This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2017.2664808, IEEE Communications Letters

A Radio Resource Virtualization based RAT Selection Scheme in Heterogeneous Networks

Shaoshuai Fan, Hui Tian, and Weidong Wang

***Abstract*—In the future heterogeneous wireless networks, heterogeneity of radio resources from different Radio Access Technologies (RATs) still exists. The heterogeneity, especially for networks with the coexistence of non-orthogonal and orthogonal resources, makes the radio resources difficult to be uniformly measured, and thus hinders the efficient utilization of radio resources. To overcome this limitation, this paper firstly proposes a radio resource virtualization approach in heterogeneous networks. Based on the accumulated historical data of resource utilization information, heterogeneous radio resources are virtualized into normalized resources using deep learning method. Secondly, the consumption difference of virtualized resources under different situations of network load and user demand is modeled. Moreover, aiming at efficiently utilizing radio resources and reducing access blocking rate, a RAT selection scheme based on the radio resource virtualization is proposed. Through simulation, the validity of the proposed scheme is evaluated.**

***Index Terms*—Heterogeneous wireless networks, radio resource virtualization, RAT selection, deep learning.**

I. Introduction

FOR coping with the increasing demands for capacity and various QoS support, a cost-effective solution is to use existing radio access technologies (RATs) [1]. Therefore, heterogeneity of wireless network architectures (e.g., the co-existence of 2.5G, 3G, 4G, Wi-Fi, etc) is an important feature of current and future wireless networks. Even the upcoming 5G networks are devised with the vision of heterogeneity [2].

From the perspective of efficient utilization of radio resources, heterogeneous RATs should be integrated and jointly managed. However, the heterogeneity arise challenges for network design and operation. Due to the nature composition difference of radio resources, the radio resources from heterogeneous RATs are difficult to be uniformly measured, and thus hinders the efficient utilization of overall radio resources [3]. Take RAT selection problem for example, mobile devices are increasingly equipped with multiple RATs that can choose to access any network. Without the knowledge of heterogeneous radio resource properties, the RAT selection decisions may cause loss of utilization efficiency of radio resources in heterogeneous networks.

For jointly managing the radio resources of heterogeneous networks, massive literatures have been proposed. In [4], the heterogeneous radio resources are formulated as bandwidth, and radio resource utilization was optimized through the bandwidth allocation where the heterogeneity of radio resource units was not reflected. In [1], [5], network-assisted approaches for RAT selection were proposed, and heterogeneous networks composed of OFDM-based radio access technologies were considered where the radio resources are orthogonal. However, the heterogeneous radio resources can be divided into two categories: orthogonal resource (e.g. resource block in OFDMA system) and non-orthogonal resource (e.g. code resource in CDMA system) [6]. Most existing literatures focus on the network with orthogonal resource. The non-orthogonal radio resource, of which the utilization efficiency is nonlinear due to non-orthogonal interference between radio resources, is difficult to be normalized and corresponding resource management strategy has not been well studied in heterogeneous networks.

Additionally, RAT selection problem [7], [8] has been well studied under certain assumptions of throughput models where the throughput of users depend on the number of connected users. However, QoS mechanism exists in most networks for realizing the guaranteed/constant bit rate of users through resource allocation, and therefore a model of evaluating the consumption of radio resource is needed for better radio resource management in heterogeneous networks.

Different from existing works, this paper focuses on radio resource virtualization and efficient radio resource utilization in heterogeneous networks, more specifically, normalizing the non-orthogonal resource, evaluating the consumption of virtualized radio resource and proposing a RAT selection scheme to reduce access blocking rate.

II. A Radio Resources Virtualization based RAT Selection Scheme

A. Virtualization of radio resources

In heterogeneous wireless networks, radio resources from different RATs have their specific characteristics and orthogonality. It will bring benefits for improving utilization efficiency of radio resources through jointly and uniformly management if we can virtualize the non-orthogonal resource into orthogonal resource [3], e.g. resource block.

To abstract heterogeneous radio resources into normalized orthogonal resources, the huge amount of historical data of resource utilization information accumulated in the network operation process can be used. Considering the fact that utilization efficiency of radio resource are different for users locating at different places or under different network load situations of serving cell and neighboring cells, we recompose the information of resource utilization situation as $\mathbf{S} = \{\mathbf{L}, \mathbf{R}\}$, where $\mathbf{L} = \{x, y\}$ is the location of user, $\mathbf{R} = \{R_s, R_1, ..., R_{N_c}\}$ are the ratios of used resource of serving cell and neighboring cells and N_c is the number of neighboring cells.

Before mapping the non-orthogonal resource into orthogonal resource, we first weight the capacity of the heterogeneous resources. Without loss of generality, we take two-RAT

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2017.2664808, IEEE Communications Letters

heterogeneous networks consisting of a network with non-orthogonal resource and a network with orthogonal resource for example, the historical data of resource utilization information $\mathbf{H}^{no} = \{\mathbf{H}_1^{no}, \mathbf{H}_2^{no}, ...\}$ in networks with non-orthogonal resource and $\mathbf{H}^{o} = \{\mathbf{H}_1^{o}, \mathbf{H}_2^{o}, ...\}$ in networks with orthogonal resource are utilized, where $\mathbf{H}_i^{\bullet} = \{\mathbf{S}_i^{\bullet}, r_i^{\bullet}\} (\bullet \in \{no, o\})$ is the ith record of resource utilization information of a certain network and $r_i^{\bullet}$ is the achieved rate of a user occupying all the residual resource under the situation $\mathbf{S}_i^{\bullet}$. Considering that the accumulated historical information may be sparse in terms of location and network load information, and the impacts of various network and user situations are different. To fully capture the characteristics of resource capacity and describe the capacity under all the continuous situations, deep neural network (DNN) [9] can be adopted for generalizing the data sets and forming the mapping function from resource utilization situation to the capacity of residual resource.

The DNN we adopted consists of one input layer, K hidden layers and one output layer as shown in Fig. 1. The input layer consists of a set of nodes representing the resource utilization situation $\mathbf{S}$ denoted as $\mathbf{O}^{(0)}$. All nodes in the previous layer are connected with all nodes in the next layer with weights $\mathbf{w}$. To better sense the input information, the sigmoid function $sig(x) = 1/(1 + e^{-x})$ is used as activation function of hidden layers. Thus, the output value of the nth node in the kth $(k = 1, 2, ..., K)$ hidden layer is $O_n^{(k)} = sig(\sum_{m=1}^{N_{k-1}} w_{mn}^{(k-1)} O_m^{(k-1)} + w_{0n}^{(k-1)})$, where N_k is the number of nodes in the kth layer. In the output layer, there exists one node and its output value $O_1^{(K+1)} = \sum_{m=1}^{N_K} w_{m1}^{(K)} O_m^{(k)} + w_{01}^{(K)}$ represents the capacity of residual resource r for $\mathbf{S}$.

To make DNN be able to approximate the objective function with arbitrary precision, a learning process is needed to determine the proper weight values using the training data consisting of the inputs $\mathbf{S}$ and outputs r. In our case, the weights are pretrained using the discriminative pretraining approach [10] in an offline manner as follows.

- Firstly, establish an one-hidden-layer neural network and train the weights of this network using the well-known back-propagation (BP) method represented as $w_{mn}^{(k)} = w_{mn}^{(k)} - \alpha \frac{\partial E}{\partial w_{mn}^{(k)}}$, where $E = (O^* - O)^2/2$ is the squared error signal between the output O of output layer and its target value O^*, and α is the learning rate.
- Secondly, insert a new hidden layer between the last hidden layer and output layer, and train the weights connecting to the newly added layer using BP method and the same training data. Repeatedly inserting new hidden layers and training the weights till the predefined number K of hidden layers is reached.
- Finally, jointly fine-tune the weights of all layers using BP method and then mark this DNN trained.

The trained DNN plays the role of mapping function from resource utilization situation to the capacity of residual resource. Note that, besides this offline pretraining procedure, the DNN will be continually trained in a online manner using BP method and newly accumulated data.

Based on the resource utilization information $\mathbf{H}^{no}$ and $\mathbf{H}^{o}$ of heterogeneous networks, the mapping functions from resource utilization situation to the capacity of residual resource can be separately learned as functions f_{no} and f_o by two DNNs, where $r^{no} = f_{no}(\mathbf{S}^{no})$ and $r^{o} = f_o(\mathbf{S}^{o})$. Based on the capacity of residual resource, residual non-orthogonal resource can be virtualized into $V_{no}(\mathbf{S}_{no})$ orthogonal resource blocks, where $V_{no}(\mathbf{S}_{no}) = (1 - R_s^{no}) R_{max}^{o} f_{no}(\mathbf{S}_{no}) / f_o(\mathbf{S}_{no})$ and R_{max}^{o} is the total number of resource blocks in the orthogonal network.

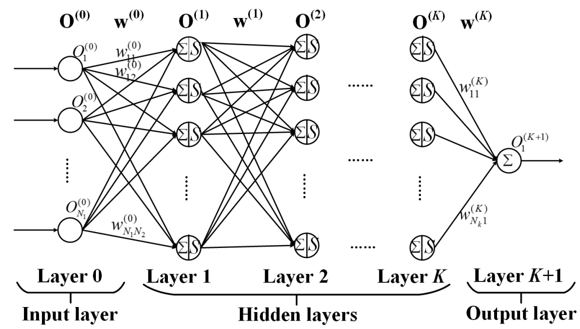

Fig. 1. diagram of a deep neural network

B. Consumption model of virtualized resource

After the virtualization of radio resources, the total capacity of residual heterogeneous resources can be measured in the same orthogonal resource units. However, the consumption of virtualized resource units for meeting the same user demand in different RAT is still different even under the same network situation. The main reason of this difference is that resources of heterogeneous networks are still practically consumed in their own physical resource units which are mapped to different numbers of virtualized resource units. Therefore, to effectively utilize the virtualized resource, a model for measuring the consumption of virtualized resources should be established.

Before a user accesses the target network, only cell reference signal can be measured and specific information of traffic channel gain can not be obtained. As a result, only the expected consumption of resource can be obtained and provide general guideline for network management. Therefore, we establish the expected consumption model based on the accumulated historical data of resource consumption information described as $\hat{\mathbf{H}}^{no} = \{\hat{\mathbf{H}}_1^{no}, \hat{\mathbf{H}}_2^{no}, ...\}$ in the non-orthogonal networks and $\hat{\mathbf{H}}^{o} = \{\hat{\mathbf{H}}_1^{o}, \hat{\mathbf{H}}_2^{o}, ...\}$ in the orthogonal networks, where $\hat{\mathbf{H}}_i^{\bullet} = \{\mathbf{S}_i^{\bullet}, d_i, p_i^{\bullet}\}$ is the ith record of resource consumption of a certain network, d_i is the achieved rate of a user and $p_i^{\bullet}$ is the corresponding consumption ratio of physical resource including the rate compensation for other affected users if the radio resource is non-orthogonal.

Still using the deep neural network mentioned above, we can firstly establish the mapping functions from rate demand d to consumption ratio p of physical resource. The inputs of the DNN are $\mathbf{S}, d$ and the output is p. Based on the accumulated historical data of resource consumption information, the mapping functions of heterogeneous networks can be learned as functions g_U and g_L by another two DNNs, where $p^{no} = g_{no}(\mathbf{S}^{no}, d)$ and $p^{o} = g_o(\mathbf{S}^{o}, d)$.

Afterwards, the expected consumption of normalized virtualized resources for satisfying the user demand can be measured as $C_{no} = V_{no}(\mathbf{S}_{no}) - V_{no}(\{\mathbf{L}, \{R_s^{no} + p^{no}, R_1^{no}, ..., R_{N_c}^{no}\}\})$ for network with non-orthogonal resource and $C_o = p^o R_{max}^o$ for network with orthogonal resource. However, in the special case that the corresponding network can not satisfy the user demand due to the lack of available resource, i.e. $R_s^\bullet + p^\bullet > 1$, only the residual resource will be consumed, and the expected consumption can be measured as $C_{no} = V_{no}(\mathbf{S}_{no})$ or $C_o = (1 - R_s^o)R_{max}^o$.

C. Resource virtualization based RAT selection

Using the above-mentioned virtualization model and consumption model, heterogeneous resources have been virtualized into normalized orthogonal resources, and the resource consumption of traffic demand can be uniformly measured. These results can provide general guideline for network management, especially for the selection of accessing network.

In this paper, considering that accessing different network will cause different consumption of available resource due to the resource heterogeneity and non-linear consumption of non-orthogonal resource, we further propose a resource virtualization based RAT selection scheme (RVRS) as follows.

- Firstly, based on the accumulated historical data, virtualization model f and consumption model g are learned in an offline manner. Then, before a user accesses a network, the network should obtain the location $\mathbf{L}$ of the target user through existing positioning technology from network side or location information (e.g. GPS information) report from user side. Moreover, the network information $\mathbf{R}_{no}$ and $\mathbf{R}_o$ of nearby cells should be exchanged locally through intra-RAT and inter-RAT interfaces.
- Secondly, based on the resource utilization situation $\mathbf{S}_{no} = \{\mathbf{L}, \mathbf{R}_{no}\}$ and $\mathbf{S}_o = \{\mathbf{L}, \mathbf{R}_o\}$, the residual resources of heterogeneous networks are virtualized using virtualization model as $V_{no}(\mathbf{S}_{no})$ and $V_o(\mathbf{S}_o)$. Moreover, based on the user demand d, the consumption of the virtualized resources of candidate heterogeneous networks are measured as C_{no} and C_o according to the situation of residual resource.
- Finally, aiming at saving the overall network resource, if any network has enough resource for meeting the user demand, i.e. $R_s^{no} + p^{no} \leq 1$ or $R_s^o + p^o \leq 1$, the user is guided to access network n_s with minimum consumption of normalized resource, i.e. $n_s = \arg\min_{\bullet \in \{no,o\}}(C_\bullet | R_s^\bullet + p^\bullet \leq 1)$. Otherwise, the user is guided to access the network with maximum residual virtualized resource for better satisfaction, i.e. $n_s = \arg\max_{\bullet \in \{no,o\}}(C_\bullet | R_s^\bullet + p^\bullet > 1)$. And obviously, if all networks are full-loaded, i.e. $R_s^\bullet = 1$, the user will be blocked.

Note that, the computation complexity of the proposed scheme is quite low because of the offline manner of learning. The proposed scheme is applicable for practical application and can be adapted to scenarios where orthogonal and non-orthogonal resources coexist and joint radio resource management is needed in heterogeneous networks.

III. Simulation Results and Analysis

For the simulation, we consider a scenario of heterogeneous networks consisting of 19 UMTS cells and overlapping 19 LTE cells. The inter-site distance of UMTS networks and LTE networks are both 500m. The bandwidth of both networks are 5MHz. The maximum transmission power of each cell is 43dBm. The location of users follows uniform distribution. The path losses are calculated as $L = 37.6log_{10}d + 128.1$ for UMTS users and $L = 37.6log_{10}d + 127.7$ for LTE users, where d measured in kilometers is the distance between the cell and user. Rayleigh fading is considered for small-scale fading. The rate demand of users are randomly initialized between 0.5Mbps to 2Mbps. The last time of each user traffic is set to 30s. Each of the four adopted DNNs for establishing virtualization models and consumption models has five hidden layers, and each layer has twenty hidden nodes. The learning rate of back-propagation method is 0.1.

After training the four DNNs, the coefficients of determination, which reflect the learning accuracy, of the mapping functions (f_{no}, f_o, g_{no} and g_{no}) are respectively 0.957, 0.999, 0.934 and 0.969. The results of the coefficients of determination imply that the accuracies of trained DNNs are at high levels. The higher accuracy means the mapping function can better describe the mapping relations reflected by the historical data. Obviously, the more accuracy the mapping function has, the better performance gain may be achieved because the efficiency of resource utilization will be more accurately described and this may lead to better RAT selection decisions.

Then, we compared the following RAT selection schemes:

- RVRS: Our proposed scheme.
- Random: Randomly select a accessible network to access.
- Minimum: Access a network with the minimum ratio of already used resource units to overall resource units.

As the main metric of network access selection, user blocking rate is evaluated as illustrated in Fig. 2. As the user arrival rate increase, more users compete for radio resource which increases the probability of being full-load for networks, and correspondingly the blocking rate will increase. Among these schemes, Random performs the worst because the heterogeneity of networks is totally neglected and causes improper utilization of radio resources. Minimum performs better than Random. However, as the user arrival rate increase, the user blocking rates under Minimum is rapidly increased to the similar level as Random. This implies that heterogeneous resources can also not be fully utilized when only the occupation ratio of resources is considered, and Minimum will also block a great number of users when the arrival rate is high. Based on the consideration of resource utilization efficiency, RVRS can guide users to access the more resource-efficient networks under specific situations of network load and user demand, thus the user blocking rate is kept to a quite low level. Compared with Random and Minimum schemes, the absolute gains in user blocking rate are 6.06% and 5.05% respectively on average, and both 7.40% at most when the arrival rate is 2.5 users/s/km^2.

To evaluate the rate performance of users under different RAT selection schemes, we define a rate satisfaction ratio

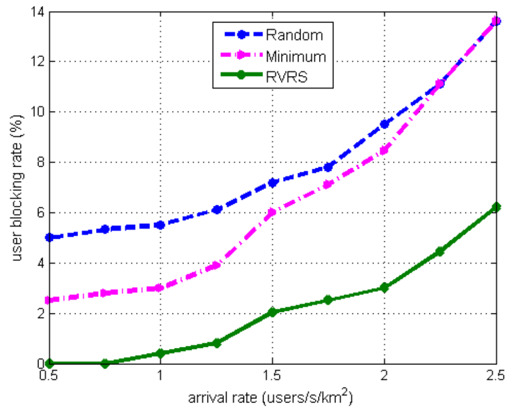

Fig. 2. User blocking rate

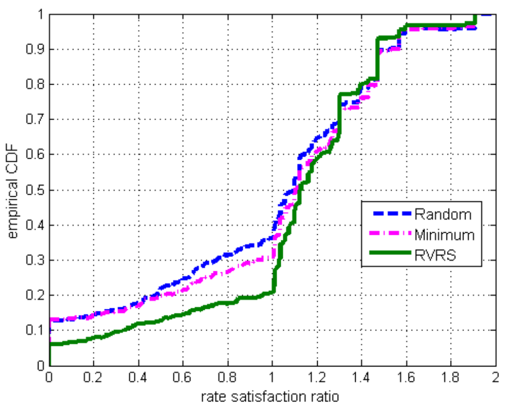

Fig. 3. CDF of rate satisfaction ratio

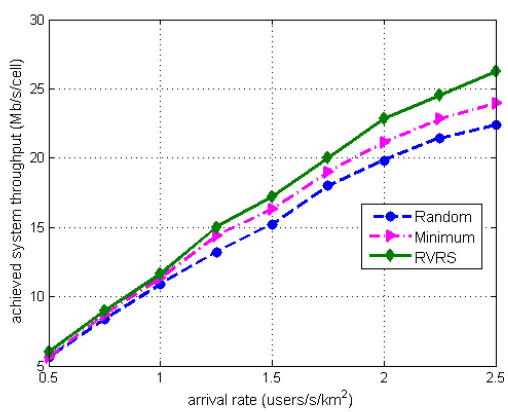

Fig. 4. Achieved system throughput

metric of a user as the ratio of the achieved rate to the rate demand, and the cumulative distribution function of rate satisfaction ratio is shown in Fig. 3, where the arrival rate is 2.5 users/s/km^2. Compared with Random and Minimum, more users achieve their demand under RVRS. The ratio of satisfied users is improved by 15.38% and 10.04%, respectively. These results verified that, on one hand, RVRS guides users to access the more resource-efficient networks and reduces the user blocking rate which makes more users get the service, and on the other hand, users are guided to access a network which has more normalized resource and is more capable of satisfying users' QoS demands when network load is high.

Fig. 4 shows the evaluation of achieved system throughput. From the results, we can observe that RVRS improves the system throughput by 11.60% and 5.71% on average compared with Random and Minimum schemes. When the arrival rate is 2.5 users/s/km^2, RVRS improves system throughput by 17.20% and 9.47%, respectively. This improvement implies that RVRS greatly improves the resource utilization efficiency of heterogeneous RATs by normalizing the heterogeneous radio resources and measuring the different resource consumptions for satisfying users' demands. As a result, users will be guided to access the more efficient network and cause less consumption of total network radio resources. Therefore, heterogeneous radio resources can be efficiently used and the limited radio resources can meet more users' demands.

IV. Conclusions

In this paper, we focused on radio resource virtualization and efficient radio resource utilization in heterogeneous networks. A deep learning based resource virtualization approach is proposed for normalizing heterogeneous radio resources, and then the consumption model is established for measuring the different consumptions of heterogeneous resources for satisfying users' demands. Moreover, a RAT selection scheme based on the radio resource virtualization is proposed. Simulation results show that, it is of benefit for heterogeneous networks to virtualize the heterogeneous radio resources and guide users to access the more resource-efficient networks under specific situations.

References

[1] M. E. Helou, M. Ibrahim, S. Lahoud, K. Khawam, D. Mezher, and B. Cousin, "A network-assisted approach for RAT selection in heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1055–1067, 2015.

[2] White paper: 5G radio network architecture. [Online]. Available: http://www.ict-ras.eu/

[3] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 358–380, 2015.

[4] J. Miao, Z. Hu, K. Yang, C. Wang, and H. Tian, "Joint power and bandwidth allocation algorithm with QoS support in heterogeneous wireless networks," *IEEE Communications Letters*, vol. 16, no. 4, pp. 479–481, 2012.

[5] M. Ibrahim, K. Khawam, and S. Tohme, "Congestion games for distributed radio access selection in broadband networks," in *GLOBECOM, 2010 IEEE*, Miami, FL, United states, Dec. 6–10, 2010, pp. 1–5.

[6] P. Wang, J. Xiao, and L. Ping, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Vehicular Technology Magazine*, vol. 1, no. 3, pp. 4–11, 2006.

[7] A. Keshavarz-Haddad, E. Aryafar, M. Wang, and M. Chiang, "HetNets selection by clients: convergence, efficiency, and practicality," *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–14, 2016.

[8] P. Naghavi, S. H. Rastegar, V. Shah-Mansouri, and H. Kebriaei, "Learning RAT selection game in 5G heterogeneous networks," *IEEE Communications Letters*, vol. 5, no. 1, pp. 52–55, 2016.

[9] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.

[10] D. Yu, L. Deng, F. Torsten, B. Seide, and G. Li, "Discriminative pretraining of deep neural networks," U.S. Patent 0 026 914, Jan. 28, 2016.