BALANCING BACKHAUL LOAD IN HETEROGENEOUS CLOUD RADIO ACCESS NETWORKS

CHEN RAN, SHAOWEI WANG, AND CHONGGANG WANG

Chen Ran and Shaowei Wang are with Nanjing University.

Chonggang Wang is with InterDigital Communications.

ABSTRACT

Inspired by the explosive growth of mobile data traffic, the severe inter-tier interference, and the fierce competition between the total cost of ownership and revenues for mobile operators, the heterogeneous cloud radio access network (H-CRAN) has been proposed as one of the most prominent ways to handle these challenges. The key idea of the H-CRAN is incorporating cloud computing into a heterogeneous network (HetNet) to enhance coordinated multipoint transmission and reception, cooperative radio resource management, and self-organizing networks, which improve both spectral and energy efficiencies of cellular systems. One of the most critical challenges that hinder the implementation of the H-CRAN is the high transmission demand on the backhauls between the baseband unit (BBU) pool and remote radio heads (RRHs). In this article we suggest that we can balance the workload of different RRHs to alleviate the pressure on the transmission links. Our proposed method is different from but compatible with existing compression techniques that have been widely investigated in the literature to lighten the transmission burden of the backhauls. We also describe the technical challenges in existence during the implementation of our proposal and give preliminary ideas of how to address them.

INTRODUCTION

In recent years mobile data traffic has experienced record growth among the world's operators as subscribers use more smart phones and mobile devices, such as tablets. Diverse data applications such as high-definition wireless video streaming, machine-to-machine communication, social networking, and Internet of Things (IoTs), have been proliferating over the past 20 years. Investigations have discovered that Internet traffic is expected to increase more than 1000 times by the year 2020. In order to satisfy growing user demand, mobile network operators must increase network capacity steadily. As the spectral efficiency for the Long Term Evolution (LTE) standard is approaching the Shannon limit, further improvements in system spectral efficiency can be achieved only by increasing the density of access nodes. In a relatively sparse deployment of macro base stations (BSs), adding more BSs does not severely increase intercell interference, and solid cell splitting gains are easy to achieve. However, as the density of BSs is already quite large today, cell splitting gains are greatly reduced due to the intercell interference that is severe under such circumstances. In addition, costs concerning site acquisition in a capacity-limited dense urban area can become extremely expensive. Furthermore, the current cellular network architecture is 40 years old and was primitively designed for coverage and mobility considerations, instead of achieving high energy efficiency (EE) and spectral efficiency (SE) performance. To deal with such challenging goals, revolutionary approaches concerning novel wireless network architectures supported by advanced signal processing and networking technologies are urgently needed [1].

Challenges associated with the dense deployment of traditional macro BSs can be partly avoided by the utilization of low power nodes (LPNs), which are normally classified into pico BSs, femto BSs, and relays. A network that consists of macro cells and LPNs, where some of them are configured with restricted access and some may be lacking in wired backhaul, is referred to as a heterogeneous network (HetNet) [2]. LPNs are usually deployed in the coverage of the macro BSs to offload the traffic of these macro BSs and support users with high data rate requirements. As a result, HetNets can serve customers in traffic hot zones by deploying dense LPNs and providing ubiquitous coverage to all user equipment through the powerful macro BSs. However, they still face the problem of severe interference when the distribution of LPNs is too dense, which is expected in 5G. Therefore, it is of great significance to suppress interference through advanced signal processing techniques to fully exploit the potential of HetNets. Coordinated multiple point (CoMP) transmission and reception is increasingly seen as a promising solution to improve the performance of HetNets. Unfortunately, it also has disadvantages in practical cellular networks as its performance gain relies heavily on the backhaul constraints and even degrades as the density of LPNs increases

1536-1284/15/$25.00 © 2015 IEEE

[1]. Moreover, such a proposal also increases the complexity of the cellular network and always incurs high capital expenditure (CAPEX) and operating expenditure (OPEX) [3, 4].

In order to enhance SE and EE performance and decrease energy consumption in HetNets with densely deployed LPNs, a novel architecture for improving both SE and EE through restraining inter-tier interference and improving cooperative processing capabilities is urgently needed. Cloud computing technology has emerged as a prominent scheme for providing high gigabit data rates at the cost of lower energy across software defined wireless communication networks. As a consequence, heterogeneous cloud radio access networks (H-CRANs) are proposed as a cost-effective alternative to suppress inter-tier interference and enhance cooperative processing gains in HetNets accompanied by cloud computing technology. The motivation behind H-CRANs is to improve the capabilities of macro BSs with massive multiple-input multiple-output (MIMO) technology and simplify LPNs by building connections between LPNs and a signal processing cloud (BBU pool in the H-CRANs) through high-speed optical fibers. To be more specific, the baseband data processing and radio resource control in traditional LPNs are transferred to the BBU pool to take advantage of cloud computing capabilities [5].

H-CRANs have been researched in the literature in the areas of resource assignment and interference coordination. These works mainly focus on mitigating inter-tier interference among remote radio heads (RRHs) and macro BSs, and allocating resources in an energy-efficient manner. However, the critical problem concerning backhauls that hinders the implementation of H-CRANs has not been fully investigated. Previous solutions concerning backhauls mostly concentrate on various compression techniques and the single fiber bi-direction technique. These are effective in decreasing data transmission in backhaul for the given amount of data. However, is it possible to decrease the amount of data to be transferred at the source for the RRHs with a heavy load, so that all backhauls only carry a moderate load? In this article we will show that the answer to this question is "Yes" and propose a novel scheme for alleviating the high pressure on backhauls in H-CRANs during centralization based on workload balancing. In the following sections the system architecture of H-CRANs is briefly presented, and its bottleneck, high demand on the data transmission through backhauls, is discussed, as well as current proposals concerning backhauls of H-CRANs. Then a novel method for alleviating data transmission pressure on the backhauls is proposed, where the key idea is based on workload balancing. The challenging issues related to the proposal and possible research directions are discussed.

H-CRANS ARCHITECTURE AND CRITICAL CHALLENGES

As mentioned earlier, even though HetNets are good at providing seamless coverage and high capacity, their commercial development is blocked by the huge amount of signaling in backhauls caused by the intracell and intercell CoMPs and the low EE performance with hyper-dense LPNs. On the other hand, cloud radio access networks (C-RANs) have been recognized as a promising architecture for providing high transmission data rates with excellent EE performance, and most importantly, decreasing the CAPEX and the OPEX of cellular networks. The C-RAN architecture mainly consists of three parts: first, the distributed radio units that can be referred to as RRHs with antennas located at the remote site; second, the baseband unit (BBU) pool which is composed of high-performance programmable processors and real-time virtualization and management technologies; and third, the high bandwidth low-latency transport network that links the RRHs and the BBU pool. Most of the processing procedures are aggregated into the centralized virtual BBU pool, like the synch, radio resource control, transport-media access control, and the baseband processing, while the RRH is only for radio functionalities, such as digital processing, frequency filtering, and power amplification [3, 6–8]. However, backward compatibility with existing cellular systems, together with the support for seamless coverage, make macro cells still critical for a practical cellular network, since RRHs are mainly designed for high capacity in special zones without much consideration in coverage. With the help of macro cells, multiple radio networks can work in harmony with each other, and system control signals throughout the network can be easily delivered. Because of these advantages, H-CRANs are proposed to incorporate macro cells into C-RANs, and benefit from the strengths of both HetNets and C-RANs, where cloud computing technologies are fully exploited to cope with the challenges in HetNets [1].

H-CRANs have many similarities with C-RANs [6], as shown in Fig. 1, with a large number of RRHs of low power consumption cooperating with each other and a centralized BBU pool facilitated to achieve high performance gains. As mentioned earlier and specifically marked in Fig. 1, only simple symbol processing and the front radio frequency (RF) are conducted in the RRHs, while other baseband processing and upper layer procedures are conducted in the BBU pool.

On the other hand, H-CRANs differ from C-RANs with macro cells. Macro cells play a significant role in H-CRANs as they connect with the centralized BBU pool to alleviate cross-tier interference between RRHs and macro cells with centralized cooperative processing techniques based on cloud computing. Moreover, with the help of macro cells, the backhaul requirements are alleviated as the control signals and data are decoupled in H-CRANs. To be more specific, all control signals are delivered through macro cells to user equipment (UE), which weakens the capacity and delay constraints in backhauls between RRHs and the BBU pool. Furthermore, as has been discussed in [9, 10], services with a small amount of data, moving at high speed or requiring a low data transmission rate, can be supported efficiently by macro cells, while applications that require a high data transmission rate or have low mobility are allocated

With the help of macro cells, multiple radio networks can work in harmony with each other, and system control signals throughout the network can be easily delivered. Because of these advantages, H-CRANs are proposed to incorporate macro cells into C-RANs, and benefit from the strengths of both HetNets and C-RANs.

The most critical challenge that hinders the implementation of H-CRANs is the high demand on the data transmission through optical or wireless links. Insufficient backhaul capacity prevents RRHs from making full use of available radio resources, which becomes even more worrysome when various cooperative techniques are applied.

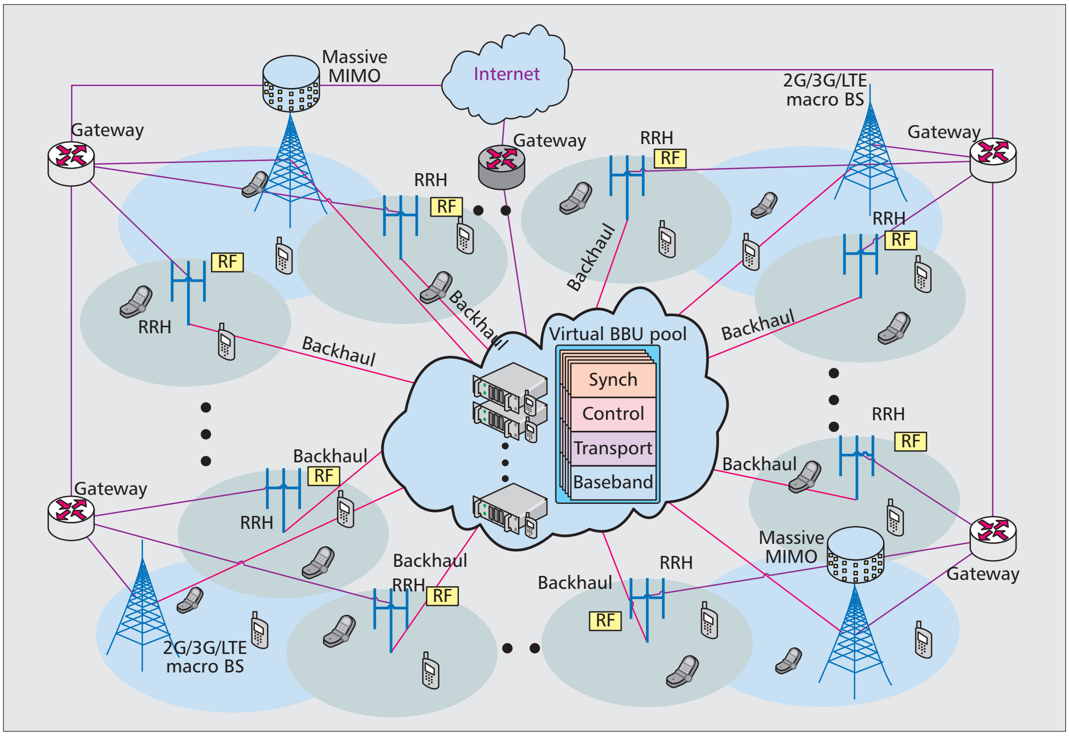

Figure 1. Heterogeneous C-RAN.

to RRHs. To enhance coverage and capacity, MIMO serves as an effective technology to be applied to macro BSs. With the help of centralized processing in the BBU pool, cooperative processing techniques inherited from virtual MIMO can achieve multiplexing gains. As with C-RANs, inter-RRH interference can be restrained by the cooperative processing techniques in the BBU pool. Moreover, the interference among macro cells and RRHs can be mitigated through cooperative resource management via an interface between macro cells and the BBU pool [1].

Despite the potential advantages of H-CRANs, they still require long-term development before they van be widely deployed. The most critical challenge that hinders the implementation of H-CRANs is the high demand on the data transmission through optical or wireless links. Insufficient backhaul capacity prevents the RRHs from making full use of available radio resources, which becomes even more of a problem when various cooperative techniques are applied.

Centralization is the first key step required to implement all the other features of H-CRANs, which aggregates various BBUs into one central office with shared resources and facilities. The critical challenge for centralization is the need for a large number of fiber resources when using a dark fiber solution, which means a direct fiber connection. For example, in a TD-LTE system with 20MHz bandwidth where the RRHs are equipped with eight antennas, the Common Public Radio Interface (CPRI) data rate between one BBU and one RRH for one TD-LTE carrier transmission reaches up to 9.8Gbps. Considering both uplink and downlink, we need four fiber connections of 6 Gbps optical modules to realize possible transmission. As one site usually consists of three sectors that support at least one carrier separately, as high as 12 fiber connections are required to enable valid communication, which concerns most operators due to limited backhaul resources [4]. Furthermore, with the explosive growth of data demand, ultra dense RRHs will be required in the coming 5G. If no remedial measures are taken, the ultra high demand caused by hyper-dense RRHs on backhauls between RRHs and the BBU pool will be intolerable in the future.

RETHINKING CURRENT PROPOSALS FOR BACKHAULS OF H-CRANS

Various data compression techniques have been investigated to cope with the BBU-RRH signal transmission bandwidth problems [11].

In time domain, reducing the signal sampling rate is a solution of low complexity and has minimal impact on protocols while conducting efficient compression of up to 66 percent with some performance degradation. Non-linear quantization serves as an alternative. More quantization levels are specified in a magnitude where more values are likely to be presented, which improves quantization signal to noise radio (QSNR). In addition, mature algorithms such as μ-Law and A-Law are available to specify the step size. Moreover, it can achieve compression efficiency of up to 53 percent. However, it also brings additional Open BBU RRH interface (OBRI) complexity. Interface-in-phase and quadrature (IQ) data compression is another well-developed compression technique. It is based on reducing the signals' dynamic range by normalizing their power to the average power reference. Even though high compression efficiency can be achieved, it has high complexity, real-time and compression distortion issues, and there is no mature algorithm available [3, 6].

An example in the frequency domain scheme is to perform sub-carrier compression. It does reduce Ir interface load by 40 percent to 58 percent by implementing FFT/IFFT. However, due to drawbacks such as high system complexity, extra processing ability on optical chips and thermal design, high device cost and difficulty of maintenance, etc., its application is limited [3, 6].

Apart from compression techniques, another solution for alleviating transmission pressure is single fiber bi-direction, which allows simultaneous transmission of uplink and downlink on the same fiber. Moreover, wavelength-division multiplexing and microwave transmission further reduce fiber consumption [4].

Considering the methods mentioned above, we can discover that they are all designed to decrease data transmission in backhaul for the given amount of data. These compression techniques are effective in alleviating the data transmission pressure. However, is there any other way to facilitate these compression techniques to make the implementation of the centralization of H-CRANs more achievable? To be more specific, is it possible to decrease the amount of data to be transferred at the source for the RRHs with heavy load so that all backhauls only carry a moderate load? In cellular networks, a user scans all available channels and associates itself with an RRH that has the strongest received signal strength indicator (RSSI) by default, while being ignorant of the load of data to be transmitted by the RRHs. As users are typically not evenly distributed, some backhauls of RRHs tend to suffer from heavy load, which requires high transmission bandwidth, while their adjacent backhauls of RRHs may carry only light load, as can be seen in Fig. 2. The differences in the widths of backhaul lines reflect the gaps between transmission bandwidths required by different RRHs. Besides, as users are typically shifting between different areas, e.g. residential and office, during the day, peak transmission bandwidth requirements may be as much as 10 times higher than during off-peak hours. In order to guarantee robust data transmission at any time and any place, the transmission bandwidth of the backhaul for the RRH is designed for the peak transmission bandwidth required. As a result, transmission bandwidth resources are largely wasted, because most of the time these RRHs do not need to communicate with the BBU pool at full rate. The waste of transmission bandwidth is partly reflected in the waste of transmission links such as optical fibers that have been constructed. Obviously, if we balance the data transmitted in backhauls among RRHs in real time, the peak transmission bandwidth required can be greatly decreased. As a consequence, balancing the workload among backhauls followed by compression techniques can greatly reduce the amount of data to be transmitted. Besides, costs spent on construction of backhauls can be decreased, as the backhauls of RRHs can now be equipped with lower transmission bandwidth. Furthermore, thanks to the aggregation of BBUs into a BBU pool, the distribution of traffic demands can be easily obtained, making it possible to balance the workload among backhauls. Finally, as is known to all, the BBU pool is capable of workload balancing due to the aggregation of processing units and its natural capability of central management. However, the function of workload balancing executed in the BBU pool is mainly designed to balance the workloads among different processing units while being ignorant of the workloads of backhauls. For practical cellular networks, workload balancing among backhauls must be executed in the user association part among RRHs, which relies on the geometrical partitioning of the practical service area for the purpose of workload balancing, not just the BBU pool.

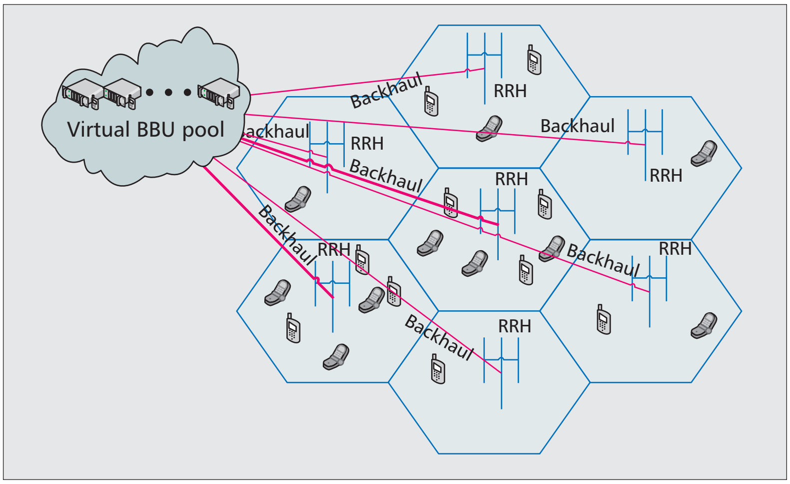

Figure 2. Illustration of unbalanced backhauls.

WORKLOAD BALANCING TO ALLEVIATE DATA TRANSMISSION PRESSURE ON BACKHAULS OF H-CRANS

Workload balancing has been studied widely in the literature. It is a way to balance the workload among various servers and machines to optimize factors such as resource utilization, fairness, waiting/processing delays, or throughput [12, 13]. The problem of balancing load among backhauls can be seen as the equitable location problem in a broad sense. By assigning the density of workload to each spot, the weighted equitable location problem is exactly the workload balancing problem [14]. The equitable location problem on a plane has been studied in the operational research field, which is generally designed to locate M facilities on a unit square so as to minimize the maximum demand faced by any facility subject to closest assignments and coverage constraints. The proposed strategies are usually based on local or global adjustments (depending on which strategy designers practically adopt) and a Voronoi diagram [12]. Given some number of points in the plane, their Voronoi diagram divides the plane according to the nearest-neighbor rule: each point is associated with the region of the plane closest to it. It can be observed that the strategy works quite effectively in continuous cases where facilities

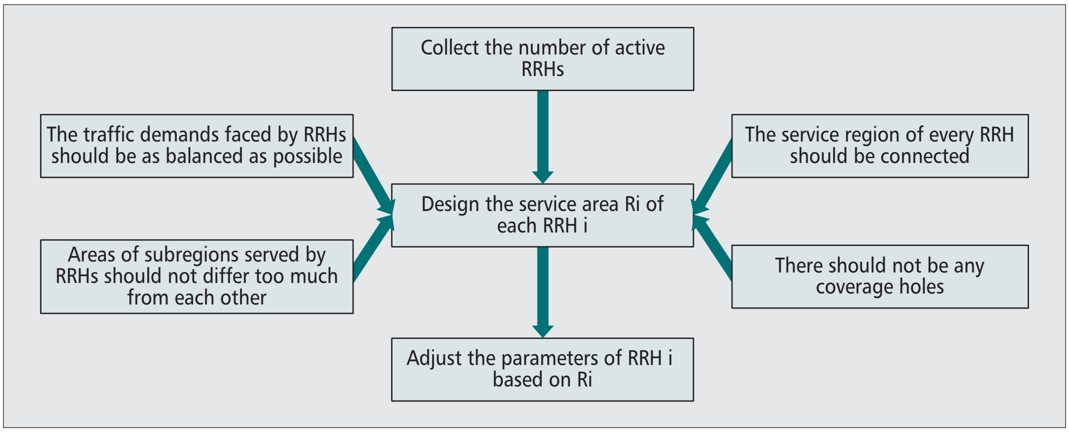

Figure 3. Flow chart of our proposal.

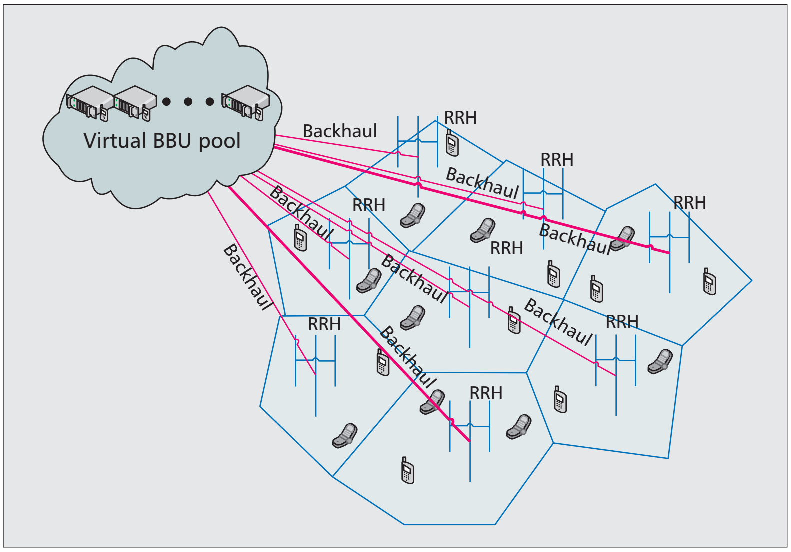

Figure 4. Illustration of workload balancing.

(referring to RRHs in H-CRANs) can move continuously in any direction or where the density of candidate sites for facilities is quite large. However, when the distribution of candidate sites cannot be approximated continuously, as is always the case with RRH locations in H-CRANs, it will not be so satisfactory or possibly even lose effect. In [14] a given region *R* is divided into *n* subregions so as to balance the overall utilities on the subregions, and it shows us that we can balance the workload among backhauls with similar algorithms.

It is found that if we model the service regions of RRHs as geometry regions on the plane and abstract the distribution of traffic demands at any time to be a normalized matrix whose elements represent the densities of traffic demands across the service regions, we can model the workload balancing task as a convex optimization problem [14].

To be more specific, the proposal to balance workload among backhauls can be summarized as follows:

- Collect the number of active RRHs, denoted as *n*, that serve the users in the objective region *R*.
- Design the service area R_i of each RRH *i*. Denote the distribution of traffic demands as $f(x)$. First, the traffic demands faced by the RRHs $\iint_{R_i} f(x)dA$ should be as balanced as possible, and this task is transformed into the objective function of the optimization problem, which minimizes the maximum traffic demands faced by the RRHs. This objective function aims to balance data transmissions on the backhauls of the RRHs as much as possible. Second, areas of subregions served by the RRHs $\iint_{R_i} dA$ should not differ too much from each other to avoid the case where some subregions are too large to be covered by only one RRH per subregion. It also help avoid yielding ill-shaped subregions. This task is transformed into a constraint in the optimization problem. Third, it is our natural desire that the service region of every RRH should be connected, which introduces a penalty function to punish our objective function by preventing it from getting to its optimality when subregions are far from connected. Fourth, in order to ensure that users at any place are provided with reliable communication service, there should not be any coverage holes, which serves as another constraint in the optimization problem [14]. After the formulation of the problem, we can either adopt mature algorithms such as the cutting plane method, or software such as cvx, to solve it, thus obtaining the service region of every RRH, which ensures that the data transmissions on every backhaul of the RRH do not differ too much from each other.

The flow chart and sketch map of workload balancing are presented in Fig. 3 and Fig. 4, respectively. The differences in the widths of backhaul lines reflect the gaps between transmission bandwidths required by the RRHs in Fig. 4. It can be easily determined that the balance of data transmission on backhauls results in the decrease of peak data transmission, which alleviates the high demand on the bandwidth of backhaul requested by data transmission. Experiments have proven that the bandwidth required can be decreased to less than 1/3 of the bandwidth needed before workload balancing, as shown in Fig. 5. The red line in Fig. 5 represents the ratio of peak transmission bandwidth required between schemes with and without workload balancing in 100 cases. Each case is tested with a scenario whose distribution of traffic demands differs greatly from other cases. Two backhaul allocation schemes are applied to the scenario in each case. First, a traditional scheme is applied to the case, where a user associates itself with the RRH that provides the strongest RSSI while being ignorant of the load of data to be transmitted by RRHs, and the peak transmission bandwidth required without workload balancing is calculated. Second, our proposal for workload balancing is applied to the case. The ratios between two peak transmission bandwidths required certify the feasibility of workload balancing on backhaul. Moreover, it can be seen from Fig. 5 that the ratios between two transmission bandwidths required are stable and they stay below 1/3 during the whole experiment, which indicates the reliability of our proposal. Furthermore, it needs to be especially noted that due to the aggregation of BBUs into the BBU pool in H-CRANs, the distribution of traffic demands at any time and any place can be obtained without much effort, and the feasibility of centralized processing and integrated planning further makes the proposal effective.

Finally, it is worth mentioning that our proposal to balance workload among backhauls is completely compatible with existing compression techniques. Compression techniques are all designed to decrease the data transmission on the backhaul for the given amount of data, while the workload balancing proposal is designed to decrease the amount of data to be transferred at the source for every RRH. To sum up, the workload balancing among backhauls proposed in this article can work in harmony with existing compression techniques to handle the BBU-RRH signal transmission bandwidth problems.

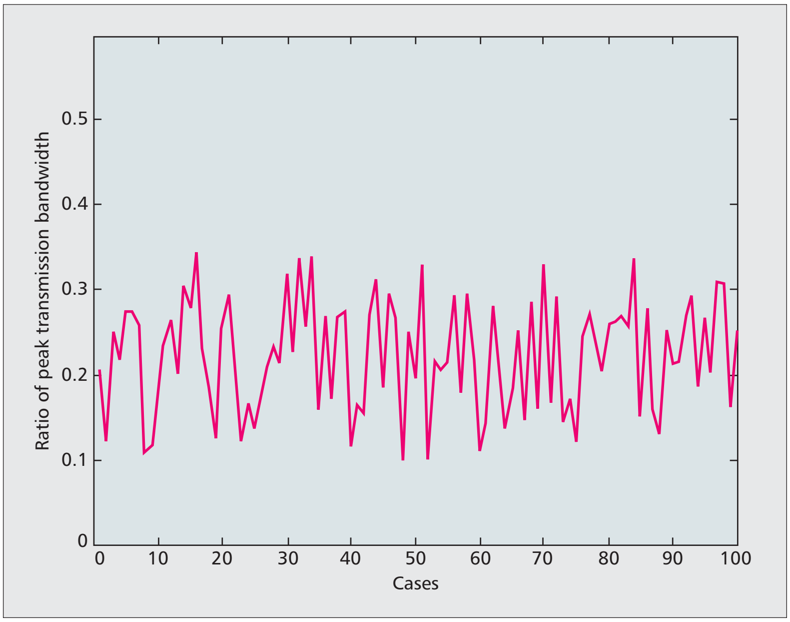

Figure 5. Ratio of peak transmission bandwidth required between schemes with and without workload balancing in 100 cases.

CHALLENGES AND RESEARCH DIRECTIONS

CHALLENGES

Based on the understanding of the workload balancing scheme proposed above, we would ask a question. H-CRANs are made up of RRHs as well as macro cells, and they can be seen as two-tier networks. Considering that our proposal above is mainly focused on one tier of H-CRANs (the tier of RRHs), we wonder why conducting workload balancing in one tier without considering the influence of other tiers (the tier of macrocells) makes sense. In practice, due to differences in deployment and propagation environments, RRHs and macro BSs have separate path loss exponents and spatial densities. RRHs together with macro BSs not only cooperate with each other to provide service, but also bring each other interference, which makes it more difficult to analyzing the characteristics of each tier and then balance the workload among backhauls. However, if we simplify the system based on a model that characterizes the two tiers of a H-CRAN by transmit power, node spatial density, path loss exponent, and bias factor, it has been proven that the outage probability of each tier is the same for all tiers, and it is even the same as the outage probability of the overall network. This indicates that we can balance the workload among backhauls of the RRHs without being concerned that macro cells would largely influence the plan's effectiveness [15]. However, how would we decide specifically whether the data transmission of a user is assigned to a macro BS or a RRH? That is to say, how should data transmissions be allocated among different tiers?

Besides, as the distribution of users changes dynamically in a day, and the RRH-user association is not applicable at various times time and the balance among data transmissions on backhauls is broken, new RRH-user associations need to be set up to maintain balance. During the process of changing RRH-user associations, much handover would be created, causing a significant increase in signaling overhead and harming the user experience. If no special actions are taken, robust communication cannot be guaranteed among users.

RESEARCH DIRECTIONS

Motivated by the challenges mentioned earlier, we can propose the following research directions.

First, how would we allocate data transmission between RRHs and macro cells in H-CRANs? As RRHs are usually deployed in traffic hot zones or dead zones to provide reliable communication service, a scheme that takes the density of traffic demands into consideration when allocating workload between RRHs and macro cells should be researched in the future.

Second, how would we handle the handover created during the change of RRH-user associations? When the original balance among backhauls is broken, a new balance needs to be set up based on the new distribution of traffic demands. If, when forming a new balance, we consider the distribution of traffic demands as well as the former RRH-user associations, the amount of handover can be decreased. Besides, with BBU centralization in H-CRANs, handover can be realized in a simplified procedure and improved performance can be achieved.

From a technical standpoint, the implementation of the proposal faces other challenges such as workload allocation between macro cells and RRHs, and handover incurred by the changes of the RRH-user associations to maintain balance, which should be investigated in detail in the future.

Last but not the least, it is worth mentioning that the workload balancing method proposed in this article does not only apply to backhauls in H-CRANs; it will also apply to ultra dense small cell networks in 5G, where the problem of backhauls remains a bottleneck in the implementation of densification of cells.

CONCLUSION

H-CRANs have been recommended as one of the most promising ways to handle the challenges faced by today's HetNets, such as explosive data growth, low energy efficiency, severe inter-tier interference, etc. Due to the assistance of various cloud computing-based cooperative techniques and resource management techniques, together with the MIMO technique to be installed on macro BSs, H-CRANs have the advantages of higher energy efficiency, much larger capacity, suppressed inter-tier interference, and adaptability to scalability. On the other hand, large costs of data transmission in backhauls caused by centralization make H-CRANs less feasible, and they will even become intolerable when the density of RRHs becomes ultra dense, which is expected in the coming 5G. Due to the unevenly distributed traffic and the huge gap between peak transmission rate and the transmission bandwidth required during off-peak hours, transmission bandwidth is not fully used most of the time since mobile operators must provide users with the bandwidth required during peak hours to guarantee robust communication quality. Such waste is partly reflected in the costs spent on the construction of backhauls. Various compression techniques have been proposed to decrease the amount of data to be transmitted on backhauls.

In this article we proposed to balance the data to be transmitted on backhauls of the RRHs so as to decrease the bandwidth of backhaul required by the RRHs or achieve better performance with given backhaul capacity in a different way from compression techniques. The key challenges of our proposal were discussed, as well as ideas to deal with them with some preliminary results. From a technical standpoint, the implementation of the proposal faces other challenges such as workload allocation between macro cells and RRHs, and handover incurred by changes in the RRH-user associations to maintain balance, which should be investigated in detail in the future.

REFERENCES

[1] M. Peng *et al.*, "Heterogeneous Cloud Radio Access Networks: A New Perspective for Enhancing Spectral and Energy Efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, Dec. 2014, pp. 126–35.

[2] A. Damnjanovic *et al.*, "A Survey on 3GPP Heterogeneous Networks," *IEEE Wireless Commun.*, vol. 18, no. 3, June 2011, pp. 10–21.

[3] A. Checko *et al.*, "Cloud RAN for Mobile Networks — A Technology Overview," *IEEE Commun. Surv. Tut.*, vol. 17, no. 1, 1 Quarter 2015, pp. 405–26.

[4] C.-L. I *et al.*, "Recent Progress on C-RAN Centralization and Cloudification," *IEEE Access*, vol. 2, Sept. 2014, pp. 1030–39.

[5] M. Peng *et al.*, "System Architecture and Key Technologies for 5G Heterogeneous Cloud Radio Access Networks," *IEEE Network*, vol. 29, no. 2, Mar. 2015, pp. 6–14.

[6] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN," Tech. Rep., Oct. 2011.

[7] P. Rost *et al.*, "Cloud Technologies for Flexible 5G Radio Access Networks," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014, pp. 68–76.

[8] D. Wubben *et al.*, "Benefits and Impact of Cloud 17 Computing on 5G Signal Processing: Flexible Centralization Through Cloud-RAN," *IEEE Signal Process. Mag.*, vol. 31, no. 6, Nov. 2014, pp. 35–44.

[9] W. H. Chin, Z. Fan, and R. Haines, "Emerging Technologies and Research Challenges for 5G Wireless Networks," *IEEE Wireless Commun.*, vol. 21, no. 2, Apr. 2014, pp. 106–12.

[10] P. K. Agyapong *et al.*, "Design Considerations for a 5G Network Architecture," *IEEE Commun. Mag.*, vol. 52, no. 11, Nov. 2014, pp. 65–75.

[11] S.-H. Park *et al.*, "Robust and Efficient Distributed Compression for Cloud Radio Access Networks," *IEEE Trans. Vehic. Tech.*, vol. 62, no. 2, Feb. 2013, pp. 692–703.

[12] O. Baron *et al.*, "The Equitable Location Problem on the Plane," *Eur. J. Oper. Res.*, vol. 183, no. 2, Dec. 2007, pp. 578–90.

[13] J. G. Andrews *et al.*, "An Overview of Load Balancing in HetNets: Old Myths and Open Problems," *IEEE Wireless Commun.*, vol. 21, no. 2, Apr. 2014, pp. 18–25.

[14] J. G. Carlsson and R. Devulapalli, "Shadow Prices in Territory Division," University of Minnesota, available at: http://menet. umn. edu/˜jgc/shadow-prices-rev2. pdf, 2013.

[15] H. S. Dhillon *et al.*, "Modeling and Analysis of K-Tier Downlink Heterogeneous Cellular Networks," *IEEE JSAC*, vol. 30, no. 3, Apr. 2012, pp. 550–60.

BIOGRAPHIES

SHAOWEI WANG (wangsw@nju.edu.cn) is a professor at Nanjing University. He received his Ph.D. degree from Wuhan University, China, in 2006, and joined the School of Electronic Science and Engineering at Nanjing University in the same year. From 1997 to 2001 he was with China Telecom as an R & D scientist. From 2012 to 2013 he was with Stanford University and The University of British Columbia as a visiting scholar/professor. His research focuses on wireless communications and networking. In these areas he has published more than 60 papers in leading journals and conference proceedings. He organized a special issue on Enhancing Spectral Efficiency for LTE-Advanced and Beyond Cellular Networks for *IEEE Wireless Communications*, and a feature topic on Energy-Efficient Cognitive Radio Networks for *IEEE Communications Magazine*. He is on the editorial board of *IEEE Communications Magazine and IEEE Transactions on Wireless Communications*, and serves/has served on the technical or executive committee of reputable conferences including *IEEE INFOCOM*, *IEEE GLOBECOM*, *IEEE ICC*, *IEEE WCNC* etc.

CHEN RAN (rchnju@gmail.com) received the BS degree in electronic science and engineering from Nanjing University, Nanjing, China, in 2013. She is currently pursuing the M.S. degree at the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. Her research interests include wireless communications and optimization. Currently, her research focuses on cellular network planning in wireless networks.

CHONGGANG WANG (cgwang@ieee.org) received the Ph.D. degree from Beijing University of Posts and Telecommunications (BUPT), China, in 2002. He is currently a member of technical staff at InterDigital Communications with a focus on Internet of Things (IoT) R&D activities, including technology development and standardization. His current research interest includes IoT, mobile communication and computing, and big data management and analytics. He has (co-)authored more than 100 journal/conference papers and book chapters. He is the founding Editor-in-Chief of *IEEE Internet of Things Journal* and is on the editorial boards of several journals, including *IEEE Access*.