53-59, Dec. 2015.

[5] H. K. Nguyen, Y. Zhang, Z. Chang, and Z. Han, "Parallel and distributed resource allocation with minimum traffic disruption for network virtualization," *IEEE Trans. on Commun.*, vol. 65, no. 3, pp. 1162-1175, Mar. 2017.

[6] A. Moubayed, A. Shami, H. Lutfiyya, "Wireless resource virtualization with device-to-device communication underlaying LTE network," *IEEE Transactions on Broadcasting*, vol.61, no.4, pp.734-740, Dec. 2015.

[7] Z. Chang, Z. Han and T. Ristaniemi, "Energy efficient optimization for wireless virtualized small cell networks with large scale multiple antenna,"*IEEE Trans. on Commun.*, vol. 65, no. 4, pp. 1696-1707, Apr. 2017.

[8] S. Parsaeefard, R. Dawadi, M. Derakhshani and T. Le-Ngoc, "Joint user-association and resource-allocation in virtualized wireless networks," *IEEE Access*, vol. 4, pp. 2738-2750, 2016.

[9] C. Liang, F. R. Yu, H. Yao and Z. Han, "Virtual resource allocation in information-centric wireless networks with virtualization," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9902-9914, Dec. 2016.

[10] Y. Sun, M. Peng and S. Mao, "Deep reinforcement learning-based mode selection and resource management for green fog radio access networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1960-1971, April 2019.

[11] N. Zhao, Y. Liang, D. Niyato, Y. Pei, M. Wu and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5141-5152, Nov. 2019.

[12] Y. He, N. Zhao and H. Yin, "Integrated networking, caching, and computing for connected vehicles: A deep reinforcement learning approach," *IEEE Trans. Veh. Tech.*, vol. 67, no. 1, pp. 44-55, Jan. 2018.

[13] X. Zhang, M. Peng, S. Yan and Y. Sun, "Deep-Reinforcement-Learning-Based Mode Selection and Resource Allocation for Cellular V2X Communications," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6380-6391, July 2020.

[14] H. A. Shah and L. Zhao, "Multiagent Deep-Reinforcement-Learning-Based Virtual Resource Allocation Through Network Function Virtualization in Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3410-3421, March, 2021.

[15] Y. Xu et al., "Constrained Reinforcement Learning for Resource Allocation in Network Slicing," *IEEE Communications Letters*, vol. 25, no. 5, pp. 1554-1558, May 2021.