Traffic Scheduling in Software-defined Backhaul Network

Hao Li

Department of Electronic and Electrical Engineering
University of Sheffield

This thesis is submitted for the approval of the
Doctor of Philosophy

July 2018

I would like to dedicate this thesis and everything I do to my parents. I would not be who I am today without their love and support.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Jie Zhang for the continuous support of my PhD study and related researches, for his patience, encouragement and guidance. He always encouraged me to go deeper in interesting research topics. Besides my supervisor, I would like to thank Dr. Xiaoli Chu for her great support in my entire study.

My sincere thanks also goes to Dr. Jiliang Zhang, who set a good example for me to be a researcher. His passion and earnest attitude affected me deeply. He also contributed in some collaborated works in this thesis. Without his valuable suggestions and comments, I can barely complete my researches.

A very special gratitude goes out to my girlfriend Ms. Huey Teng Koh, who have provided me emotional support in my daily life. She has always been there for me even in my difficult and trying times.

I thank my colleagues, Mr. Qi Hong, Ms. Hui Zheng, Mr. Baoling Zhang, Mr. Weijie Qi, Mr. Kan Lin, Ms. Fan Wu, Mr. Haonan Hu, Mr. Yuan Gao, and Mr. Bozhong Chen, for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last four years. A special mention goes to Mr. Lewis Chung, Ms. Lin Chen, and Ms. Tiantian Ding, who have shared many happy memories in my leisure time.

In addition, please allow me to express my sincere appreciation to the examiners of this thesis, Prof. Zhili Sun and Dr. Petros Karadimas, for the most valuable suggestions and corrections to improve the quality of the thesis.

I would also want to send my special thanks to my faculty for all the supports throughout my studies.

Last but not the least, I would like to thank my family: my parents, my aunt and uncle, for supporting me spiritually throughout my life.

Abstract

In the past few years, severe challenges have arisen for network operators, as explosive growth and service differentiation in data demands require an increasing number of network capacity as well as dynamic traffic management. To adapt to the network densification, wireless backhaul solution is attracting more and more attentions due to its flexible deployment. Meanwhile, the software-defined network (SDN) proposes an promising architecture that can achieve dynamic control and management for various functionalities. In this case, by applying the SDN architecture to wireless backhaul networks, the traffic scheduling functionality may satisfy the ever-increasing and differentiated traffic demands.

To tackle the traffic demand challenges, traffic scheduling for software-defined backhaul networks (SDBN) is investigated from three aspects in this thesis. In the first aspect, various virtual networks based on service types are embedded to the same wireless backhaul infrastructure. An algorithm, named VNE-SDBN, is proposed to solve the virtual network embedding (VNE) problem to improve the performance of the revenue of infrastructure providers and virtual network request acceptance ratio by exploiting the unique characteristics of SDBNs. In the second aspect, incoming traffic is scheduled online by joint routing and resource allocation approach in backhaul networks operated in low-frequency microwave (LFM) and those operated in millimetre wave (mmW). A digraph-based greedy algorithm (DBGA) is proposed considering the relationship between the degrees of vertices in the constructed interference digraph and system throughput with low complexity. In the third aspect, quality-of-service is provided in terms of delay and throughput with two proposed algorithms for backhaul networks with insufficient spectral resources. At last, as a trial research on E-band, a conceptual adaptive modulation system with channel estimation based on rain rate for E-band SDBN is proposed to exploit the rain attenuation feature of E-band.

The results of the research works are mainly achieved through heuristic algorithms. Genetic algorithm, which is a meta-heuristic algorithm, is employed to obtain near-optimal solutions to the proposed NP-hard problems. Low complexity greedy algorithms are developed based on the specific problem analysis. Finally, the evaluation of proposed systems and algorithms are performed through numerical simulations. Simulations for backhaul networks with respect to VNE, routing and resource allocation are developed.

List of Publications

Published

1. **H. Li**, J. Zhang, Q. Hong, H. Zheng, and J. Zhang, "Exploiting adaptive modulation in E-band software-defined backhaul network," in *2018 IEEE 8th IEEE Annual Computing and Communication Workshop and Conference(CCWC)*, pp. 1009-1013, 2018.

2. **H. Li**, J. Zhang, Q. Hong, H. Zheng, and J. Zhang, "Digraph-based joint routing and resource allocation in software-defined backhaul networks," in *2017 IEEE 22nd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 1-5, 2017.

3. **H. Li**, J. Zhang, Q. Hong, H. Zheng, and J. Zhang, "QoS-aware channel-width adaptation in wireless mesh networks," in *2016 IEEE International Conference on Communications (ICC)*, pp. 1-6, 2016.

Submitted

4. **H. Li**, J. Zhang, and J. Zhang, "Virtual network embedding in software-defined backhaul networks," submitted to *IEEE Transactions on Network and Service Management*, 2017.

Table of contents

List of figures

List of tables

List of Abbreviations

BS	Base Station
CSI	Channel State Information
GA	Genetic Algorithm
HetNet	Heterogeneous Network
InP	Infrastructure Provider
IoT	Internet of Things
ISP	Internet Service Provider
LFM	Low Frequency Microwave
LOS	Line-of-Sight
$LTE-A$	Long Term Evolution Advanced
MIMO	Multi-Input Multi-Output
mmW	Millimetre Wave
NLOS	None-Line-of-Sight
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
P2MP	Point-to-MultiPoint
P2P	Point-to-Point

QoS	Quality of Service
RB	Resource Block
SDBN	Software-defined backhaul network
SDN	Software-Defined Network
SINR	Signal to Noise Ratio
SN	Substrate Network
SP	Service Provider
TDMA	Time Division Multiple Access
UE	User Equipment
VLiM	Virtual Link Mapping
VN	Virtual Network
VNE	Virtual Network Embedding
VNoM	Virtual Node Mapping
VNR	Virtual Network Request
VoIP	Voice over IP

Chapter 1

Introduction

Overview

In this chapter, the background of the thesis will be introduced. From mobile networks to backhaul networks, the explosive traffic data-increment issue presents severe challenges to network operators. Service providers also demand tailored requirements from the wireless networks. Software-defined network (SDN) proposes a promising architecture to tackle these challenges. Thus, the motivation of the thesis is to investigate the traffic scheduling in wireless backhaul networks taking advantage of the SDN features. Then, the principle objectives will be proposed. Finally, the structure of the thesis will be presented together with an overview of the contributions this thesis has made.

1.1 Background and Motivation

In this section, the trends in mobile network development will be presented. Challenges and problems will be analysed generally. The necessity of wireless backhaul network and software-defined network will be explained, together with basic concepts of both aspects. The motivation for this work will also be explained.

1.1.1 Mobile Networks

Mobile phones have become an essential tool in daily life all over the world. With the development of modern phones, the features of mobile networks are not only providing voice calls but also data services like web browsing, video streaming and mobile games among other facilities. Moreover, the data services are accounting for a growing proportion of mobile phone usage. As a consequence, the last decade has witnessed an exponential increase of data demands [1]. The huge amount of data traffic has created a severe challenge

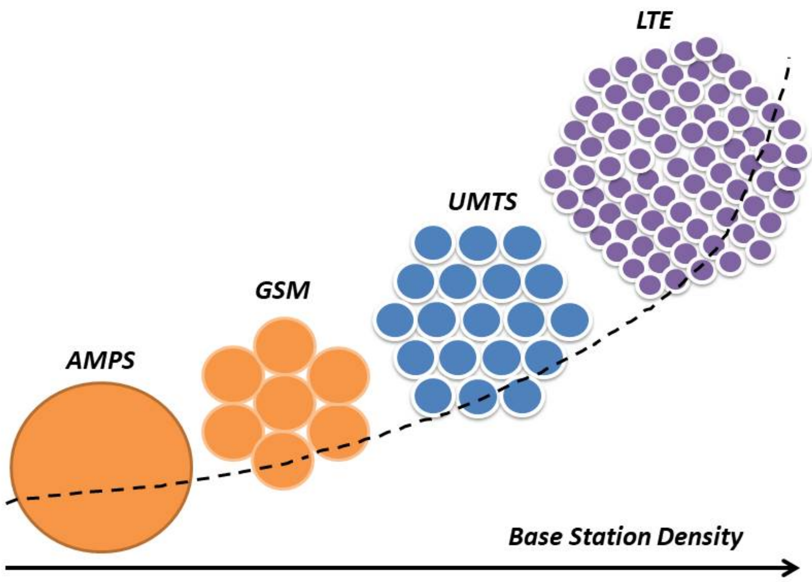

Fig. 1.1 Increasing base station density

for network operators as they aim to satisfy all services with limited spectral resources. In particular, with the development of the Internet of Things (IoT) [2], the data that goes through mobile networks is expected to continue to grow rapidly. Based on the network functions, the mobile network can be divided into two categories: the access network and the backhaul network. The access network is responsible for serving user equipments (UEs) with the base stations. The backhaul network is deployed to connect the base stations to the core network.

To solve the conflicts between data demand and limited spectral resources, small cell techniques are accepted as one of the promising solutions in the access networks [3]. To relieve the pressure of macro base stations (BSs) , small cell base stations are deployed to increase the total capacity and offer better quality-of-service (QoS) . This multi-tier mobile network structure is known as heterogeneous network (HetNet) , which has been widely used around the world as an important concept in Long Term Evolution Advanced (LTE-A) . From the capacity point of view, when increasing the number of small cells that are deployed, the network capacity can be raised nearly linearly if interference is well-managed. Thus, small cell densification is leading a trend in the future generation of network planning as shown in Figure 1.1.

However, the dense deployment of small cells has created new problems. In conventional mobile networks, the base stations backhaul the traffic data to the core network through wired

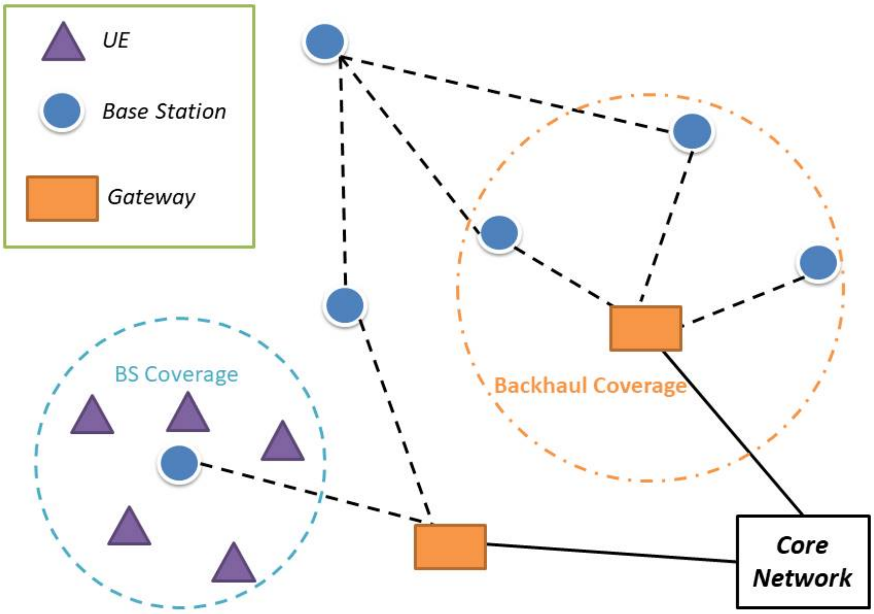

Fig. 1.2 Backhaul network architecture

links, such as copper and fibre. The cost of deploying wired links was acceptable in early years as the number of base stations was limited. The planning of backhaul network is not an issue. However, with increasing numbers of small cell base stations being deployed, the cost and difficulty of installing wired links also rises. More importantly, with the development of small cell techniques, the requirements of small-cell base station deployment on simplicity and flexibility have been improved. Self-backhauled small-cell base stations are taking over greater proportions of the deployments, and wireless backhaul solutions are attracting increasingly more attentions from researchers and mobile network operators.

1.1.2 Wireless Backhaul Networks

The wireless backhaul network is comprised of radio links that connect the base stations to the core network. Figure 1.2 shows an example of the wireless backhaul architecture. Some principle components play important roles in the backhaul networks. The definitions of the components and related concepts are specified as follows.

Gateway

The gateway is a central aggregation point that collects the traffic from different base stations and delivers it to the core network. The gateway is installed with wired backhaul links that directly connect to the core network. In general, the capacity of the wired link is considered to be sufficient for most cases. The gateway can also improve the scalability of the network and provide control and user plane functionalities [4].

Backhaul Coverage

The backhaul coverage is referred to the coverage of the gateway or the relay. The coverage of the gateway is the area where the base stations can directly communicate with the gateway. The coverage of the relay is the area where the base stations have direct links connecting to the relay. Compared with the access network coverage, we can regard the gateways or relays as the base stations in the access network. The base stations in the backhaul network can be seen as the UEs in the access network.

Backhaul Topology

Multi-hop backhaul is required if a base station is not in the backhaul coverage of any gateway. The traffic data can be relayed by other base stations to reach the gateway. There may exist various relay choices for one specific base station, as it may be located in the backhaul coverage of several relays. Thus, the topology of a wireless backhaul network can be similar to wireless mesh networks due to the web-like connectivity.

Carrier Frequency

Wireless backhaul solutions can be mainly classified into two categories by carrier frequency: low frequency microwave (LFM) and millimetre wave (mmW) . Although LFM techniques have matured after years of development, the capacity that it can offer cannot guarantee to satisfy the increasing traffic demand completely. Meanwhile, mmW has the potential to deliver high capacity with large undeveloped bandwidth. However, compared to LFM, mmW has much larger atmospheric attenuation, which limits the transmission range of the link. Among different frequency bands of mmW, E-band, which is commonly referred to $71-76$ GHz and $81-86$ GHz (International Telecommunication Union stipulation), has a relatively low atmospheric attenuation, which makes a promising choice for future wireless backhaul networks.

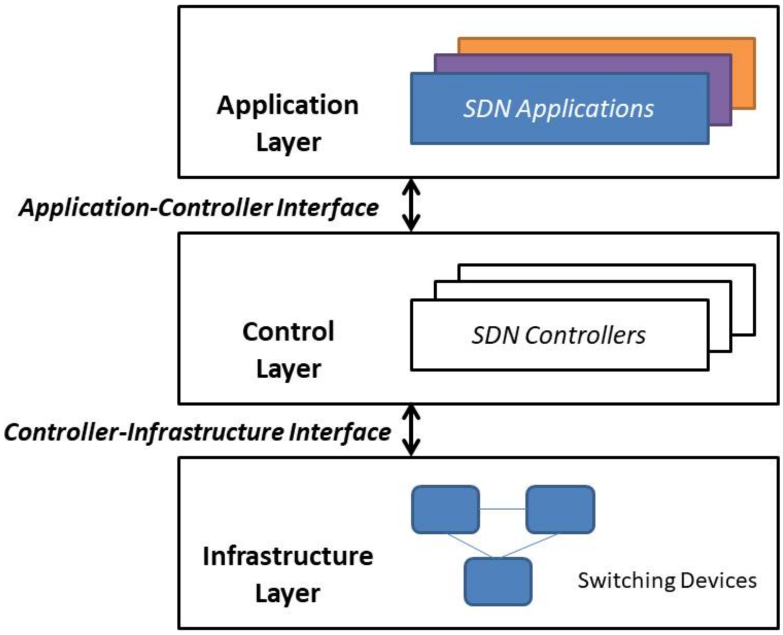

Fig. 1.3 Software-defined network three-layer model

Backhaul Planning

When planning a wireless backhaul network, requirements such as capacity, latency, coverage and availability should be strictly examined [4]. As backhaul networks are providing connections between the access networks and the core network, the backhaul performance should not limit the QoS of the access networks. Furthermore, with the trend being that backhaul units are often co-located with small cells, coverage and connectivity should be carefully designed.

1.1.3 Software-defined Networks

Emerging trends like network densification and service differentiation are commanding new challenges to future network architecture. Owing to its simplified and dynamic management, high flexibility and improved performance, SDN has drawn considerable attention in recent years [5–7]. The typical three-layer model of SDN is shown in Figure 1.3. By separating the control plane and the data plane, SDN can offer logical centralisation of network management over distributed switching devices and introduce programmability, which opens up new approaches to control functions in the application layer.

Control Plane

The control plane is responsible for the network management with the aid of SDN controllers. The SDN controller serves as a logically centralised intelligence in the SDN structure. Thus, the SDN controller has a global view of the entire network. Thanks to programmability [8], the SDN controller also has the ability to control the various functions in the application layer individually and dynamically. The SDN applications are programs that directly make requests and report their behaviors to SDN controller. Different functionalities can be provided with separate programmed applications, such as network protocols, network monitoring, and network reconfiguration.

Data Plane

The data plane is generally responsible for forwarding the traffic flow between switching devices based on the rules that are provided by the control plane. The switching devices are also responsible for collecting the network state information and reporting to the control plane.

Research Gaps

Many aspects of the mobile networks have been studied combining the SDN architecture. Function virtualisation has been comprehensively stated in survey [9]. Cellular networks based on SDN structure have been studied in [10]. Researches on software-defined optical networks have also been thoroughly summarized in [11]. Wireless sensor networks combining SDN have been investigated in the study conducted by [12]. From the operators' point of view, advanced SDN has been researched in [13]. However, research works in SDN-based wireless backhaul network are rarely seen.

1.1.4 Motivation

The main function of backhaul network is to direct the traffic flow from the access network to the core network. Traffic scheduling is essential in wireless backhaul networks. In wireless backhaul networks, traffic scheduling is basically consisted of a routing strategy, resource allocation and QoS provisioning, which aims at optimising the performance of the network [14]. Nowadays, the traffic requirements from different services are becoming multifarious. For example, mobile games are strict with network latency. Video streaming is demanding large bandwidth. It is therefore a severe challenge for the backhaul network to load the tremendous amount of traffic data while satisfying various traffic demands.

To deal with the large amount of data and tailored requirements of various services, the software-defined network architecture stands out as a promising solution due to its control and data plane separation, and programmability. By applying software-defined network architecture to wireless backhaul network, different services can be separately managed. Furthermore, the SDN controller has a global view over all the traffics flowing in the backhaul network, which means that all services in the network share all the information and can cooperate to improve the network performance while meeting their own requirements.

Thus, this thesis is mainly concerned with the traffic scheduling in software-defined backhaul networks (SDBN) . In order to satisfy various requirements from different services, network virtualization is an effective technique, which can form different virtual networks based on various service requirements. The problem of virtual network embedding (VNE) should be investigated in the SDBN. Then, to satisfy the traffic demands and improve the throughput of the network, the routing and resource allocation problem is studied in both LFM-based and mmW-based wireless backhaul networks. To guarantee the QoS of the diverse services, service delay and bandwidth allocation should be carefully handled in the backhaul network. At last, as a promising frequency band for future backhaul network, E-band has the potential to build large-capacity and high-quality backhaul links. The features of E-band can be taken advantage of to enhance the network capacity.

1.2 Objectives of the Thesis

As a result, the principal objectives of the thesis can be summarised as follows:

- To overview the current VNE, routing, resource allocation and QoS provisioning approaches in the mobile networks, discuss their applicability in SDBNs, and introduce the motivation for new approaches that are more suited for SDBNs.

- To build models for SDBNs in terms of VNE, routing, resource allocation and QoS provisioning aspects. This objective should take advantage of the SDN features to reduce the difficulty of corresponding problems.

- To propose novel algorithms that can improve the network performances for traffic engineering in SDBN.

- To implement the models and algorithms in a simulation tool to verify the enhanced performances in SDBN with respect to other approaches.

1.3 Contributions of the Thesis

Chapter 3: Virtual Network Embedding in Software-defined Backhaul Networks

1) The VNE problem is modelled in the SDBN. Based on backhaul network features, the virtual node mapping problem is divided into serving nodes, gateway nodes and caching nodes mapping. 2) Caching technique is considered in VNE problems. Caching gain metric for backhaul network is proposed. 3) For virtual link mapping problem, two metrics, interference properties and available resources, are proposed to evaluate the substrate link. 4) The VNE-SDBN algorithm is proposed to solve the developed problem for SDBN.

Chapter 4: Joint Routing and Resource Allocation in Software-defined Backhaul Networks

1) Routing and resource allocation functions module is proposed for SDBN. LFM and mmW scenarios are both considered. 2) System throughput is given in two forms of equations. 3) Two cost function based on digraph are proposed to evaluate the routing process. 4) The DBGA algorithm is proposed to improve the system throughput of SDBN.

Chapter 5: Quality of Service Provisioning in Software-defined Backhaul Networks

1) QoS module for SDBN is proposed. 2) The queueing sequential order adjustment is introduced to improve the routing successful rate. 3) Delay-aware routing algorithm and QoS-aware bandwidth allocation greedy algorithm are proposed to provide QoS support.

Chapter 6: Exploiting Adaptive Modulation in E-band Software-defined Backhaul Networks

1) Adaptive modulation is introduced to enhance the capacity of networks operated in E-band, countering its rain attenuation problem. 2) Taking advantage of IoT, rain rate data based channel estimation is proposed. 3) Rain rate data update algorithm is proposed based on the Markov chain of rain rate. 4) System performance is evaluated under various error levels.

1.4 Structure of the Thesis

The thesis is organised by the following layout:

Chapter 2: State of the Art and Research Challenges

This chapter introduces the background knowledge and related works of VNE, routing, resource allocation, QoS provisioning and E-band characteristics. Corresponding analysis and comments are presented about the applicability in SDBNs.

Chapter 3: Virtual Network Embedding in Software-defined Backhaul Networks

The VNE problem in OFDMA-based SDBN, where caching technique is employed, is investigated. Virtual node mapping and virtual link mapping are separately modelled based on the features of SDBN. In the proposed VNE-SDBN algorithm, serving nodes and gateway nodes of the backhaul network are mapped based on location information of virtual network request (VNR) , while the caching nodes are located to maximise the caching gain of the backhaul network. For the virtual link mapping problem, interference property and available resources of links are taken into consideration. Numerical results have shown the improvements in the revenue and the VNR acceptance ratio with the proposed algorithm.

Chapter 4: Joint Routing and Resource Allocation in Software-defined Backhaul Networks

The joint routing and resource allocation in an OFDMA-based SDBN is investigated. Based on the proposed SDBN system model, the joint routing and resource allocation problem is formulated as a ‘system throughput optimisation problem’. To solve the problem in a decomposition manner, an interference digraph was constructed. A low complexity greedy algorithm is proposed based on the indegree and outdegree of vertices in the digraph. Simulation results are presented comparing with benchmark genetic algorithm.

Chapter 5: Quality of Service Provisioning in Software-defined Backhaul Networks

The QoS provisioning problem in OFDMA-based SDBN is investigated. As two main QoS aspects in backhaul networks, delay and bandwidth are emphatically analysed. Delay-aware routing algorithm is proposed to guarantee various end-to-end delay requirements. A greedy algorithm for the QoS-aware bandwidth allocation problem is proposed to improve the QoS in resource-limited scenarios. Numerical results also validate the performance of the proposed algorithms.

Chapter 6: Exploiting Adaptive Modulation in E-band Software-defined Backhaul Networks

Adaptive modulation is introduced into E-band SDBN. With accurate channel model, a system model that where the rain rate data is utilised as channel estimation in the adaptive modulation scheme has been proposed. The system performance is evaluated under various error levels caused by channel model and rain rate data inaccuracy.

Chapter 7: Conclusions and Future Work

The concluding remarks are presented with directions for the future work.

Chapter 2

State of the Art and Research Challenges

Overview

In this chapter, the topics related to traffic scheduling in software-defined backhaul networks are reviewed. Specifically, state of the art in virtual network embedding, routing, resource allocation and QoS provisioning are presented. Related works are discussed and further research possibilities and directions are given. The characteristics of E-band are also discussed due to its potential to establish high-capacity backhaul links.

2.1 Reviews of Virtual Network Embedding

In this section, the basic concepts and problems of virtual network embedding will be introduced. Then, the related works in this research field will be presented with comments and discussions. At last, the research gap for virtual network embedding problem in software-defined backhaul network will be pointed out.

2.1.1 Fundamental Concepts

Based on the Infrastructure as a Service [15] business model, the role of the current Internet Service Provider (ISP) will be decoupled into the infrastructure provider (InP) and the service provider (SP) . InPs are responsible for deploying and maintaining the physical hardware of the network. SPs rent part of the infrastructure according to its own demands. Thus, the components and resources of the network infrastructure need to be abstracted for SPs to utilise.

Network virtualization [16–18] has been proposed as enabling technologies for the next generation of network services. In network virtualisation, two fundamental entities are the virtual network (VN) and the substrate network (SN) . All the nodes and links in the

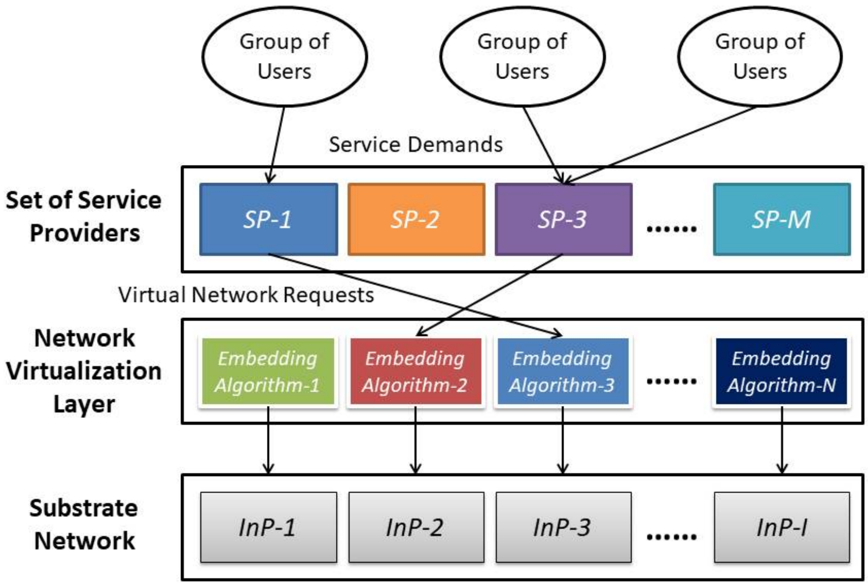

Fig. 2.1 Virtual network embedding diagram

network infrastructure form the SN, which is the set of physical hardware, while the VN is a combination of certain nodes and links. In VNs, virtual topologies are produced by connecting virtual nodes with virtual links according to the varying VNRs from different SPs. By embedding various VNs into the SN, different services with widely varying characteristics can co-exist in the same physical infrastructure. Fig. 2.1 shows a diagram of virtual network embedding to the substrate network.

However, the VNE problem, which is to embed various virtual networks onto the substrate network, remains a severe resource allocation challenge in network virtualisation [19]. Based on the objectives of embedding, the VNE problem can be divided into two sub-problems: virtual node mapping (VNoM) problem and virtual link mapping (VLiM) problem. In VNoM procedure, the virtual nodes are located to the substrate nodes. In VLiM procedure, the virtual links are mapped onto paths that connect the corresponding mapped substrate nodes. In both procedures, the resources of the candidate substrate nodes and links should be able to satisfy the requests of the virtual resources, where in most cases the substrate resources are partitioned to host several virtual resources of the same type.

Generally, in virtual network embedding problem, the virtual network request is modelled by node and link demands, while the substrate network consists of node and link resources. Then the embedding procedure is to satisfy the demands with resources. Once all demands have been met, the virtual network embedding procedure ends with a successful solution.

The VNE problem is related to the multi-way separator problem [20]. Thus, it is *NP*-hard to solve. If the VNoM solution has been given, the VLiM problem is reduced to the unsplittable flow problem [21, 22], which is also *NP*-hard. Thus, exact solutions for VNE only exist in problems with simple scenarios. Most of works in this research field are based on heuristic or meta-heuristic approaches [23].

VNE aims to find optimal solutions for VNoM and VLiM in consideration of single or multiple objectives. Some common objectives is listed as follows: 1) Maximise the profit of InPs; 2) Provide survivability; 3) Enhance QoS.

2.1.2 Related Works and Discussions

As an enabling technology for future Internet, the VNE area has attracted attention from many academic researchers.

A survey-based study [23] has comprehensively presented the research efforts before 2013 in the VNE area. The classification of current VNE algorithms was provided and research opportunities were discussed. The literature pointed out that solving VNE problem in wireless networks remains a challenge, as the interference caused by the broadcast nature of wireless communication needs careful consideration.

Exact solutions for VNE before 2016 was summarised in another survey-based study [24]. The formulation, implementation and performance of VNR exact solution were presented in detail. The authors pointed out that the exact solution is the foundation of heuristic and meta-heuristic solutions. However, the limit of the exact solution is conspicuous, in that it is not suitable for on-line dynamic VNE in complex scenarios.

Authors in [25] proposed that the topological attributes of neighbourhood nodes impact the performance of VNE. Based on the node degree and clustering coefficient information, VNE-DCC was proposed to enhance the resource utilisation of the substrate network. Through simulations, the proposed algorithm was verified to improve the Revenue/Cost and acceptance ratio, and the metric of node importance was also validated as a useful tool to solve VNE problems. The limitation of the work is that the authors did not consider the wireless network scenario.

Besides node capacity and link bandwidth, the location constraint was taken into consideration in [26]. Based on Integer Linear Programming, an exact VNE algorithm ILP-LC was proposed, which aimed at improving the acceptance ratio of VNRs. The performance of the algorithm outperformed typical heuristic algorithms by at least 15%. However, the work was based on wired VNE and the computational complexity of the algorithm was large.

The VNE is investigated on substrate network resources sharing basis in [27]. Dynamic resource splitting and sharing, which means that the resource blocks of each substrate node/link are split and shared by multiple virtual nodes/links, was introduced. The formulation of an optimisation problem was presented and an LP-based optimal solution RBS-LP was proposed, with which higher node/link utilization was achieved. However, the work was also based on wired networks.

Authors in [28] studied the VNE problem in mobile wireless networks. Countering the mobility and migration problem of mobile user equipment, an approximation VNE algorithm BIRD-VNE was proposed. The algorithm aimed at narrowing the solution search space, avoiding backtracking while maintaining the acceptance ratio of the VNE solutions. This work takes the node mobility and link instability into consideration, which are common in mobile networks. The work made great progress towards the application of VNE to mobile networks.

The researchers in [29] also considered the VNE to support the actual user mobility in mobile networks. The impact of mobility on the performance and efficiency of the networks that applied VNE was revealed by numerical investigations. The proposed mobility aware algorithm utilised the mobility information and would take the next mobility anchor point where the user was moving towards into consideration. The work contributed to VNE research in the mobility aspect of the mobile network. However, this work emphasised routing and ignored the resource allocation aspect.

Survivability was mainly investigated in VNE problem in Cloud's backbone network [30]. The authors proposed a new reactive VNE algorithm Advanced-CG-VNE based on game theory. The algorithm took advantage of reliable physical resources and would then re-embed the failed virtual resources. Both unsplittable and splittable virtual link mapping were considered. The numerical results showed the improvement of the proposed algorithm in terms of rejection rate of VNs and rate of VNs impacted by physical failures. Wired backbone was considered in this work, which is not applicable in our research objectives.

To maximise the profitability of advertisement targeting, a novel design framework for location-recommendation-aware VNE in optical-wireless hybrid networks was proposed by [31]. The interdependency among user groups was utilised to decide into which data centre the embedded virtual network should be pushed. As a result, the advisor could push the corresponding products to the user groups in the single DC. Greater energy efficiency and more profit were achieved by the proposed algorithm. The location recommendation concept can be used for reference.

Node and link mapping were solved in one-shot based on optimisation theory and path generation in [32]. The VNE solution first obtained an initial solution based on node and link coordination. Then the solution was improved in the final solution by performing pricing for the dual variables. The acceptance ratio and VNE revenue results were obtained with extensive simulations and the proposed approach performed better over other algorithms. To solve the VNE problem in the one-shot manner is innovative.

An approach to generate simulation scenarios for VNE algorithm evaluation was proposed in [33]. The authors also introduced concepts to control the randomness in evaluation scenarios and discussed solvable scenarios with two generation elements. This work helps researchers to evaluate the quality of VNE algorithms accurately with less time, which may be applied to the simulations in the current study.

Location-constrained virtual network embedding was investigated in [34]. The *NP*-completeness of the problem was proved by utilising graph bisection. Two heuristic algorithms were proposed based on compatibility graph to one-shot node and link mapping. Reduced time complexity, lower blocking probabilities and higher time-average revenue were provided. The location constraint can be employed in backhaul network scenario where user equipment is served by base stations located in a fixed location.

The VNE problem was studied in the integrated wireless and wired domain 5G network infrastructure [35]. A novel heuristic VNE algorithm was proposed based on the layered-substrate-resource auxiliary graph and six-quadrant service-type-judgement method. Compared with benchmark, the proposed algorithm performed better in average blocking rate, average latency and substrate resource efficiency. The layered network structure may be applicable for various VNs of different service types in the backhaul network.

Authors in [36] proposed a general distributed auction mechanism for VNE problem. The proposed consensus-based auctions include three procedures: discovery, virtual network mapping and allocation. The convergence and performance proved to be bounded with CAD. The auction mechanism can be applied to the competition for resources among different services. However, the work is based on distributed manner.

To conclude, many efforts have been made to solve the VNE problem in wired networks. However, the investigation into wireless networks is rarely seen. In particular, as far as we know, no work has been conducted on wireless backhaul networks considering dynamic resource allocation and interference between links. Besides, with the proposal of caching technique benefiting the backhaul network, new challenges has occurred in VNE.

2.2 Reviews of Routing

As a basic sub-problem of traffic scheduling, routing has been studied for many years in different network architectures [37]. Routing is the procedure of determining a path for traffic in networks. If more than one path exists, the best path is selected with the aid of routing metrics. In this section, some typical routing metrics will be introduced and related works will be presented and discussed.

2.2.1 Typical Routing Metrics

Some typical routing metrics are presented in details in the following.

Hop count

Owing to its simplicity, hop count is the most widely used metric. It is based on the assumption that every link counts as one equal unit, ignoring any characteristics of the link. The least-hop routing algorithm is a straightforward implementation of this metric, which selects the path with least hop count from source to destination. If several paths all have the minimal number of hops, an arbitrary path is chosen. Widest Shortest Path [38] was proposed to deal with situation, where the path with largest bandwidth would be chosen.

Geographical Information

In ad hoc networks, usually the geographical position of nodes is assumed as known to neighbours. Such information can be used to design routing algorithms with particular applications.

Another metric known as geographical distance is also widely used, which is often combined with other metrics, as distance information is easy to acquire and has an impact on signal strength and delay. The least-distance routing algorithm is based on this metric.

Bandwidth

Bandwidth is an important metric in throughput-demanding applications [39]. The minimal available bandwidth of links in the path can indicate the route capacity that is the upper bound of the traffic throughput. Thus, selecting paths with larger bandwidths can guarantee the traffic throughput to be maximised. Sometimes, the bandwidth metric can be employed to avoid traffic congestion in the network because the available bandwidth can tell if the link is overloaded.

Traffic Load

Another metric that can help relieve the traffic congestion situation in the network is the load in links. The larger the load of one link is the more chance of traffic congestion occurring. Thus, the path with least loads on links should be selected. Load balancing can also be performed based on this metric. Moreover, sometimes higher network capacity may be achieved because the load of certain links is diverted and more traffic can be loaded through those links.

Besides the load of links, queue length is a metric that can also indicate the traffic load from the nodes perspective [40]. Each node is equipped a processor and buffer to deal with traffic packets. More traffic can be assigned to nodes with shorter queues.

Delay

The delay metric is a measurement of the transmission time for one data packet from the source node to the destination node. It is a vital metric in latency-crucial applications. This metric can also be applied to provide QoS for various services that have requirements on network delay.

Delay variation is a metric to indicate the jitter situation of paths. Applications with strict real-time restrictions (e.g. video streaming) may utilise this metric to smooth transmissions.

Survivability

In multi-path routing, survivability is an evaluation of fault tolerance of the multi-path solution if link failure occurs. This metric is usually employed to design routing algorithms to provide guaranteed transmissions.

Other Metrics

Some other metrics are also proposed in network scenarios with special interest. Transmission energy is an accepted metric in energy-efficiency networks and battery capacity is considered in wireless sensor networks [41]. Reliability is emphasised in vehicular ad hoc networks [42]. Link quality is presented to contribute to QoS management. Security is especially crucial in military networks.

2.2.2 Routing in Wireless Backhaul Networks

In backhaul networks, the traffic mainly goes between base stations and gateways . Within networks where caching techniques are enabled, some traffic flows may go from nodes with caching capacity to other base stations to provide caching contents, e.g. the red route shown in Fig. 2.2. With wireless connectivity, the base stations and relays can build up a mesh topology. Multiple paths may exist from one particular base station to the gateway. For example, as shown in Fig. 2.2, three possible paths between the node at the top and the

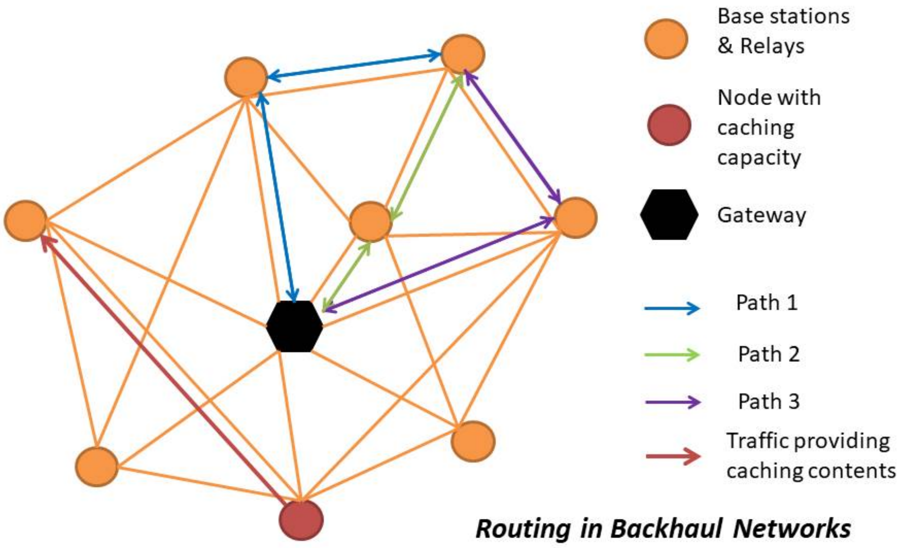

Fig. 2.2 Routing in backhaul networks

gateway are marked. Thus, routing problem needs investigation in this network topology.

The backhaul network takes the responsibility to connect the access network to the core network. As the performance of the backhaul links limits the access network functionality, many metrics related to service requirements have to be considered in the routing procedure in the backhaul network. For example, bandwidth should be considered to guarantee the service throughput. The delay metric should be employed to meet the latency requirements of specific services. Moreover, the routing in the backhaul network should attach primary importance to improve the capacity of the backhaul network to satisfy the huge amount of traffics. Although interference management is more related to resource allocation, the routing should also consider the interference states between links, which may potentially impact the network capacity.

2.2.3 Related Works

A few research studies have been conducted focusing on the routing problem in the backhaul networks.

Focusing on routing and scheduling in time division multiple access (TDMA) based wireless mesh backhaul networks, the authors in [43] proposed a linear programming optimization to provide QoS guarantees to real-time services with unrestricted topologies. A model for performance evaluation was also developed. The numerical results showed that the proposed mechanism outperforms the least-hop routing and conventional TDMA scheduling. Bandwidth metric is also considered in this work.

In paper [44], an adaptive cross-layer routing scheme was proposed to select the most reliable path in IEEE 802.16 WiMAX based backhaul networks. Evolutionary game theory was employed to model the routing procedure. The speed of convergence of the proposed scheme was also investigated. In this work, the packet error ratio was chosen as the main metric in routing.

Routing and scheduling were investigated in [45] in wireless backhaul networks with smart antennas. Interference aware tree construction problem was defined by the authors. A routing algorithm was proposed to solve it in polynomial time with the aid of adaptive antennas that could suppress the interference with the degrees of freedom. Scheduling was also presented with both optimal algorithm and effective heuristic algorithm. Efficiencies of the algorithms were justified by simulation results. Interference metric is considered in this work.

Wireless self-backhaul was assumed in [46], where the access network and the backhaul network share the same spectral resources. To maximise the throughput of routes, mutual interference among wireless links was taken into account in the centralised joint routing and resource allocation algorithm in an ultra-dense indoor network scenario. Multi-hop self-backhaul was shown to be a viable solution with limitations on system throughput through simulations. A joint design approach was adopted in this work considering interference management.

The authors in [47] focused on employing mmW backhaul solution based on LTE architecture. A method of dual connectivity establishment for heterogeneous wireless backhaul network was proposed, which enabled LFM and mmW backhaul in the same infrastructure. Additionally, a self-organised mmW multi-hop backhaul establishment procedure was presented to take advantage of the bidirectional beam alignment. Dynamic routing was enabled by the flexible backhaul reconfiguration. The dual connectivity is innovative to embed mmW and dynamic routing into current broadband systems.

Time-varying routing, of which the concept is that routing changes in certain time interval to adapt to dynamic traffic loads, was investigated in mobile backhaul topologies in [48]. The benefit of time-varying routing was proven through comparison with static routing. A new metric, capacity variation, was proposed to minimise capacity waste. The approach to adapt to dynamic traffic loads can be used for reference in future design.

Based on Cloud-RAN central unit, the problem of joint routing and backhaul link scheduling in mmW multi-hop backhaul network was studied in paper [49]. Considering various channel conditions and QoS requirements, the problem was formulated as a generalised vehicle routing problem, which is comprised of channel-aware path selection and queue-aware link scheduling. Solutions were given to achieve higher spectral efficiency and utilization. Interference and Delay metric are considered in this work.

Considering distributed caching in mobile networks, the authors proposed a novel in-RAN caches orchestration strategy was proposed in the work of [50]. Minimising content transmission delay was set as the objective in the joint wireless and backhaul routing algorithm. Numerical results showed that the strategy and algorithm can significantly reduce the delay and backhaul load. Delay metric is employed in this work.

An optimisation model to minimise the total power consumption in 5G HetNets with mmW backhaul was presented by [51]. User association, multi-hop backhaul routing and small cell sleeping strategy were jointly considered in the mixed integer linear problem. Simulations were performed to gain insights of the system model. Energy-based metric is utilised in this work.

To conclude, few works have been performed to improve the backhaul capacity while satisfying various traffic demands. Further research should be conducted in this direction. In addition, the interference management is important in routing procedure for wireless backhaul networks.

2.3 Reviews of Resource Allocation

Resource allocation is a fundamental problem for wireless communications. The resource in wireless domain mainly refers to spectral resource. Due to spectrum scarcity, how to optimally allocate the available spectral resources to various communication links remains a challenge.

2.3.1 Basic Concepts

The resource allocation problem mainly considers the following aspects: 1) Spectral resource model; 2) Model of resource request for traffics; 3) Mechanism of assigning and updating the resources.

The spectral resource model can vary, with diverse technologies applied. For example, with TDMA technique, the spectral resources are sliced based on time slots, while, in orthogonal frequency division multiple access (OFDMA) systems, the spectral resources can be divided into resource blocks according to time slot and frequency sub-carriers.

The resource request model of traffics can be mainly categorised into two classes: resource-based models and throughput-based models. The resource requests in resource-based models are straightforward, consisting of a certain amount of resource units (e.g. time slots in TDMA and resource blocks in OFDMA). As the resource units can be directly put into the allocation procedure, this kind of models has advantages in allocation implementation. However, the disadvantage of such requests is also obvious, in that it is hard for service providers to guarantee the data rate by requesting resource units due to various channel conditions. In the throughput-based models, the requests are given in the measure of minimum data rates. This kind of models make the allocation procedure more strict and complex because the request has to be transformed to resource units based on current channel condition and dynamic control has to be performed due to time-varying data rate per resource unit.

To accomplish the requests, a mechanism to assign the resources to various transmission links in the network should be designed based on various metrics. Several metrics are mainly considered in wireless communications: energy efficiency, throughput maximisation, fairness, robustness and QoS. The resource allocation mechanisms can vary in order to fit for the requirements in various network scenarios.

2.3.2 Resource Allocation in OFDMA-based Backhaul Networks

With the advantages of scalability, robustness to multipath, downlink multiplexing, uplink multiple access and multi-input multi-output (MIMO) benefits [52], OFDMA has been accepted as the multiple access technology for LTE, and also as an outstanding candidate for future networks. Thus, OFDMA-based backhaul network is mainly considered in this thesis.

The principle of orthogonal frequency division multiplexing (OFDM) is to divide the data stream into several sub-streams that can be transmitted over different orthogonal sub-carriers. The rate and power on each sub-carrier can be individually adjusted so that OFDM can resist frequency selective fading. Moreover, any combination of sub-carriers can be accepted to transmit the data, which makes OFDM flexible for resource allocation. OFDMA extends OFDM to multiple transmission scenarios. For example, in the access networks, multiple UEs can build transmission links with the base station through OFDMA. Different UEs are assigned with different orthogonal sub-carriers to avoid interference.

By adopting OFDMA, the spectral resources can be divided into resource blocks (RBs) in time-frequency plane . RB is the smallest unit of resources that can be assigned by the network. Typically, RB is 180 kHz wide in frequency and 1 slot long (0.5 ms) in time. The resource request model can be resource-based or throughput-based. As the topology of backhaul network is stable, the channel conditions are relatively steady compared with the

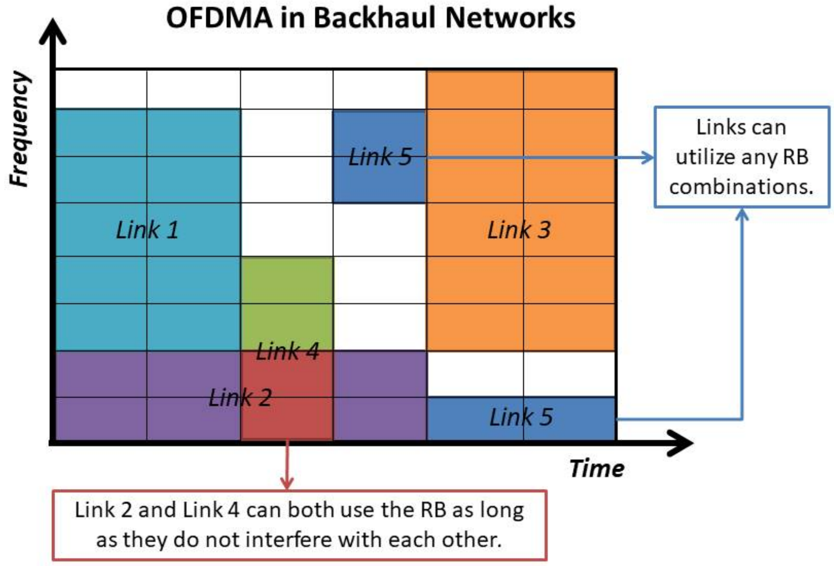

Fig. 2.3 OFDMA in backhaul networks

access networks. Thus, both models are simpler to implement.

In multi-hop backhaul networks, the nodes may build multiple transmission links with other nodes. Thus, links are the research objects that need to be assigned with resource blocks. Throughput maximisation and QoS are the main metrics to be considered in backhaul networks. With the assumption that omnidirectional antennas are equipped to the nodes, there are two basic rules that can apply to all resource allocation mechanisms from the interference management point of view. Different outward links from the same node cannot occupy the same resource blocks. Different inward links to the same node cannot be assigned with the same resource blocks.

As shown in Fig. 2.3, any combination of RBs can be utilized by any link. For example, Link 5 are assigned with 4 RBs that are not adjacent. Also, the OFDMA enables the reuse of the same RB. For example, as long as Link 2 and 4 do not interfere with each other, they can utilize the same RBs.

2.3.3 Related Works

Many researches in the resource allocation area in backhaul networks have been conducted.

In [53], the authors proposed a cognition based resource allocation strategy for multi-hop backhaul networks with directional antennas. A weighting factor based reinforcement learning scheme was investigated to apply to interference weighted channel selection policy. As a result, the spectrum efficiency and QoS were improved. In this novel work, cognition and reinforcement learning are introduced to resource management in backhaul networks.

Resource allocation in multi-channel multi-radio wireless mesh backhaul network was investigated in [54], where physical-layer interference was concerned. A two-stage radio allocation scheme based on a game-theoretic approach was proposed to increase operative links in the network. In the proposed game, a utility function was designed to minimise co-channel interference among players. The game-theoretic approach is proven to be effective in resource management.

A mmW-based unified access and backhaul network was investigated in [55]. As the same mmW bands were shared between the radio access and backhaul networks, resource allocation problem was studied to adjust the allocated resource ratio between them. Three dynamic algorithms were developed in order to maximise the network capacity and the fairness among base stations. Network capacity is analysed from both access and backhaul perspectives in this work.

Self-backhauling and caching were introduced into resource allocation problem in ultra-dense networks in the work of [56]. To solve the resource allocation problem in the proposed network architecture, a tri-stage fairness algorithm was developed considering fairness, efficiency, overhead and complexity. A traffic model taking caching impact into account was also presented. Numerical results indicated that flexible resource allocation between access and backhaul link enhances performance gain. Caching technique is employed in the system model in this work.

A joint optimisation of resource allocation in wireless backhaul links and user-centric clustering in the access links was performed in [57]. To solve the problem iteratively, proper approximation and transformation were employed. The proposed scheme aimed to assign users to BSs considering the backhaul link constraint. Simulation results verified the performance of the proposed algorithm in various network settings. In this work, the novelty is that the serving BSs are dynamically assigned to users combining the backhaul network states.

E-band was employed in wireless mesh network for 5G backbone in the research conducted in [58]. Simultaneous transmission and reception with divergent beams was described for the system. A graph approach was proposed to distribute the time-frequency resources in the backbone links. The efficiency of the heuristic distribution algorithm was verified through simulations. Antenna technique is introduced to aid the resource management in E-band mesh networks.

The authors in [59] focused on the problem of backhaul resource allocation in converged wireless-optical network architecture. The scenario where two base stations belonging to different operators compete for resources in the shared optical backhaul network was investigated. An evolutionary game theory based approach was designed to assign backhaul resources dynamically to base stations considering QoS and backhaul limitation. The results indicated that the system design was stable even with various time delays. The game theory that base stations competing for backhaul resources can be referenced in further backhaul resource allocation studies.

HetNets with macro-cell base station providing wireless backhaul for small cell access points were investigated in paper [60]. Massive MIMO was assumed in the macro-cell base station and the same spectrum resources were shared in the backhaul and access networks. Deterministic formulations for ergodic uplink and downlink sum rates were developed depending on statistical channel information. Based on the expressions, a resource allocation method was proposed to improve the system performance in terms of sum rates. The sum rates expressions are presented in the proposed system architecture, which can be further transformed to apply to other network settings.

To conclude, resource allocation problem in wireless backhaul networks still need further investigation in the scenario where backhaul network operates in exclusive frequency bands. Interference management and performance enhancement should be achieved to satisfy various service requirements.

2.4 Reviews of QoS Provisioning

The definition of quality of service may vary based on different applications. Generally speaking, QoS provisioning can be defined as satisfying the stated and implied needs of the users of the service [61], which can be achieved through cross-layer approaches. In the following, basic challenges and approaches of QoS provisioning in wireless backhaul networks will be presented and related works will be introduced with discussions.

2.4.1 Main Challenges

Implementing QoS mechanisms to wireless backhaul networks not only faces traditional challenges like interference management but also confronts unique challenges due to the speciality of the network characteristics. In this section, some of these challenges are highlighted and discussed:

Limited Resources

Until now, the mmW-based backhaul network has not been commercially available. By sharing the same spectrum with the access networks, the wireless backhaul network suffers from limited spectral resources. Even if mmW bands are utilised in the future networks, the data demand is also experiencing explosive growth. There is no guarantee that the spectral resources can completely satisfy the future traffic demand.

Dynamic Traffic

Not only the randomness of different users accessing the network from various locations but also the random time durations of various services contribute to the dynamic behaviour of the traffics in the backhaul network. Furthermore as pointed out by [62], user plane traffic from the access networks can be summarised as busy time traffic and quiet time traffic. During busy times (e.g. noon and evenings), many UEs are distributed across the whole network generating large amount of traffics. During quiet time (e.g. midnight), only a few UEs are accessing the network, which means that few traffics are required in this scenario. However, users may demand higher QoS with services like video streaming and mobile gaming. Thus, the backhaul network has to satisfy the dynamic traffics while providing diversified QoS requirements.

Heterogeneous Traffic

The diversification of applications in mobile phones is pushing the service differentiation. Various services have different requirements on network performances. Especially in 5G networks, all kinds of functions need to be supported in the same architecture. For example, mission-critical communications have a steep demand on reliability and resilience. Autopilot systems cannot tolerate delays in vehicular communications. To satisfy the heterogeneous traffics, the backhaul networks have to support differentiated QoS.

2.4.2 Basic Approaches

QoS provisioning is a complex cross-layer function, which requires several QoS management components to collaborate. The QoS management components include routing, spectral resource allocation, transmission power control among others. QoS metrics, such as throughput, delay, packet loss rate, packet error rate and reliability should be optimised according to service requirements.

Generally speaking, there are two types of approaches for network QoS management:

global offline QoS networking and greedy online QoS networking [63].

Global offline QoS networking jointly considers all QoS management components to find a global optimal QoS solution for all current traffic flows in the network. This approach typically employs optimisation tools, such as integer or mixed integer programs, within which computational complexity may be very high.

Greedy online QoS networking aims to find a local optimal QoS solution for incoming traffic flows while maintaining the QoS of current traffic flows in the network. In this approach, the estimated QoS metric values of different components are required prior to decision making. This approach is more suitable for dynamic networks.

In wireless backhaul networks, the transmission links are stable due to the stationary topology of the backhaul nodes. Thus, transmission parameters in the nodes are configured with appropriate values, which mean that the physical layer of QoS is not a concern. Therefore, delay and throughput are two main metrics to consider in wireless backhaul QoS provisioning.

2.4.3 Related Works

Few works have been conducted that focus on QoS provisioning for backhaul networks, although different QoS metrics have been considered in the related works.

The work performed by [64] addressed both throughput and delay requirements in multi-hop wireless backhaul networks. Several admission control schemes were developed to guarantee the requirements of admitted connections. In the proposed algorithm, a tree-based topology was constructed and then the subset of connections that satisfy the QoS metric was selected. The topology construction and the subset selection are novel approaches to deal with admission control in backhaul networks.

The QoS problem was investigated through routing and scheduling in TDMA-based wireless mesh backhaul networks in study [43]. By integrating the routing and scheduling procedure, the authors proposed an algorithm to provide throughput and delay guarantees for real-time services. Linear programming was employed to assign non-collision bandwidth for the selected path. Minimal-hop routing was outperformed by the proposed scheme. Throughput, delay and interference management are jointly considered in this work.

By studying the scheduling of traffic flows and channel assignment, the authors in [65] proposed a recursive fair stochastic matrix decomposition algorithm and a constrained graph colouring algorithm to support QoS on a per-flow basis in wireless mesh backhaul networks. Beamforming was also employed to guarantee the transmission rate and minimise the signal to noise ratio (SINR) . Near-minimal delay and jitter were achieved while satisfying the throughput requirements with the proposed algorithm. However, the network model in this work is simplistic.

A wireless backhaul business case was studied in [66]. To achieve low cost and low power, a novel carrier-grade heterogeneous multi-radio backhaul architecture was developed. Various functions were described including topology, capacity, pipe management, terminal control and QoS-aware LANE emulation functions. Measurements were conducted to verity the proposed system. Fairness and predictability are considered as the QoS metric in this work.

To confront the unfairness problem in chain-based wireless backhaul networks, a two-level QoS scheduler was proposed by adopting the Ripple protocol. A queueing model was developed to analyse the mean delay, throughput in the network. Multiple-class traffics were supported by the proposed QoS scheduler. Mean delay, end-to-end throughput and aggregated throughput are guaranteed to provide QoS for various traffic classes.

Combining SDN and existing LTE architecture, three OpenFlow-based mechanisms were proposed to provide the required QoS in research conducted in [68]. Traffic routing, admission control and traffic coexistence mechanism were separately designed to improve the QoS in the infrastructure by controlling the bandwidth usage. Numerical results validated the performance improvement in terms of throughput, delay, jitter and packet loss. This work introduced the SDN protocol into QoS provisioning for future researches.

To reduce the end-to-end delay of high priority packets in mobile backhaul networks, packet priority scheduling was investigated in [69]. The proposed algorithm utilised dynamic packet reordering and scheduling to support QoS. The concept of priority and queue reordering can be referenced in further researches.

To conclude, throughput and delay are two main QoS metrics to consider in wireless backhaul networks. However, works that focus on satisfying diverse QoS requirements for various services are rarely seen. Especially when the network has insufficient spectral resources to satisfy all traffic demands, the QoS provisioning remains a challenge. Further researches are required to improve the QoS management in wireless backhaul networks.

2.5 Reviews of E-band

As a potential candidate for future backhaul networks, E-band frequency has unique characteristics that can be utilised for system designs. In this section, the features of E-band are reviewed and related works are also presented.

2.5.1 Characteristics of E-band

Compared to lower frequency bands, E-band signals can only transmit in shorter range and cannot penetrate physical obstructions very well. These characteristics should not be treated as disadvantages. On the contrary, benefits can be obtained with proper network designs. The characteristics of E-band are discussed in the following sub-sections.

Propagation Features of E-band

With the assumption that the transmission link is line-of-sight (LOS) and the antennas are omnidirectional, the path gain of the link with free space propagation can be formulated as:

$$G = G_T G_R \frac{\lambda^2}{(4\pi D)^2} \tag{2.1}$$

Where G_T and G_R are the gains of transmitter and receiver antennas respectively, λ is the signal wavelength and D is the link distance [70]. It can be concluded from Eq. (2.1) that with given G_T, G_R and D, the path gain is decided by λ. As the signal wavelength in E-band is much smaller than that of low frequency microwave bands, E-band transmission is experiencing less gain. Thus, omnidirectional antennas are not fit for E-band transmissions.

Directional antenna, which is made up of a number of antennas to form a “pencil beam”, is one approach to counter the addressed problem. In E-band, it is much easier to apply directional antennas because the space between antennas to form an equivalent directional beam is scaled down with the wavelength size.

Free space propagation only exists with LOS guaranteed and no obstacles around. However, the application scenario of E-band links is mainly located in urban area, where buildings and trees may block or reflect the signals. Penetration, absorption, diffraction, reflection and diffusion are phenomena that would happen in such cases. The wavelength of E-band signal is the main factor that influences all these phenomena. The penetration ability of E-band signal is extremely poor [71]. Due to the millimetre level wavelength, E-band signal is difficult to diffract when encountering obstacles. Speaking of multipath, besides the absorption part by the materials the reflection part is much smaller compared to low frequency microwave because greater diffusion would arise for E-band signal with small wavelength due to the relatively “rougher” surface of the materials.

With all the characteristics discussed above, we can summarise two main NLOS E-band transmission features compared to low frequency microwave transmission: 1) The number of multipath components is smaller; 2) The signal power of multipath component is much lower than that of the LOS component. With the sparse propagation channel of E-band, channel

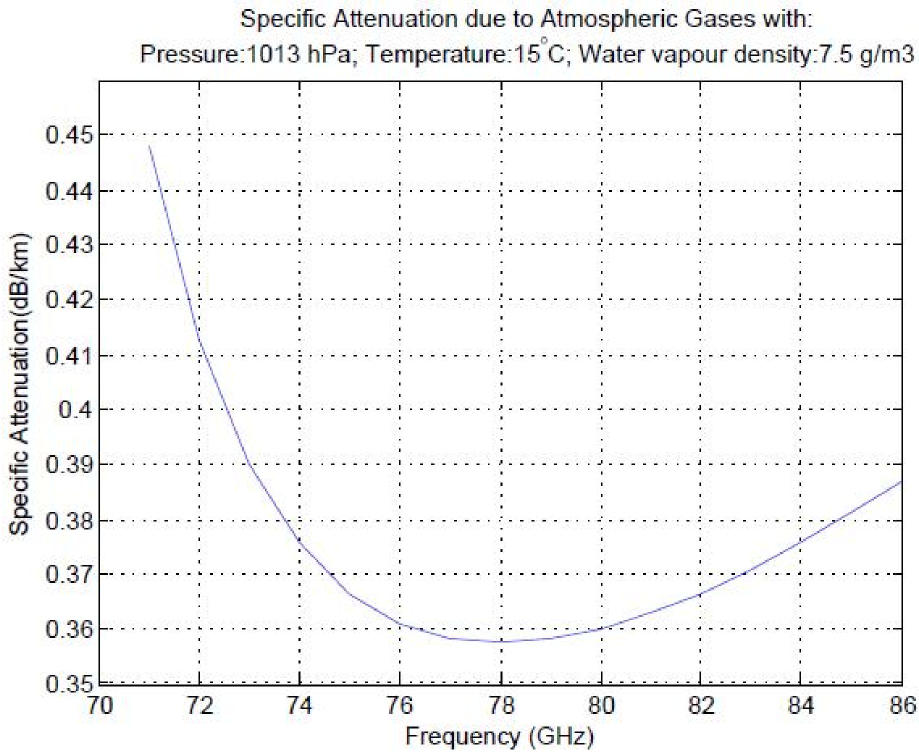

Fig. 2.4 Specific attenuation due to atmospheric gases [73]

estimation and transceiver design is much easier [72]. For example, beamforming can be used to improve the path gain significantly if the channel information is provided.

Attenuation Features of E-band

Besides the power loss during free space and reflection propagations, attenuations also impact the transmission of E-band signals.

• Atmospheric Attenuation

Absorption by water vapour, molecules of oxygen and other gaseous atmospheric constituents may take place when signal is travelling through atmosphere. Fig. 2.4 shows the specific attenuation due to atmospheric gases for a pressure of 1013 hPa, temperature of 15 $^{\circ}C$ for the cases of a water-vapour density of 7.5 g/m^3, which is derived from ITU-R Rec. 676-10 [73]. For E-band, atmospheric attenuation is less than 0.45 dB/km. It is relatively much lower than other mmW bands, which can be 5 dB/km. This characteristic makes E-band a favourable choice for longer transmissions.

• Rain Attenuation

As the wavelength of E-band is comparable to the size of the rain drop, rain can cause large attenuation. However, rain drop size is affected by rain rate. Thus, there is a relationship

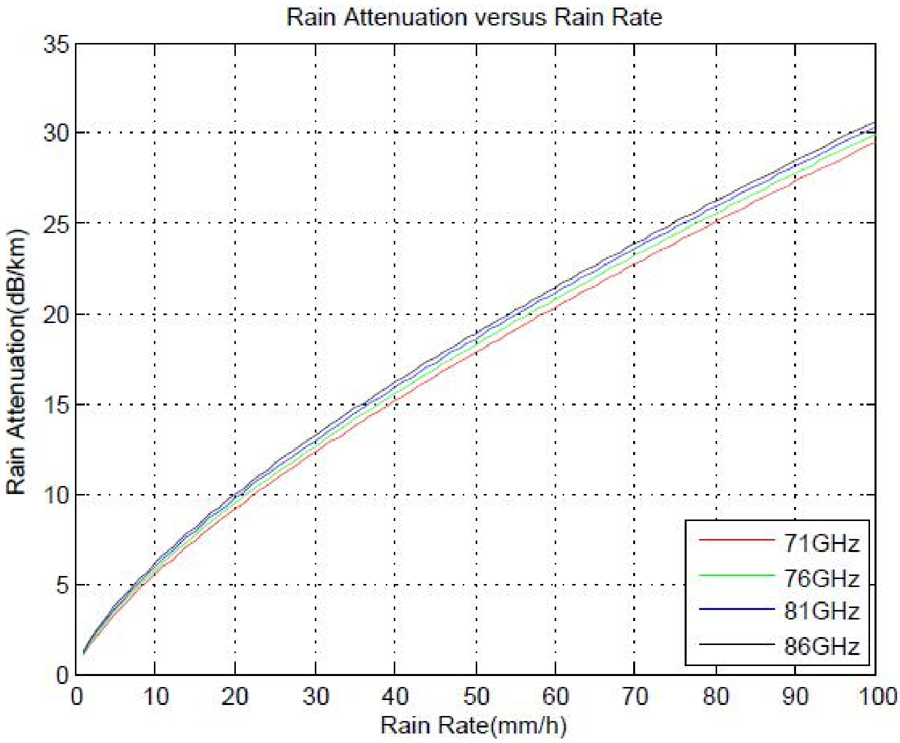

Fig. 2.5 Rain attenuation versus rain rate [74]

between rain attenuation and rain rate. Fig. 2.5 shows the specific attenuation due to rain derived from ITU-R Rec. 838-3 [74], in which the following equation is used:

$$\gamma_R = kR^{\alpha} \tag{2.2}$$

where γ_R stands for the specific attenuation (dB/km), R is the rain rate (mm/h) and values for the coefficients k and α are determined as functions of frequency, $f(GHz)$. As can be seen from the curves, the rain attenuation can be more than 30 dB/km, which is much larger than the atmospheric attenuation (up to 0.45 dB/km).

• Foliage Attenuation

While lower frequencies can penetrate more easily through solid objects, mmW bands do not penetrate very well [71]. Thus, foliage attenuation can be much more serious in E-band case, which may be a limiting factor when deploying an E-band node. As we aim to discuss the possibility of the backhaul application of E-band, we can assume that the transmission link is LOS, where no foliage exists when the backhaul links are carefully installed.

• Other Attenuations

Additionally, fog, haze and snow are also weathers that can affect the propagation of E-band. However, according to [75], fog attenuation is less than 0.4 dB/km. With the fact that fog seldom appears, the attenuation due to fog can be ignored when considering E-band backhaul link. Similar to fog, haze cause ignorable attenuation due to the small size of these particles. Snow attenuation is also negligible, even if the fall rate exceeds 125 mm/h [72].

2.5.2 Related Works

Some introductory works about mmW and E-band have been presented in early years:

As analysed in work [76], mmW communication was viewed as a potentially disruptive technology for 5G. As a key feature, adaptive antenna arrays may require new access protocols. Hardware constraints may also bring new research challenges. The conclusion was drawn that mmW requires complete changes in both component and architecture designs of the system.

An early introductory work on mmW in [77] analysed the feasibilities of different mmW frequencies and the propagation characteristics of mmW. It pointed out that interference would not be as serious as in traditional cellular networks due to the narrow beam width of mmW transmissions. A hybrid mmW mobile broadband and 4G system was put forward by the authors and relative link budget analysis showed that a 2 Gb/s data rate can be achieved over 1 km distance.

In the work aiming at multi-gigabit per second wireless communication [78], three technologies, which were 60 GHz wireless, 70-80 GHz wireless and Free Space Optics , were compared. E-band wireless can achieve 1 Gb/s connectivity with 99.999% availability over 2-3 km and with 99.9% availability over 5 km. It was also indicated that different technologies should be commercially applied depending on the needs of end users, the requirements of link performance and budget.

A review on V-band, E-band and optical fibre backhaul technologies was firstly given in [79]. Then the authors presented a successful high-capacity hybrid mmW and optical backhaul by simulation. A backhaul resource manager, of which a key feature was capacity-aware path computation, was proposed to enable dynamic usage of the large capacity in this backhaul network. This work contributed a solution to take full advantage of the high capacity of the mmW backhaul network.

The work of [80] mainly discussed the potential of E-band communications in the coming years. By introducing the background and propagation characteristics of E-band transmission, the authors suggested that providing coverage might be a key challenge due to the directional beamforming. Several techniques that can potentially solve the problem were discussed such as adaptive beamforming, sparse channel estimation and user cooperation. This work provides some effective solutions on the directional beamforming point of view to solve E-band coverage problem in access network.

In the work of [81], the authors showed the applicability of the E-band wireless technology for wideband connectivity using theoretical analysis and computation based on real urban network parameters. This work thoroughly discussed the relationship between link length and rain intensity as well as the relationship between link length and capacity when using different modulation schemes under different availability levels. This work confirms that the potential of E-band communication in urban areas needs to be subjected to detailed analysis.

Several research efforts have been made in relation to mmW area based on measurements:

Measurements in [82] study showed good agreement with the ITU-R models, which is a strong support for the application of ITU-R models in simulation works. Large amount of measurements at 28 GHz and 38 GHz of angel of arrival , angel of departure , root-mean-squire delay spread, path loss, and building penetration and reflection characteristics were presented in the work of [83]. Indoor measurements of 72 GHz mmW band were operated in [84] study. The results of E-band wideband propagation measurements around 73 GHz in [85] revealed that directional antenna is a key feature that enables future mmW communications. In [86] work, measurement results of channel attenuation and phase noise on E-band systems were compared against existing models and discussions about bandwidth efficiency of E-band systems were also presented. By operating some propagation measurements in New York, [87] indicated that mmW signals are viable even in none-line-of-sight (NLOS) scenarios over 100-200m and can provide significant capacity boost with the help of beamforming.

Some works that applied mmW to backhaul networks has also been performed on a trial basis:

Authors in [88] conducted a research that focused on throughput and energy consumption in 5G wireless backhaul networks with mmW. Under the premise of ultra-dense small cells, centralised and distributed backhaul scenarios were configured to analyse the traffic and energy efficiency in future 5G networks. The authors indicated that the distribution solution was more energy-efficient than the central one. This work contributed to the energy efficiency analysis of backhaul network.

Through analysing different backhaul technologies, work [89] indicated that the mix of fibre and mmW is the most potential choice to enable cloud-based architecture. The demand for rate and latency would increase with the degree of centralisation. Thus, a flexible functional split and a joint optimisation of access and backhaul were required to satisfy the centralised processing.

Research [90] presented a novel control and user plane split to enable mmW for backhaul and access. This concept work can provide high data rate densities and support for other technologies like C-RAN.

The work by [91] considered both point-to-multipoint (P2MP) and point-to-point (P2P) scenarios proposed a sub-6 *GHZ* and E-band links aggregation multi-hop backhaul architecture. The key idea was that through P2MP base stations were connected to aggregation nodes that were connected to gateways through P2P E-band links to form a multi-hop network. Low latency design techniques, such as adaptive relaying and dynamic routing and scheduling, were also analysed to enable the novel architecture.

To conclude, plenty of literature has pointed out E-band is a potential candidate for future backhaul network. The capacity performance has been confirmed for transmission links operated in E-band. However, there are few studies dealing with the rain attenuation problem in E-band. Further work to tackle this challenge is therefore necessary.

2.6 Summary

This chapter consists of five major overview sections: virtual network embedding problem, routing strategy, resource allocation scheme, QoS provisioning and E-band frequency band.

In the overview of virtual network embedding, basic concepts of network virtualization, virtual link mapping and virtual link mapping are introduced. The *NP*-hard nature of the VNE problem is explained by referring to related problems. Through discussing the related works, the conclusion is drawn that online VNE in wireless backhaul network with dynamic approaches needs further investigation. Caching technique should also be taken into consideration.

In the reviews of routing, several routing metrics are introduced. Specialised on routing in wireless backhaul networks, the traffic flows are analysed and design challenges are highlighted. Few related works have been conducted to improve the wireless backhaul network throughput by loading more traffic requests. Thus, further research is needed to explore the method of enhancing network capacity through routing.

In the reviews of resource allocation, basic concepts of resource model, resource request model and resource allocation mechanism have been introduced. More detail about the models and research challenges are given based on OFDMA-operated wireless backhaul networks. Most related works considered the scenario where backhaul networks and access networks share the spectrum or regard backhaul as a constraint. Resource allocation schemes should be studied in wireless backhaul networks to improve the backhaul performance.

In the reviews of QoS provisioning, the main challenges to implement QoS support in backhaul network are highlighted. Basic approaches are introduced to solve the QoS provisioning problems. Greedy online QoS provisioning is preferable for the traffic-dynamic backhaul networks. The reviewed works mainly employ routing and resource allocation schemes to guarantee QoS. However, no work has considered QoS support in resource-limited scenarios. As delay and throughput are two main factors that influence the QoS, further research to reduce delay and guarantee throughput in wireless backhaul works in resource-limited scenarios should be conducted.

Finally, the E-band frequency features and related works are presented. Rain attenuation is the main challenge in E-band backhaul designs. Further work that tackles this challenge is needed.

Chapter 3

Virtual Network Embedding in Software-defined Backhaul Networks

Overview

The VNE problem constitutes the main resource allocation challenge in network virtualisation, which is considered to be one of the most promising technologies for overcoming the network ossification problem. A number of algorithms have been proposed to efficiently exploit physical resources to facilitate virtual networks considering the NP-hard nature of VNE problem. However, existing algorithms only consider links with fixed bandwidths. Due to various applications of 5G networks, the VNE problem in OFDMA-based SDBNs is investigated in this paper. Also, the caching technique is taken into account due to its benefits of backhaul traffic relief. Virtual node mapping and virtual link mapping are separately modelled based on the features of SDBNs. In the proposed VNE-SDBN algorithm, the serving nodes and gateway nodes of the backhaul network are mapped based on the location information of the VNR, while the caching nodes are located to maximise the caching gain for the backhaul network. For the virtual link mapping problem, the interference properties and the available resources of links are taken into account. Simulations have shown that our algorithm outperforms two state of the art VNE algorithms in terms of increasing revenues and the VNR acceptance ratio in wireless backhaul networks.

3.1 Introduction

Over the past few years, not only has data demand dramatically increased across nearly all networks, but also service requirement diversity is resulting in challenges to network reliability and flexibility. To confront these problems, researchers have proposed the network virtualisation technique – specifically to overcome network ossification [16–18]. By decoupling the ISP into the InP, which deploys and maintains the network equipment, and the SP, which offers end-to-end services, ISPs can provide customised services according to user demand by creating multiple virtual networks. A main challenge of network virtualisation is that of embedding virtual networks into a substrate network; this is usually referred to as the VNE [23]. The benefits of VNE can be maximised by improving the resource utilisation of physical networks with efficient and effective VNE algorithms.

As the VNE has to map virtual nodes and links simultaneously, the problem is considered as NP-hard in nature [23]. Previous studies have taken advantage of both exact and heuristic algorithms to solve the VNE problem. In [25], the authors presented a comprehensive metric of node importance which measures the embedding potential of substrate nodes and they proposed a VNE algorithm based on a breadth-first-search algorithm. An exact VNE algorithm to solve the online VNE problem was proposed in [26] – using Integer Linear Programming methodology. The authors of [27] focused on embedding virtual network requests on a sharing basis and they put forward a two-step VNE algorithm with a resource block splitting and sharing solution. Concentrating on the VNE problem in mobile networks, the authors of [28] proposed an algorithm that minimised the number of virtual network migrations, so countering the mobility of wireless nodes. In a recent research study [29], the impact of mobility on virtual network embedding was examined and a mobility aware VNE algorithm was proposed to increase network performance.

However, most of the existing research studies only consider the VNE problem in networks with fixed bandwidths for each link. In SDNs [5–7], the bandwidths of links are considered to adapt to the dynamic traffic flows in order to improve network performance. More importantly, the broadcast nature of wireless networks brings up the issue of interference between links, and this makes the situation regarding link bandwidth even more complicated. Traditional VNE algorithms are not fit for this purpose. Hence, it is necessary to investigate the VNE problem in scenarios where the bandwidth of links is dynamic.

Besides, with the development of information centric networks researches [92, 93] in recent years, it has shown that backhaul networks can benefit from in-network caching, which can alleviate the backhaul pressure and reduce user access latencies. With caching technique employed in backhaul networks, a new problem for VNE arises, which is how to assign the cache capabilities among numerous nodes to improve the network performance from the backhaul network point of view.

In this chapter, the VNE problem is investigated in an OFDMA-based SDBN, where multi-hop transmission and mesh topology are applied. This work is distinguished from previous studies because it introduces the following major contributions.

1. We model the VNE problem in the SDBN. On the basis of backhaul network features, we divide the virtual node mapping problem into serving nodes, gateway nodes and caching nodes mapping.

2. Caching techniques are considered in relation to VNE problems.

3. For the virtual-link-mapping problem, two metrics are proposed to evaluate the substrate link according to its interference properties and available resources.

4. The VNE-SDBN algorithm is proposed to solve the problem which is thus identified.

The remainder of this chapter is organised as follows. In 3.2, we describe the SDBN in detail and introduce the basic VNE problem. Then, we analyse the VNE problem in relation to the characteristics of SDBNs in 3.3. Next, we propose our VNE-SDBN algorithm in 3.4. The simulation results and analyses of these are presented in 3.5. Finally, we conclude this paper in 3.6.

3.2 System Model and Basic Problem Statement

In this section, we first introduce the basic concept of SDBNs. Based on the SDBN architecture, we then put forward the substrate network model and the virtual network request model, which are different from traditional VNE models. At last, we formulate the VNE problem for SDBNs.

3.2.1 Software-defined Backhaul Network

A backhaul network is considered in this work, in which a basic SDN architecture is adopted as shown in Fig 3.1. The architecture consists of two parts: an SDN-controller in the control plane and switching devices in the data plane. The SDN-controller is a centralised software-defined control unit which can communicate with all switching devices through the control panel. A switching device, denoted as a node in the following for simplicity, is responsible for collecting and sending network statuses to the controller and for processing data packets based on the rules provided by the controller. In a backhaul network, the nodes are basically BSs and relays, which, alternatively, can be divided into three categories: 1) any node that is directly connected to the Internet is referred to as gateway node; 2) BSs that provide services to UEs through the access network are denoted as serving nodes; and 3) any node with caching function is defined as caching node. These three categories do not conflict with each other, which means one node can be a gateway node, a serving node and a caching node

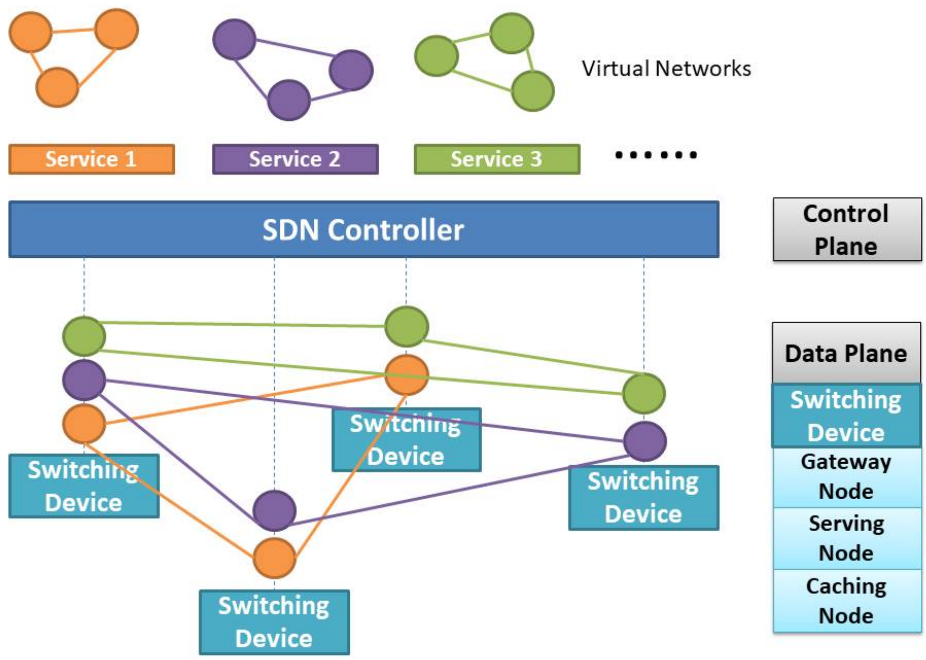

Fig. 3.1 Software-defined backhaul network architecture

at the same time. Multi-hop wireless transmission is employed to build communications between any two nodes.

Since the dense deployment of small cells becomes a trend for future networks [95], and OFDMA shows promising properties for managing interference and providing robustness to multipath [3]; OFDMA is adopted in this work. We assume that the spectral resource is divided into W sub-channels and one second is divided into T time slots. One sub-channel, w, in one time slot, t, is called a resource block $b(w,t)$. One BS can transmit and receive data in any combination of the resource blocks.

In contrast to access networks, a special characteristic of SDBNs is that the architecture is relatively fixed. Thus, it is reasonable to assume that each link is preconfigured with appropriate parameters: including transmission power, antenna beam-forming and other factors that impact interference between links. With regard to this assumption, the protocol model [96] is adopted in this work. Two conditions have to be met for a successful transmission across link l_i: 1) the receiving node must be within the transmission range of the transmitting node; 2) if the receiving node is within the interference range of other links, it must be ensured that the interfering links are not occupying the same frequency resource as link l_i.

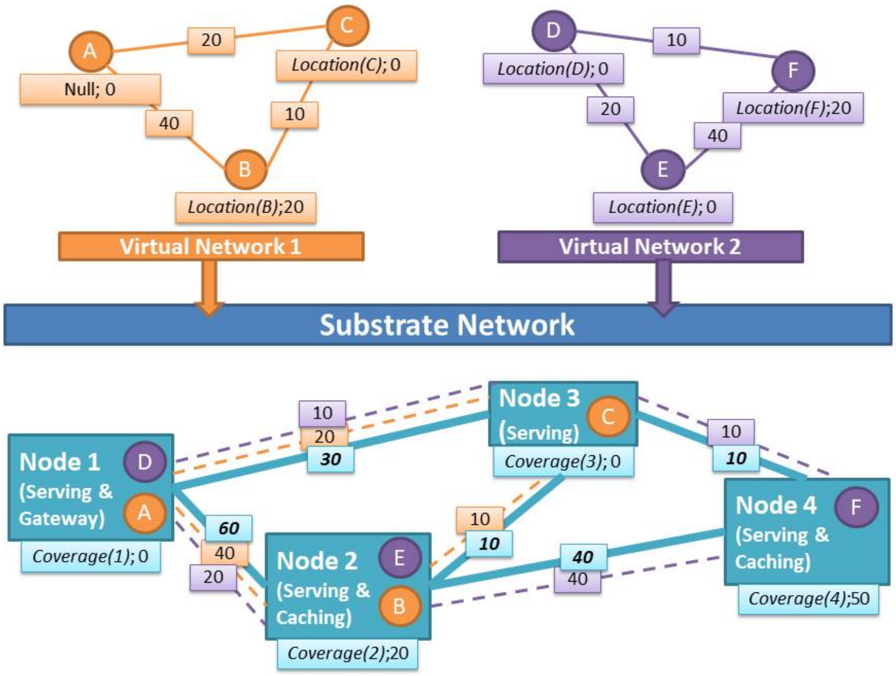

Fig. 3.2 An example of virtual network with embedded substrate network

3.2.2 Substrate Network Model

A substrate network can be abstracted as an undirected weighted graph $G_s = (N_s, L_s, A_s^N, A_s^L)$. N_s is a vertex set whose elements are nodes in the substrate network. The edge set L_s consists of undirected edges that represent established links in the substrate network. The link between node m and node n is denoted by l_{mn}. A_s^N and A_s^L stand for the attribute sets of the resources of substrate nodes and links, respectively.

In terms of the resources of substrate nodes A_s^N, location, the speed of switching, storage capacity and CPU capacity are examples of those which are often considered in studies. Departing from the norm established by some of the previous literature, we ignore the CPU capacity of nodes in this work because, as a result of the functional simplification resulting from the use of SBDN, the nodes are responsible only for switching data flows. Instead, we consider the demands of service providers in relation to backhaul networks. First, the coverage demand of the service is taken into consideration. For instance, the service subscribers should always be served through the access network. Second, each service provider has a caching capacity demand which depends on the specific service. For example, video streaming service providers have a greater demand in terms of caching capacity. To satisfy the demands of service providers, the resources of substrate nodes A_s^N include the coverage and the caching capacity of each node. The coverage of node m is denoted by set $Coverage(m)$, which consists of the location information of the area (e.g. a set of geographic coordinates) that can be served by the node. The set *Coverage* for any relay is null. Value $AvCache(m)$ is employed to quantify the available caching capacity of node m. The value *AvCache* of any node without caching function is zero. Thus, the resources of nodes are expressed as $A_s^N(Coverage, AvCache)$.

For the resources of substrate links A_s^L, in this work, only the available bandwidth is considered. However, the main challenge to overcome in terms of VNE for wireless networks results from the broadcast nature of wireless communication and the interference of one wireless link with another [97]. A resource pool matrix is introduced to help monitor and manage the available bandwidth of links. We propose that each link, l_i, has a resource pool matrix P_i. If $P_i(w,t) = 1$, this means that the specific resource block, $b(w,t)$, is already assigned to link l_i or its interfering links. Otherwise, resource block $b(w,t)$ is still available for l_i. Thus, the resources of substrate links are expressed as $A_s^L(P_i)$.

In the lower half of Fig. 3.2, four nodes form a substrate network. Nodes may play any combination of the three possible roles. For instance, node 1 acts as a serving node and a gateway node at the same time, while node 2 can serve the service subscribers and also cache popular content. The boxes below the nodes display the coverage information and the available caching capacity information. For instance, *Coverage(1)* stands for the coverage of node 1. The number indicates that node 1 has 0 available cache capacity. The numbers on the links shows the assigned RBs after the embedding procedure. For instance, the substrate link between node 1 and 3 are assigned with 30 RBs.

3.2.3 The Virtual Network Request Model

A VNR can be modelled as an undirected weighted graph $G_v = (N_v, L_v, A_v^N, A_v^L, T_a, T_d)$. Similar to the substrate network model, N_v and L_v each represents sets of virtual nodes and links. A_v^N and A_v^L stand for the resource constraint collections relating to virtual nodes and links, respectively. T_a indicates the arrival time of each VNR, and T_d denotes the duration of the VNR remaining in the substrate network.

For a request for virtual node resources, A_v^N, we consider the location recommendation of serving nodes and the cache capacity demand of caching nodes in response to the substrate network model. The location recommendation of virtual serving node m is denoted by $Location(m)$, which is a set of location data (e.g. geographic coordinates) for a specific area that should be covered. To satisfy the location recommendation request, set *Location* should always be part of $\sum_{n=1}^{N} Coverage(n)$, where N is the set of nodes that are assigned to

Symbol	**Definition**
N_s	Nodes in the substrate network
L_s	Links in the substrate network
A_s^N	Resources for substrate nodes
$Coverage(m)$	Coverage information for substrate node m
$AvCache(m)$	Available caching capacity for substrate node m
A_s^L	Resources for substrate links
P_i	Resource pool matrix for substrate link i
N_v	Nodes in the virtual network
L_v	Links in the virtual network
A_v^N	Resource constraints for virtual nodes
$Location(m)$	Location recommendation for virtual serving node m
$DeCache(m)$	Cache capacity demand for virtual caching node m
A_v^L	Resource constraints for virtual links
D_j	Resource block request for virtual link j

Table 3.1 Symbols for the virtual network embedding problem

the virtual network. The cache capacity demand of virtual caching node m is denoted by $DeCache(m)$; this should be no more than $\sum_{n=1}^{N} AvCache(n)$. In this work, we assume that each virtual network request holds one and only one cache capacity demand which is not bound to a particular virtual node. The virtual caching node only indicates the cache capacity demand of the virtual network.

For a request for link resources, A_v^L, we take the link capacity into consideration as real-time traffic is unknown at the point of VNR. Link capacity can be converted into resource blocks per second as explained for the substrate network model. The resource block request for virtual link l_j is denoted as D_j. The available resource blocks of each substrate link in the embedded path should be more than D_j.

For example, the upper half of Fig. 3.2 shows two virtual network requests. The data in the boxes below the virtual nodes represent the location recommendation and the required cache capacity, respectively. Taking node A for instance, the *Null* means that node A does not have location requirements and 0 means that the cache capacity is 0. The value in the box on the virtual link indicates the required RBs. For instance, virtual link *AC* requires 20 RBs.

In backhaul networks, the serving nodes are always linked back to gateway nodes or caching nodes. Thus, a tree structure is the typical topology of backhaul networks. Likewise, the virtual network request also shows a tree-like structure.

The symbols that are defined above are summarized in Table 3.1.

3.2.4 The Virtual Network Embedding Problem

VNE is a process in which a virtual network is mapped onto a substrate network on the premise of satisfying certain resource constraints. The mapping process is usually divided into two procedures: node mapping and link mapping.

The node mapping procedure is divided into three categories based on the three node types that were described earlier. First, the serving node mapping process is defined as finding the appropriate base stations to serve the service subscribers, based on location information. Second, the gateway node mapping process is one of embedding gateway nodes to backhaul the traffic. Lastly, the caching node mapping process assigns sufficient caching capacity among the substrate nodes that are already included in the embedded virtual network. The node mapping function is described as $Map_N : N_v \to \{N_s, \forall n_v \in N_v\}$.

The link mapping procedure determines paths between embedded nodes based on the VNR link while satisfying resource constraints. The link mapping function is described as $Map_L : L_v \to \{Route_s, \forall l_v \in L_v\}$, where $Route_s$ represents the physical paths between two nodes.

In the next section, we will provide a detailed analysis of the VNE problem and also provide a tailored solution to the problem, specifically for software-defined backhaul networks.

3.3 Virtual Network Embedding in Software-Defined Backhaul Networks

The objective of VNE has two aspects. On the one hand, from the InPs point of view, maximising the profit accrued from accepting VNRs is an inherent objective. To achieve this goal, VNE solutions should minimise the resources spent on one single VNR to improve the VNR acceptance ratio. On the other hand, the SPs have QoS requirements which they must adhere to, such as low latency and sufficient bandwidth, which results in constraints in relation to the VNE problem. In this section, we try to decompose the VNE problem for SDBN, addressing both of these aspects.

3.3.1 Virtual Node Mapping

A special issue in relation to VNoM in SDBNs specifically is that the virtual and substrate nodes are classified into three categories: the serving node, the gateway node and the caching node. The serving nodes are the ones that directly serve the service subscribers through the access network. The gateway nodes refer to those connecting to the core network through

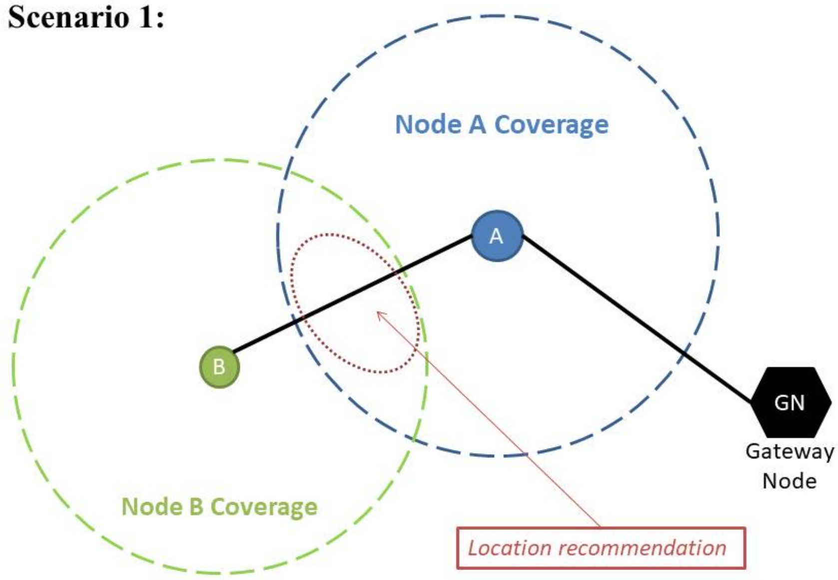

Fig. 3.3 A scenario for virtual serving node mapping

fibres. The caching nodes are the ones with caching functionality. We will analyse the virtual node mapping problem in three sub-sections, as follows.

Virtual serving node mapping

Each base station can serve the users in a certain area based on a particular user association scheme associated with the access network. The backhaul network is only responsible for connecting the serving base station to the core network. Focusing on the backhaul network, we assume that the coverage of each base station is known to the SDN-controller. In this case, the virtual serving node mapping problem for backhaul networks can be simplified to an area coverage problem. Specifically, the VNR provides the location recommendations for the service coverage which is denoted by $A_v^N(Location)$. The virtual node is then mapped onto the substrate nodes with $A_s^N(Coverage)$, which can cover the corresponding area.

However, the virtual serving node mapping process may face a multi-solution problem. If multiple combinations of nodes can provide the coverage that is required, each combination is a solution for the virtual serving node mapping problem. For example, as shown in Fig 3.3 and Fig 3.4, both nodes A and B can cover the $A_v^N(Location)$. Thus, node A and node B are two valid solutions to this serving node mapping process. As node A and node B compete to

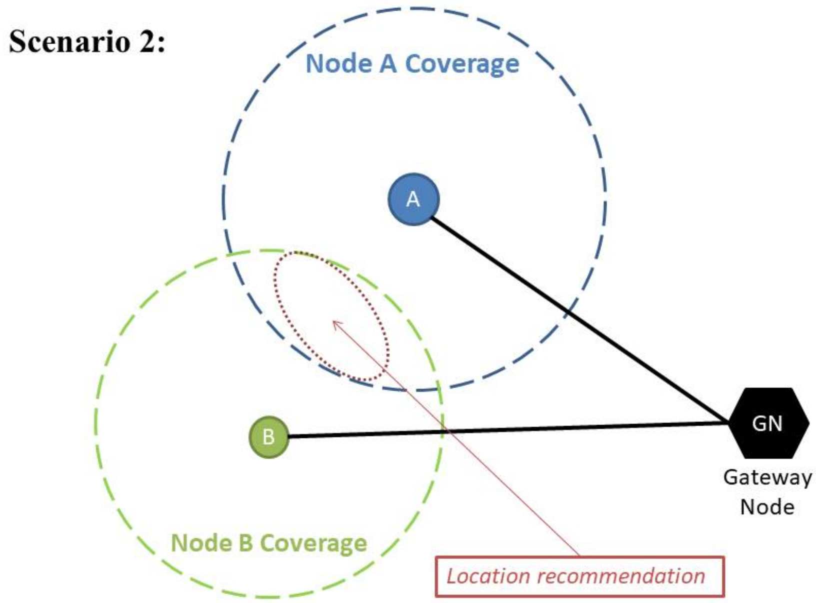

Fig. 3.4 A scenario for virtual serving node mapping

serve the UEs in the same coverage, all traffic dealt with by this procedure can be treated the same.

For the scenario in Fig 3.3, the backhaul hop count metric H is introduced to locate the optimal serving node. The backhaul hop count, $H(i)$, indicates the number of hops from node i to its nearest gateway node, which is related to backhaul performance. First, the majority of backhaul latency comes from the relay process, which is decided by the hop number of the backhaul path. Second, to backhaul the same traffic, the path with fewer hops consumes less network resources, including network links, buffers and computational power. Though routing is not considered in this work, the hop count implies the potential backhaul delay and the network resource consumption to some extent.

In the other scenario, as shown in Fig 3.4, $H(A) = H(B)$, this means that the multi-solution problem cannot be solved from the backhaul network point of view. Thus, signal strength based metrics are employed to choose the better serving node which will provide a better QoS for service subscribers.

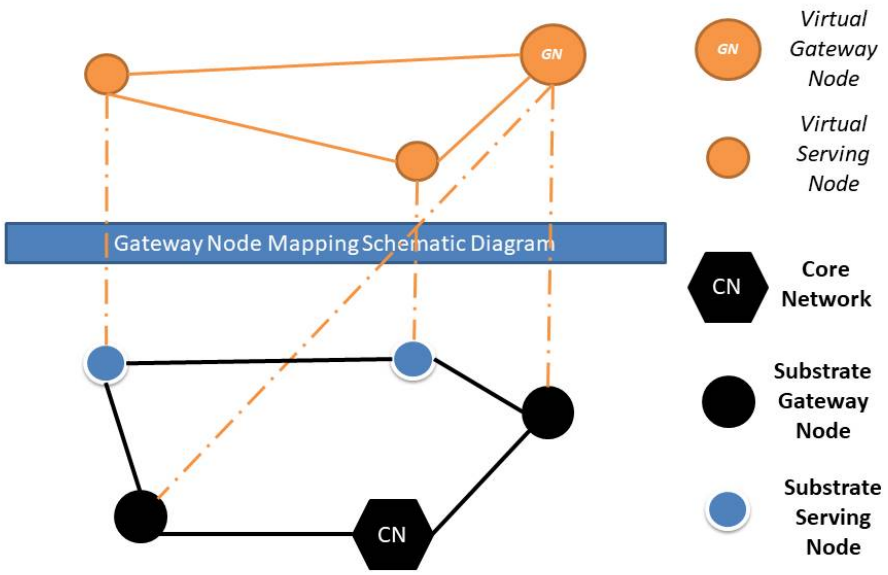

Fig. 3.5 Schematic diagram for virtual gateway node mapping

Virtual gateway node mapping

In a backhaul network, gateway nodes are all connected to the core network through wired links and the destination of most of the traffic will be the gateways; In this regard, all gateway nodes can be treated as the same except for their different locations. Thus, a virtual gateway node can be mapped to any substrate gateway node in order to backhaul the traffic with lower cost. One virtual gateway node can be mapped to multiple substrate gateway nodes, if needed. As shown in Fig 3.5, the virtual gateway node is mapped to two substrate gateway nodes as the hop count is smaller for the serving nodes to get to gateway.

Virtual caching node mapping

With the assumption that the cache capacity demand of one virtual network request is not bound to a particular virtual node, the virtual caching node mapping problem becomes a substrate caching node selection problem. First of all, the cache capacity requirement has to be satisfied. Secondly, the selection of the caching node should reduce the backhaul traffic as far as possible. To solve this problem, a metric to evaluate the caching nodes from the backhaul perspective should be introduced.

The function of a caching node is to cache popular content. The service subscribers

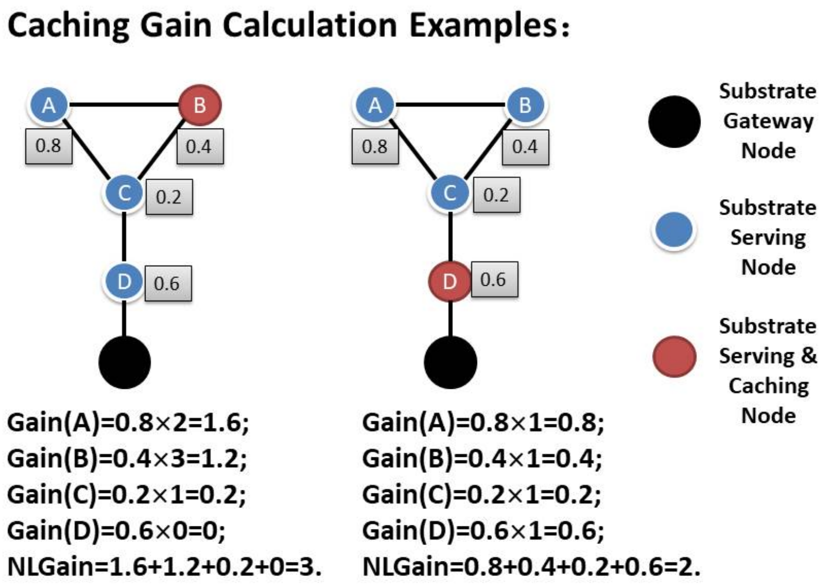

Fig. 3.6 An example for caching gain calculation

can access these contents instead of demanding them from the backhaul gateway node. In a sense, the caching node assists the backhaul network by acting as a gateway node in relation to particular content. Lacking information concerning real-world, real-time traffic, the utilisation ratio regarding content cached by different serving nodes are not emphasised in this work. We assume that the utilisation ratio $\beta(n) \in [0,1]$ is based on the SP prediction of numbers of service subscribers served by serving node n. The utilisation ratio, β, is used to distinguish serving nodes by their cache utilisation. In this case, the caching gain metric *Gain* is proposed to evaluate the caching nodes according to their locations in the network. For serving node i, we define the caching gain $Gain(i)$ as a product of the utilisation ratio, $\beta(i)$, and the difference between the number of hops from node i to its nearest gateway node and the number of hops from node i to the caching node. If the difference is negative, $Gain(i)$ equals zero. $Gain(i)$ can only take non-negative values. Then, the network-level caching gain, *NLGain*, is calculated as the summation of the caching gains of all serving nodes. The assumptions and symbols that are discussed above are summarized in Table 3.2. Two examples of caching gain calculations are shown in Fig. 3.6. A substrate network is randomly generated. The boxes below the nodes stand for the utilisation ratio β. Node B and Node D are chosen as caching node in two situations. The *NLGain* are calculated based on above assumptions. As the *NLGain* are different, it is clear that the location of the caching

Symbol	**Definition**
$\beta(n)$	Utilisation ratio of the cached contents by serving node n (based on the SP prediction of numbers of service subscribers served by serving node n)
$Gain(i)$	Caching Gain of serving node i (a product of the utilisation ratio, $\beta(i)$, and the difference between the number of hops from node i to its nearest gateway node and the number of hops from node i to the caching node)
$NLGain$	Network-level caching gain (summation of caching gain of all nodes)

Table 3.2 Symbols and definitions for caching gain calculation

node has a strong impact on the caching gain.

With the aid of the caching gain metric, *Gain*, the virtual caching node is mapped to the optimal substrate caching node that maximises the network-level caching gain.

3.3.2 Virtual Link Mapping

After the virtual nodes have been settled, the problem of VLiM becomes how to connect the nodes as requested in the VNR with minimal resources while satisfying QoS requirements. To achieve this goal, the weights on the links in the undirected weighted graph of the substrate network can be specified using two metrics as discussed below.

For wireless networks, system spectral efficiency is an important metric due to the interference resulting from the broadcast nature of wireless communications. System spectral efficiency is a measure of the quantity of transmissions that can be simultaneously performed with a limited frequency bandwidth in one particular network. The more the links interfere with each other in the network, the smaller the system spectral efficiency will become as interfering links cannot reuse the same spectral resources. Thus, the interference properties of links may potentially impact the system spectral efficiency, which is the number of links that are interfered with by link l_i. We denote this link interference property as Q_i for link l_i. Without real -world, real-time traffic information, routing and resource allocation algorithms, which may impact on system spectral efficiency, are not taken into consideration. From the InP point of view, it is always of great interest to achieve high system spectral efficiency. Thus, with the link resource demand, D_j, for virtual link l_j, we define the spectrum weight for substrate link l_i as $Weight_{spectrum} = D_j \cdot Q_i$.

From a load balancing point of view, the available resource blocks of links need to be considered when dealing with incoming VNRs. The resource pool matrix, P_i, reflects the available resource status of link l_i. Thus, $\sum_{w=1}^{W} \sum_{t=1}^{T} P_i(w,t)$ is the total number of resource blocks that have been already assigned to link i or its interfering links. In this work, we denote $\sum_{w=1}^{W}\sum_{t=1}^{T} P_i(w,t)$ by $|P_i|$. The larger $|P_i|$ is, the less the available resources which are left for the link. When one link is loaded with too much traffic or its interfering links occupy too much of the shared resource, the link may not be able to hold more incoming VNRs, which brings down the VNE performance of the network. Even worse, overloaded links may lead to traffic jams. Thus, we define the resource weight for substrate link, l_i, as $Weight_{resource} = |P_i|$

In relation to this point, the weights on the substrate links have been defined in the undirected weighted graph. With serving, gateway and caching nodes located, the VLiM problem can be solved by searching for the tree of minimum weight that contains all nodes; this can be treated as a constrained Steiner tree problem. The detailed algorithm to solve the problem is described in the next section.

3.4 The Algorithm of VNE for SDBN

In this section, we propose a algorithm, named VNE-SDBN, which is tailored for the virtual network embedding problem in software-defined backhaul networks. This algorithm exploits the unique characteristics of VNR for backhaul networks to fully utilise the substrate resources. The node mapping stage and the link mapping stage of VNE-SDBN is described as follows.

3.4.1 Node Mapping Stage

In this section, we will give detailed solutions for serving node mapping, gateway node mapping and caching node mapping. The proposed algorithms and discussions follow.

For serving node mapping, it is straightforward to find solutions based on location information. However, to overcome the multi-solution problem, hop counts need to be considered. Thus, we divide all substrate nodes into clusters and hop groups. Each serving node is assigned to a specific hop group in a cluster according to its least hop count to one gateway node. For example, if the hop count of serving node sn_1 to gateway node gn_1 is three and to gateway node gn_2 it is 1, then sn_1 is assigned to hop group one in the cluster that is organised by gateway node gn_2. A demonstration of clustering results related to substrate nodes is shown in Fig. 3.7. The clustering algorithm for substrate nodes is shown in $Algorithm$ 1. HG_0 consists of gateway node indices. Hop index HI_n is the index of the hop group that node n is assigned to. Cluster index, CI_n, indicates the cluster that node n falls into. Hop group matrix HG_h contains the indices of the nodes that are assigned to the h^{th} hop group, in which all nodes can reach the gateway nodes in h hops.

Algorithm 1: Substrate Node Clustering Algorithm

Input:
• Undirected graph: $G(N,L)$
• Gateway node array: HG_0
• Hop index for node n: HI_n
• Cluster index for gateway nodes: CI_n
Output:
• Hop group matrix: HG_h
• Hop index for node n: HI_n
• Cluster index for node n: CI_n
• Maximum hop count: H
Initialisation:
• Node Counter: $NC = 0$
• $HI_n = null$

```
while NC < N do
    for h = 0 to ∞ do
        forall the n ∈ HG_h do
            if HI_n = null then
                NC = NC + 1
                HI_n = h
                H = h
                forall the node m in the transmission range of node n do
                    if HI_m = null then
                        Assign node m to HG_{h+1}
                        CI_m = CI_n
                end
        end
    end
end
```

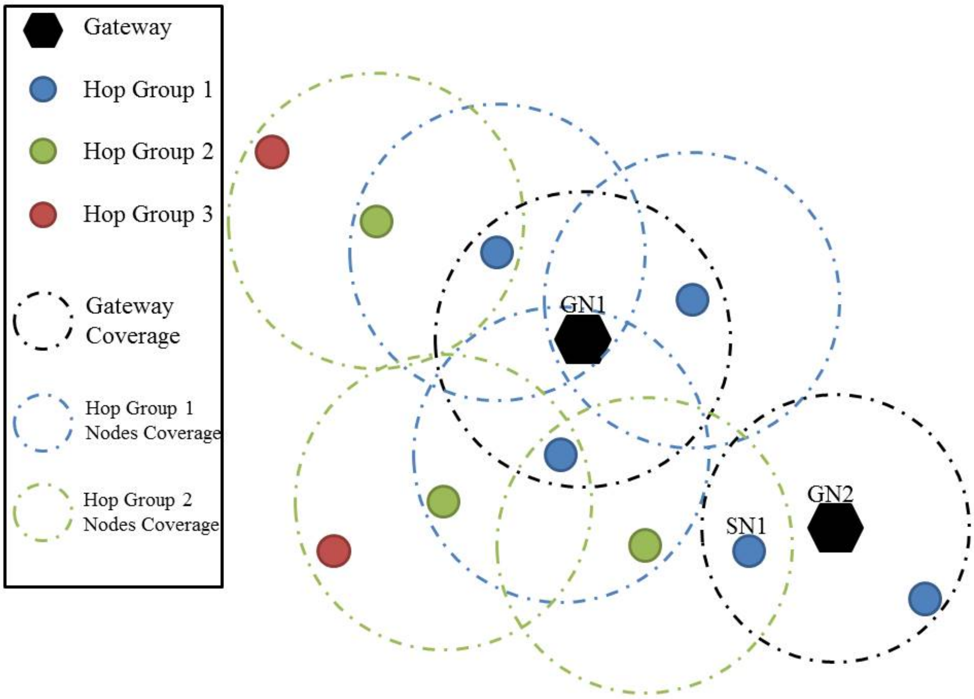

Fig. 3.7 Clustering demonstration

After the clustering procedure, we carry out the serving node mapping procedure as shown in *Algorithm* 2. First, candidate substrate serving nodes are picked out for each location recommendation area. If one serving node cannot cover the whole area, the algorithm will add another serving node to the solution candidates. The algorithm prefers solutions with less serving nodes – from a resource saving point of view. Secondly, the hop index of candidate solutions are compared. The solution with the least hop-count is kept. If there remain multiple solutions, the algorithm will access the signal strength based metrics in the SDN controller and pick the serving nodes that can provide better QoS.

The gateway node mapping procedure needs to consider network performance and resource savings for subsequent procedures; these considerations are related to the hop counts from serving node to gateway node. Since we have grouped all the nodes into clusters by hop counts to gateway nodes, the gateway nodes in the same cluster as the mapped serving nodes are the ones reachable with the least hops. Thus, the virtual gateway node mapping procedure becomes straightforward with the aid of clusters.

For the caching node mapping procedure, we propose the caching node mapping algorithm shown in *Algorithm* 3. As discussed in the last section, two conditions have to be met to embed a caching node. First, the available cache capacity should be larger than the virtual cache capacity demanded. Thus, in the first part of the algorithm, all caching nodes that

---

Algorithm 2: Serving Node Mapping Algorithm

Input:
- Undirected graph: $G(N,L)$
- Serving node coverage: $A_s^N(Coverage)$
- Location recommendation: $A_v^N(Location)$
- Hop index for node n: HI_n
- Signal strength based metrics

Output:
- Serving node mapping solution: $Solution_{serving}$

Initialisation:
- Solution Counter: $SC = 0$
- Candidate Serving Node Set: $CSN = \emptyset$
- Serving node mapping solution: $Solution_{serving} = \emptyset$

forall the *Location recommendation* $A_v^N(Location)$ **do**
    forall the *Serving node* $n \in N$ **do**
        if $A_s^N(Coverage(n))$ *totally covers* $A_v^N(Location)$ **then**
            $SC = SC + 1$;
            $Candidate(SC) = \{n\}$;
        if $A_s^N(Coverage(n))$ *partly covers* $A_v^N(Location)$ **then**
            Add node n into CSN;
    end
    for $i = 2$ *to* $|CSN|$ **do**
        forall the *Node combinations with i nodes* **do**
            if $A_s^N(Coverage)$ *totally covers* $A_v^N(Location)$ **then**
                $SC = SC + 1$;
                Add the combination of i nodes to $Candidate(SC)$;
        end
        if $SC \neq 0$ **then**
            End Loop.
    end
    for $i = 1$ *to* SC **do**
        $HopCount(i) = \sum_{n \in Candidate(i)} HI_n$;
    end
    Rank $Candidate$ by $HopCount$;
    Keep $Candidate$ with the least $HopCount$;
    while *there are multiple* $Candidate$ *remaining* **do**
        Access the signal strength based metrics for each $Candidate$ for the location recommendation area;
        Remove the $Candidate$ with worse metrics;
    end
    $Solution_{serving} = Candidate$;
end

---

satisfy this requirement are listed as candidates. Second, the caching node should assist the backhaul network to reduce traffic. Caching gain, *Gain*, is employed to evaluate the nodes in relation to this. To get the hop count information from the caching node to other nodes, *Algorithm* 1 is run with the caching node set as the only gateway node in the network. Then the network-level caching gain, *NLgain*, is calculated for each candidate caching node. The caching node with the largest *NLgain* is then returned as the solution from this caching node mapping procedure.

Algorithm 3: Caching Node Mapping Algorithm

Input:
- Undirected graph: $G(N,L)$
- Solution for serving node mapping: $Solution_{serving}$
- Hop index for node n: HI_n
- Cache utilisation ratio for serving node n: $\beta(n)$
- Available cache capacity of substrate nodes: $A_s^N(AvCache)$
- Cache capacity demand of virtual caching node: $A_v^N(DeCache)$

Output:
- Caching node mapping solution: $Solution_{caching}$

Initialisation:
- Candidate Caching Node Set: $CCN = \emptyset$
- Caching node mapping solution: $Solution_{caching} = \emptyset$

```
for Cache capacity demand A_v^N(DeCache) do
    forall the Caching node n ∈ N do
        if A_s^N(AvCache(n)) > A_v^N(DeCache) then
            Add node n to CCN;
    end
    forall the Caching node n ∈ CCN do
        Regard the caching node n as the only gateway node in the network
        and run algorithm1 to get the hop index of serving node m: HI*_m.
        forall the Serving node m in Solution_serving do
            if HI_m − HI*_m > 0 then
                Gain(m) = β(m) ∗ (HI_m − HI*_m);
            else
                Gain(m) = 0;
            NLGain(n) = NLGain(n) + Gain(m);
        end
    end
    Rank NLGain by value;
    Return the caching node index with max NLGain to Solution_caching
end
```

3.4.2 Link Mapping Stage

Algorithm 4: Next Hop Group Algorithm

Input:
- Undirected graph: $G(N,L)$
- Mapped serving node of virtual link l_i: N_s^i
- Hop group matrix: HG_h
- Hop index for node n: HI_n
- Cluster index for node n: CI_n
- Resource demand for virtual link l_i: D_i

Output:
- Mapped substrate links for virtual link l_i: L_s^i

$Tx = N_s^i$
for $h = HI_n$ *to* 1 **do**
  Update P for interferences links
  $Weight \to \infty$
  forall the $\{m | m \in HG_{h-1} \wedge CI_m = CI_n\}$ **do**
    $Rx = m$
    if $\{L(Tx,Rx)$ *exists* $\wedge |P| - |P_{L(Tx,Rx)}| > D_i \wedge Weight_{L(Tx,Rx)} < Weight\}$ **then**
      $L_s^i(h) = L(Tx,Rx)$
      $Weight = Weight_{L(Tx,Rx)}$
      $Tx = Rx$
  end
end

In the link mapping stage, we aim to connect all the selected serving nodes to gateway nodes and to the caching node, using minimal-weighted trees. The most serious issue related to virtual link mapping in SDBNs is that the available bandwidth is dynamic. Due to interference and resource allocation schemes, the available bandwidth for one link may vary after every link mapping decision. Thus, the weights on links need to be updated after each substrate link is chosen. In this case, a global optimal solution cannot be achieved in polynomial time – the situation is too complex. Thus, we propose the hop group algorithm shown in *Algorithm* 4 as a means to search for the minimum weight in terms of the route from the serving node to the gateway node in the same cluster or to the caching node, hop by hop. To compare the weight of the direct link between any two nodes with that of multi-hop links, we provide and use *theorem* 1. Taking advantage of this theorem, only direct links between hop groups are considered in the proposed algorithm.

Theorem 1. *The weight of the direct link between any two nodes is smaller than that of multi-hop links connecting them.*

Proof. We start from a two-hop scenario. First, we assume that the link from node m to node n is l_i, the link from node m to node o is l_j, and the link from node o to node n is l_k. The weight of link l_i, denoted by $Weight_{l_i}$ is $D \cdot Q_i + |P_i|$. Similarly, the weight of link l_j and link l_k, is denoted by $Weight_{l_j+l_k}$ is $D \cdot Q_j + |P_j| + D \cdot Q_k + |P_k|$. As link l_i and link l_j share the same transmission node, m, it can be derived that $Q_i = Q_j$, from the definition of Q. Link l_i and link l_k share the same receiving node, n. It is obvious, then, that $|P_i| < |P_k|$ because the resource allocation of link l_j reduces the available resources for link l_k. Thus, $Weight_{l_j+l_k} - Weight_{l_i} = |P_j| + D \cdot Q_k + (|P_k| - |P_i|) > 0$. The basic verification process for multi-hop scenarios is the same as it is for two-hop scenarios.

□

For virtual links between serving nodes and gateway nodes, *Algorithm* 4 can be directly applied to map the virtual link to substrate links. However, for virtual links between serving nodes and the caching node, we need to run *Algorithm* 1 again with the caching node set as the only gateway node. Then, such virtual link requests can be solved with the aid of *Algorithm* 4.

3.4.3 The Proposed Algorithm

Based on *Algorithm* 1, *Algorithm* 2, *Algorithm* 3, *Algorithm* 4 and the discussions above, we can propose the virtual backhaul network embedding algorithm called VNE-SDBN, as *Algorithm* 5.

To start with, the algorithm divides the substrate nodes by cluster and hop count groups, using *Algorithm* 1. Then, a single VNR must be processed. The virtual serving nodes are mapped with *Algorithm* 2 one by one and subsequently a gateway node within the same cluster as the serving nodes is chosen as one of the embedded gateway nodes. After all serving nodes and gateway nodes are located, the caching node is selected with *Algorithm* 3. Finally, all the virtual links are mapped with *Algorithm* 4, where the caching node is treated as a gateway node.

3.4.4 Time Complexity Analysis

The time complexity is the computational complexity that estimates the time taken for running an algorithm [98]. As the running time of an algorithm may vary due to different inputs, worst-case time complexity, which is the maximum amount of time taken on inputs of a given size, is commonly considered.

The time complexity of *Algorithm* 1 is $\mathcal{O}[HN^3]$, where H is the maximum hop count

Algorithm 5: VNE-SDBN

Group and cluster the substrate nodes with $Algorithm$ 1.
forall the *Incoming VNRs* **do**
  forall the *virtual serving nodes* **do**
    Map the nodes with $Algorithm$ 2;
    Find the *gateway node* gn_s in the same cluster as the mapped serving nodes
    $Map_N : gn_v \rightarrow gn_s$
  end
  for *virtual caching node* **do**
    Map the node with $Algorithm$ 3;
  end
  forall the *virtual links between serving nodes and gateway nodes* **do**
    Map these virtual links with $Algorithm$ 4.
  end
  forall the *virtual links between serving nodes and the cache node* **do**
    Regroup the substrate nodes by regarding the caching node as the only gateway node with $Algorithm$ 1;
    Map these virtual links with $Algorithm$ 4.
  end
  Return the VNE result.
end

of all nodes and N is the total number of nodes. For one virtual serving node mapping in $Algorithm$ 2, the time complexity is $\mathcal{O}[2^N]$. The time complexity of gateway node mapping is $\mathcal{O}[1]$. The time complexity of the caching node mapping procedure in $Algorithm$ 3 is $\mathcal{O}[HN^4]$. For $Algorithm$ 4, the time complexity is $\mathcal{O}[HWTNL]$, where WT represents the total resource blocks and L, the total number of links. Thus, the time complexity of our proposed VNE-SDBN is $\mathcal{O}[HN^3] + \mathcal{O}[IN(2^N + HN^3 + HWTL^2)]$, where I is the number of VNRs.

The main time-consuming element of the algorithm is running $Algorithm$ 1 repeatedly. In practice, once the network topology is fixed, the SDN controller can run $Algorithm$ 1 for all caching nodes, once, and then store the results for later use. Thus, the proposed algorithm can be faster when dealing with new VNRs.

3.5 Performance Evaluation

The simulation settings and results will be presented in the following section. In the first part of the simulation, we verify the relationship between the proposed caching gain and the backhaul traffic reduction. In the second part of the simulation, two vital metrics for

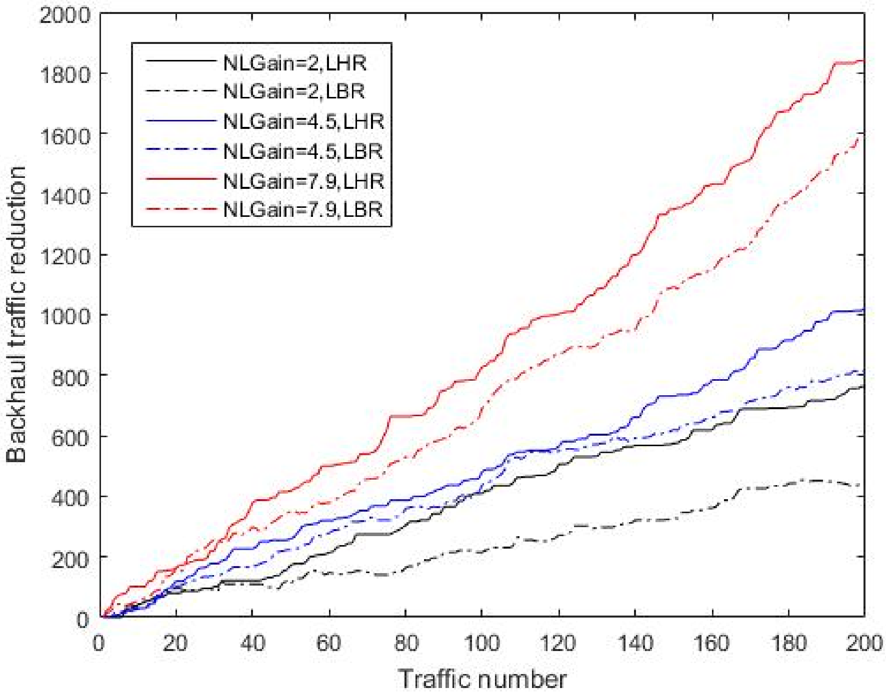

Fig. 3.8 Network-level caching gain verification

VNE problems, the revenue and the VNR acceptance ratio, will be employed to evaluate our VNE-SDBN algorithm. Due to the particularity of the VNE model in SDBN, the virtual node mapping procedure is not comparable with those in previous studies. In this work, we focus on virtual link mapping performance and compare our algorithm with two other state of the art algorithms, the shortest-path based link mapping algorithm, VNE-GRC, [99] and the load balancing guaranteed mapping used in FELL [100].

3.5.1 Caching Gain Metric Verification

A caching gain metric was proposed in section 3.3; this is used to evaluate the caching node according to its ability to relieve the backhaul traffic pressure. In this section, we perform a simulation to see if the caching gain metric is effective in terms of caching node evaluation.

A substrate network with ten nodes is randomly generated. One node is selected as a gateway node. The other nine nodes are serving nodes which are each randomly assigned a caching utilisation ratio $\beta \in (0,1)$. We choose three caching nodes with network-level caching gains (*NLGain*) of 2, 4.5 and 7.9 respectively. Backhaul demands which can take advantage of the caching contents are randomly generated from random serving nodes. The throughput demand of each item of traffic follows a uniform distribution between 0 and 10.

The backhaul traffic reduction is calculated as the cumulative traffic throughput reduction when serving nodes obtain content from the caching node instead of the gateway node. Two routing algorithms are considered in this simulation: least-hop routing (LHR) and load-balancing routing (LBR).

As shown in Fig. 3.8, all metrics ascend with increasing traffic, which means that the backhaul traffic is relieved by the caching node. In both routing algorithms, the line with the larger *NLGain* ends up higher than the ones with smaller *NLGain*. This result indicates that the caching gain metric can evaluate the caching node in terms of its ability to reduce backhaul traffic. Also , the caching gain metric is related to hop numbers; this indicates that the effect of backhaul traffic reduction is better when LHR is applied for the same caching node.

3.5.2 VNE-SDBN Performance Analysis

In this section, the performance of VNE-SDBN is evaluated through simulations. First, the parameters of the simulations are stated. Then, two metrics, which are the revenue of InPs and the VNR acceptance ratio , will be employed to evaluate the performance of the proposed algorithm.

Simulation Setup

The substrate network is randomly generated in an area of 1000 m$\times$1000 m with 20 nodes in total, two of which are gateway nodes. The transmission and interference ranges are set as 250 m and 400 m, respectively. We assume that VNRs arrives in a Poisson process with one request per timeslot on average and have durations that follow an exponential distribution with a mean of ten timeslots. In each VNR, the number of serving nodes follows a uniform distribution between 1 and 3. For the resource block settings, we divide the whole spectral resource into 50 sub-channels and 20 time slots across one second. We assume that the spectral resource requirement of VNRs follows a uniform distribution between 20 and 40.

Revenue of InPs

As the network is based on OFDMA, we consider that the InPs have fixed cost in terms of total spectral bandwidth and infrastructure construction. The revenue comes from VNRs satisfied with sufficient spectral resources. Thus, we take the total resource blocks of successfully accepted VNRs as the revenue. Fig. 3.9 shows the real-time revenues for the three algorithms and the cumulative revenues are shown in Fig. 3.10. For the real-time revenues, in the beginning, th curves ascended quickly because the network had enough resources to accept

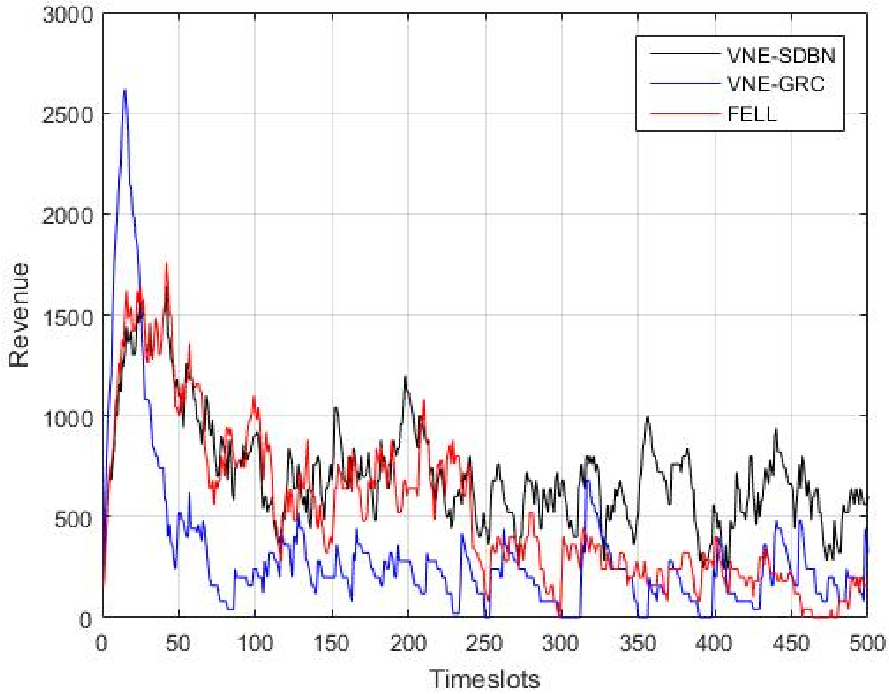

Fig. 3.9 Comparison of real-time revenues

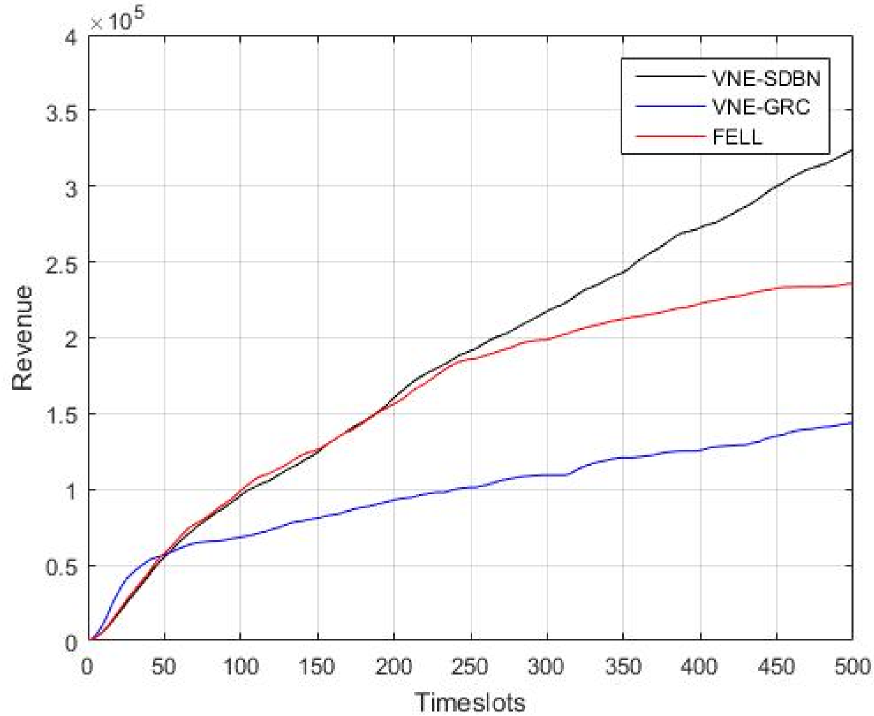

Fig. 3.10 Comparison of accumulative revenues

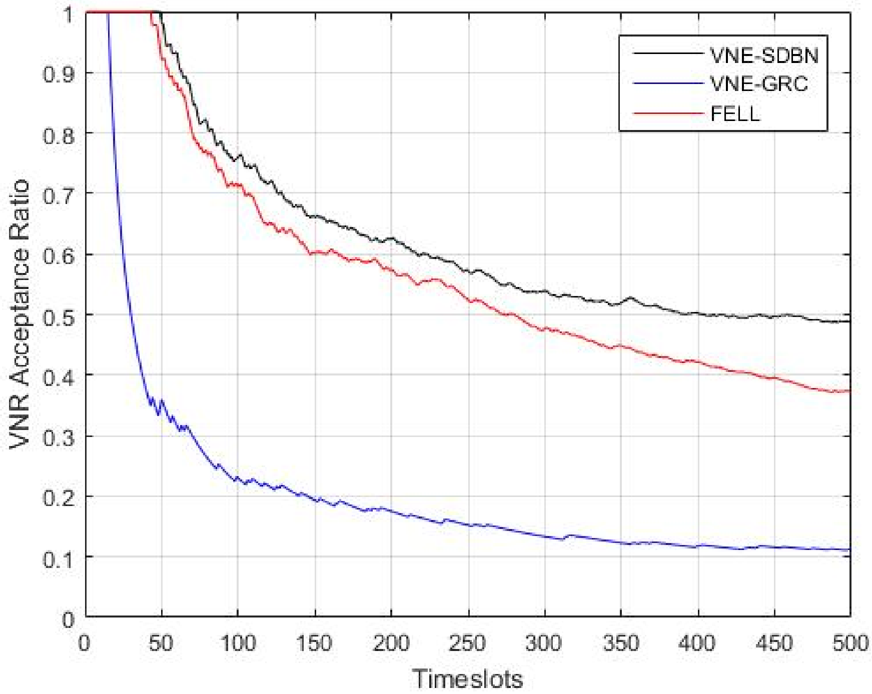

Fig. 3.11 Comparison of VNR acceptance ratio

incoming VNRs. As more and more VNRs was embedding to the substrate network, the virtual network embedding procedure can not be carried out until the network had enough resources to satisfy them. Thus, the curves fluctuated up and down due to the duration of the accepted VNRs. For the cumulative revenues, the curve ascended all the time as the revenue was increased with the number of accepted VNRs. Our algorithm, VNE-SDBN, clearly improves revenues compared with the other two algorithms.

Acceptance ratio of VNR

The acceptance ratio of VNRs is the ratio between the number of VNRs which have been successfully mapped and the total number of requests. Fig. 3.11 shows the VNR acceptance ratios of all three algorithms. In the beginning, the substrate resources are sufficient so that the three ratios start with high values. As the substrate resources are consumed, the acceptance ratios decline over time. Our algorithm, VNE-SDBN, maintains higher VNR acceptance ratios than the other two algorithms.

3.6 Conclusions

In this chapter, the VNE problem in relation to OFDMA-based SDBNs is investigated. Based on the SDBN system model, the virtual-node-mapping problem and the virtual-link-mapping problem are separately discussed. The virtual node mapping problem can be divided into the serving node mapping, gateway node mapping and the caching node mapping problems, based on the node types in the backhaul network. For the serving node mapping problem, substrate nodes are selected based on the serving coverage and the location recommendation information of the VNR. Gateway nodes are mapped based on the principle of resource saving. To solve the caching node mapping problem, we proposed the caching gain metric, *Gain*, which evaluates the caching nodes by their ability to reduce backhaul traffic. Then, the virtual-link-mapping problem is analysed on the basis of two metrics: the interference property metric and the resource pool metric. To solve the NP-hard mapping problems, we proposed the VNE-SDBN algorithm in accordance with our deliberations. At last, a verification of our proposed caching gain metric and simulations of the VEN-SDBN algorithm alongside two classic VNE algorithms were carried out. Numerical results show that the proposed algorithm outperforms state of the art algorithms in terms of improving revenues and the VNR acceptance ratio, compared with the two other algorithms.

Chapter 4

Joint Routing and Resource Allocation in Software-Defined Backhaul Networks

Overview

By decoupling the control plane from the data plane and providing programmability for network applications, SDNs are positioned to offer more efficient management, higher flexibility and better performance. Routing and resource allocation are two closely related applications in terms of wireless networks. With close cooperation, better performance and lower complexity can be achieved using a SDN architecture. However, work that jointly studies routing and resource allocation is rarely seen. In this work, the joint routing and resource allocation problem is investigated in OFDMA-based SDBNs. Two scenarios, which are networks operated in LFM and those operated in mmW, are considered. To exploit the SDN programmability, an SDBN system model is proposed, where the control panel can use highly complex algorithms in the configuration phase in order to simplify the algorithms which must be used in the operation phases. Then, the joint routing and resource allocation problem is formulated as an achievable system throughput optimisation problem. As a benchmark, a genetic algorithm (GA) is employed to obtain a near-optimal solution. By constructing the interference digraph of the network and analysing the vertex degree characteristics, a digraph-based greedy algorithm (DBGA) is proposed. Simulation results have shown that the proposed algorithm can efficiently increase the system throughput.

4.1 Introduction

Owing to their simplified management, high flexibility and improved performance, SDNs have drawn considerable attention in recent years [5–7]. By separating the control plane and the data plane, SDNs can offer logical centralisation of network management and introduce programmability, which opens up new approaches to control functions in the application layer. Routing and resource allocation are the main functions of wireless networks. Traditionally, routing and resource allocation designs are based on distributed approaches, which may confront problems such as high complexity and performance limitations. With the introduction of SDN, problems can be broken into tractable pieces and decisions can be made based on global network information.

Several papers have been presented which feature designs for utilising SDN features for better routing and better resource allocation separately. In [101], the authors built a model and developed algorithms for load balancing with packet forwarding rules. By migrating traffic from heavy loaded switching devices to lightly loaded ones, the proposed algorithm in [102] achieved load-balancing routing reactively. Multipath routing, where traffic can be forwarded with multiple paths, was exploited in [103] to provide load balancing and improve network throughput. In [104], a novel method was proposed to build a multicast topology, using SDN. In [105], a price-based joint allocation model of network resources in SDN is built by introducing a price for each of the resources; this can result in the fair, proportional allocation of link bandwidth and in minimum global delay at the same time. However, work that jointly considers both routing and resource allocation is rarely seen in SDNs.

A strong interaction exists between routing and resource allocation. For instance, when routing traffic , the available spectral resource is always a metric that needs to be considered. In return, when all routes are settled, the spectral resources need to be assigned based on the routed traffic in each transmission link. Especially when interference exists in a network, the amount of traffic in each link may impact the resource allocation strategy and further affect the capacity of the network. Hence, jointly considering routing and resource allocation is of fundamental importance. More importantly, to deal with the explosive growth of traffic over wireless networks [1], it is crucial to gain higher network throughput, and this may be achieved if routing and resource allocation cooperate well to increase the reusability of spectra .

In this work, we investigate joint routing and resource allocation in an OFDMA-based SDBN, where multi-hop transmission and mesh topology are applied. We have proposed an SDBN system model for routing and resource allocation functions which takes account of the features of SDN. Based on our system model, we formulated the joint routing and resource allocation problem as an achievable system throughput optimisation problem. The genetic algorithm was employed to obtain a near-optimal solution as a benchmark. To solve the problem in a decomposition manner, an interference digraph was constructed. Through analysing the in-degree and out-degree of the vertices in the digraph, a low complexity greedy

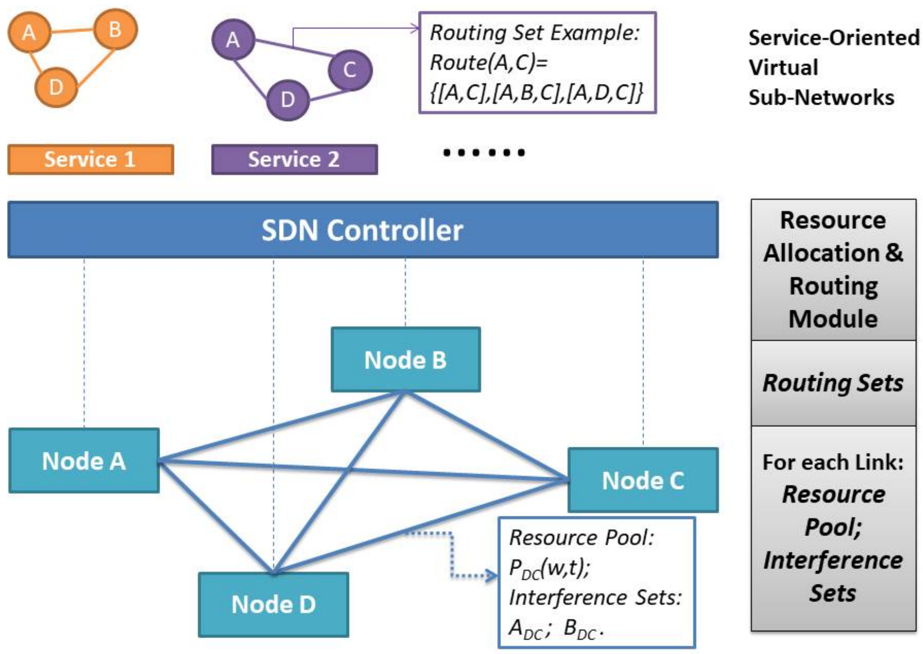

Fig. 4.1 Software-defined backhaul network with resource allocation and routing module

algorithm can be proposed.

The remainder of this chapter is organised as follows. In Section 4.2, we propose the SDBN system model in detail. Then, we formulate the joint routing and resource allocation problem and give a near-optimal solution using genetic algorithm in Section 4.3. Next, we analyse the problem on an interference digraph basis and propose our digraph-based greedy algorithm in Section 4.4. Simulation results and analyses are presented in Section 4.5. Finally, we conclude this chapter in Section 4.6.

4.2 System Model

In this section, we will introduce the software-defined backhaul network and some components that are related to resource allocation and routing, as shown in Fig. 4.1. A mathematical model will be developed for each subsection.

4.2.1 Software-defined Backhaul Network

A backhaul network is considered in this work, in which a basic SDN architecture is adopted. It consists of two major components: an SDN controller in the control plane and a set of switching devices in the data plane. The SDN controller is a centralised software-defined control unit, which can communicates with all the switching devices through the control panel. The switching devices, which are denoted as nodes in the following for simplification, are responsible for collecting and sending network statuses to the controller and for processing data packets based on rules provided by the controller. In the networks which we consider here, the nodes are small cell BSs, among which gateways are randomly placed. It is worth noting that communication is not limited to that which occurs between BSs and gateways. With caching techniques [93] in place, communication between any two BSs may be required. N nodes and L directional links are employed in this paper.

Two scenarios are considered in this work: a backhaul network operated in the LFM spectrum and one operated in the mmW spectrum. Due to the distinctive characteristics of these two frequency bands, the models for the two scenarios will be different. More details will be given in the following sections.

4.2.2 Service-oriented Virtual Sub-network

In SDNs, network virtualisation is a key technique which enables distinct control over different services (e.g. VoIP and download) within the same architecture [106]. For one specific service, the SDN-controller assigns nodes and links based on the demand (e.g. throughput and latency) of the service; this forms a service-oriented virtual sub-network. For more details about virtual networks, please refer to Chapter 3.

In the configuration phase of one sub-network, the nodes and links can be processed to produce feasible paths between any two nodes, using a depth-first-search algorithm [107]. All feasible paths from node m to node n are recorded in set $Route(m,n)$, the size of which is $F(m,n)$. The f^{th} feasible path is denoted as $Route^f(m,n)$ whose elements are link indices.

In the operation phase of the sub-network, we consider traffic requests as multi-hop node to node transmission rate demands. For one traffic request from node m to node n, the demand is denoted by $D(m,n)$ bits/s. Further, we use D_i to denote the traffic that passes through link l_i. If traffic request $D(m,n)$ routes through link l_i, $D_i(m,n) = D(m,n)$. Otherwise, $D_i(m,n) = 0$. We denote the total number of traffic requests with M.

4.2.3 Resource Pool

In SDBN, the controller is responsible for monitoring the usage of spectral resources and controlling the resource allocation and the routing strategy of the network. We assume that the spectral resource is divided into W sub-channels, and that there are T time slots in each one second. One sub-channel w in one time slot t is called a resource block $b(w,t)$. The total

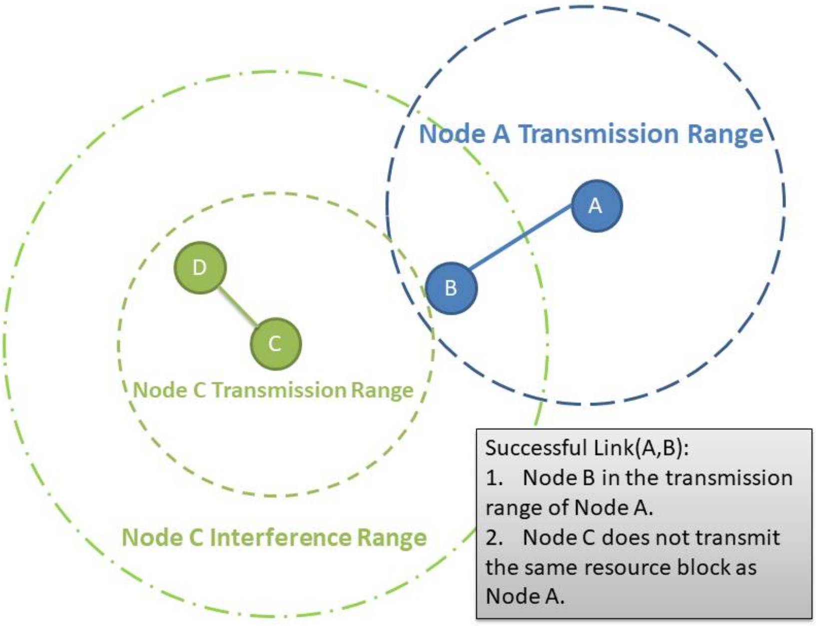

Fig. 4.2 Protocol interference model illustration in LFM-based backhaul network

number of resource blocks in one second is wt. One BS can transmit and receive data on any combination of the resource blocks. We propose the construction of a resource pool, P, in the controller to help manage the resource usage. Each link, l_i, has its resource pool matrix $P_i(w,t)$. If $P_i(w,t) = 1$, This means that the resource block $b(w,t)$ is assigned to The link l_i. Otherwise, resource block $b(w,t)$ is not in use in the corresponding link.

4.2.4 Interference Model

In contrast to an access network, a special issue of SDBNs is that the architecture and topology are stationary. Thus, it is reasonable to assume that every node and link is preconfigured with appropriate parameters including transmission power, antenna beam-forming and other factors that impact interference between links. Because of this precondition, the protocol model [96], is adopted in this work. Two conditions have to be met for a successful transmission across link l_i: 1) the receiving node must be in the transmission range of the transmitting node; and 2) if the receiving node is in the interference range of any other link, it must be ensured that that link is not occupying the same resource block as link l_i.

As shown in Fig. 4.2, we assume that the antennas in the LFM-based backhaul network scenario are omnidirectional. Thus, interference management is crucial in this scenario.

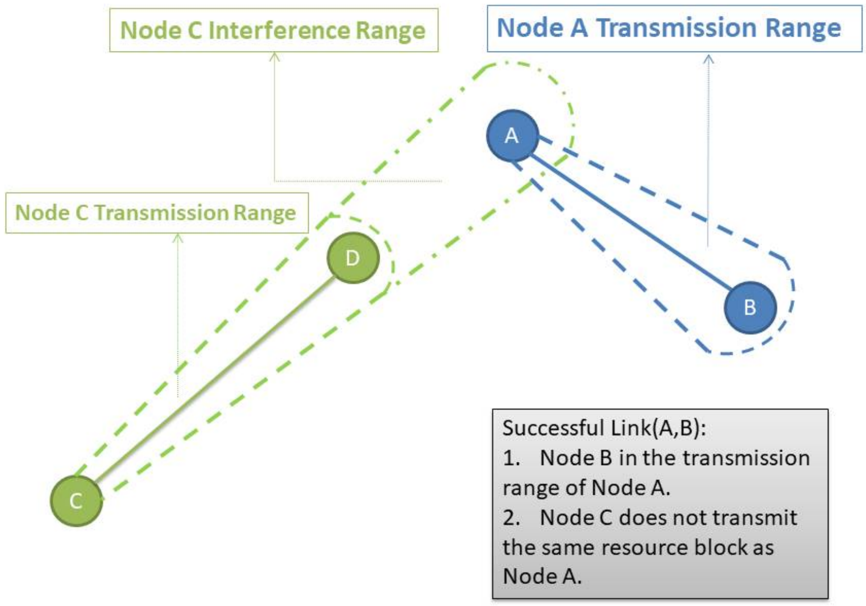

Fig. 4.3 Protocol interference model illustration in mmW-based backhaul network

In the mmW-based backhaul network scenario, the antennas are directional due to the special transmission features of mmW [77]. As shown in Fig. 4.3, in contrast to an omnidirectional antenna, the antenna transmits a beam with a particular radiation angle. Only the nodes along the beam will be interfered with, which decreases the potential interference in the network.

With a stationary network topology and configuration, there are two inherent interference properties for each link; these must be taken into consideration, and defined and recorded in the SDN-controller. The set of links that can interfere with link l_i is denoted by A_i. Likewise, the set of links that are interfered with by link l_i is denoted by B_i. A_i and B_i are called the interference sets of link l_i.

4.2.5 Link Transmission Rate

As OFDMA is adopted in this work, RB is the smallest unit of carriers. Due to different channel conditions of various links, the bit rates of data conveyed by one RB in different links may vary [108]. Thus, we define link transmission rate as the bit rate per RB for a certain link. In mathmatical notation, the link transmission rate of link l_i is denoted by r_i bits/s per RB assigned.

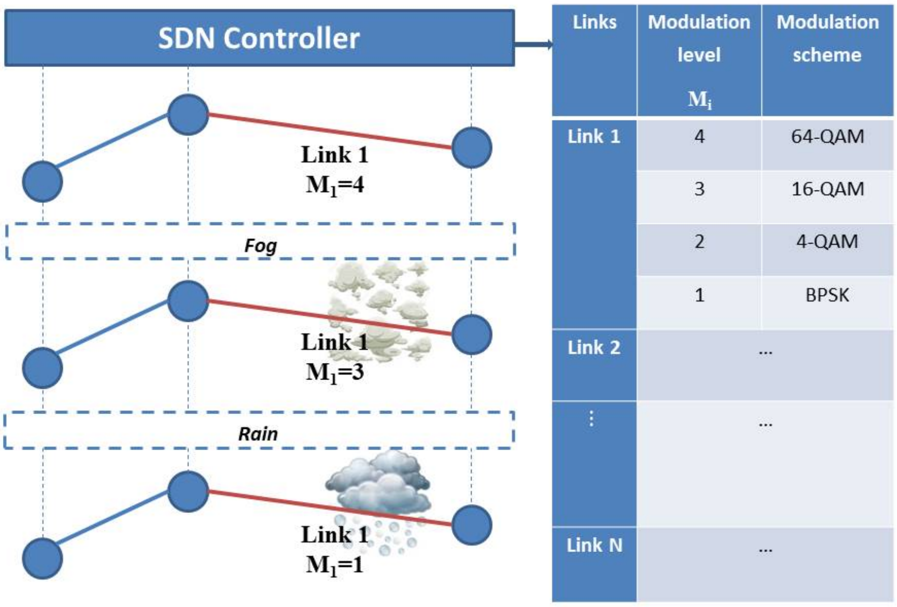

Fig. 4.4 Adaptive modulation illustration in mmW-based backhaul network

In mmW-based backhaul networks, the weather conditions significantly affect the attenuation of the signals [74, 75]. Thus, the adaptive modulation technique [109] can be adopted to increase the data rate by adjusting the modulation scheme according to channel state information. In this case, the link transmission rate is significantly impacted by the modulation scheme that is applied to the link. The modulation level, M_i, is employed to indicate the modulation scheme that is applied to link l_i, as shown in 4.4. For the same link, the larger the modulation level M_i, the greater the link transmission rate r_i. In this scenario, the link spectral efficiency caused by different modulation schemes should be considered in resource allocation and routing algorithms. Further discussion of this will be undertaken in Section 4.3.

However, low-frequency microwave is practically free from the effects of weather conditions. Moreover, the nodes are located at fixed locations in backhaul networks. In the configuration phase of the backhaul network, the transmission parameters of the nodes and links can be adjusted to their optimal state to achieve a high level modulation scheme. Thus, we can further assume that each link has a fixed link transmission rate and the rate difference between the links is not significant. In this scenario, the impact of link spectral efficiency on system throughput can be ignored when routing traffic and allocating spectral resources.

4.2.6 System Throughput

Throughput is defined as the rate of successful message delivery over a communication channel [108]. Thus, link throughput can be defined as the bit rate of data that is transferred through a specific link. For link l_i, given the link transmission rate r_i and the number of RBs that are assigned, the link throughput is computed by:

$$R_i = r_i \sum_{w=1}^{W} \sum_{t=1}^{T} P_i(w,t). \tag{4.1}$$

At the network level, the system throughput is given by the sum of the throughputs of all the links [108]. With the assumption that S is the resource allocation scheme applied to the network, the system throughput is computed by:

$$R_S = \sum_{i=1}^{I} R_i = \sum_{i=1}^{I} (r_i \sum_{w=1}^{W} \sum_{t=1}^{T} P_i(w,t)). \tag{4.2}$$

We assume that all resources are fully utilised in the network, in which case any resource block in addition to those that have already been assigned will cause interference. From an interference free point of view, the system throughput can be rewritten as:

$$R_S = \sum_{i=1}^{I} r_i \sum_{w=1}^{W} \sum_{t=1}^{T} max(0, 1 - \sum_{j \in A_i} P_j(w,t)). \tag{4.3}$$

4.3 Joint Routing and Resource Allocation Problem

In SDNs, routing and resource allocation are two basic and closely related issues. Different routing strategies may lead to different resource allocation schemes, which may further seriously impact system performance. In turn, resource allocation algorithms may also impact the routing decisions. In the following, we try to jointly consider the routing and resource allocation problems involved in order to improve the achievable throughput of SDBNs.

4.3.1 Problem Formulation

Considering the explosion in traffic which is currently being experienced, it is always going to be of great interest to obtain the optimal achievable system throughput so as to be able to load more traffic onto SDBNs. Thus, the objective function of the joint resource allocation and routing problem becomes:

$$\text{Maximise } R_S = \sum_{i=1}^{I} R_i = \sum_{i=1}^{I} (r_i \sum_{w=1}^{W} \sum_{t=1}^{T} P_i(w,t)) \tag{4.4}$$

Meanwhile, we have the following resource block constraint:

$$\begin{aligned} & P_i(w,t) \in \{0,1\} \\ & \forall i = 1,2,\ldots,L \quad \forall w = 1,2,\ldots,W \quad \forall t = 1,2,\ldots,T. \end{aligned} \tag{4.5}$$

A link interference constraint is required to avoid interference between links:

$$\begin{aligned} & P_i(w,t) + P_j(w,t) \leq 1 \\ & \forall i = 1,2,\ldots,L \quad \forall j \in A_i \\ & \forall w = 1,2,\ldots,W \quad \forall t = 1,2,\ldots,T \end{aligned} \tag{4.6}$$

where A_i is one of the interference sets of link l_i.

To guarantee the resources are fully exploited, the following achievable throughput constraint is defined:

$$\begin{aligned} & P_i(w,t) = max(0, 1 - \sum_{\forall j \in A_i} P_j(w,t)) \\ & \forall i = 1,2,\ldots,L \quad \forall j \in A_i \\ & \forall w = 1,2,\ldots,W \quad \forall t = 1,2,\ldots,T. \end{aligned} \tag{4.7}$$

Though many feasible paths are predefined for any two nodes, only one path needs to be selected. Thus, the following routing constraint should be obeyed:

$$\begin{aligned} & f \in F(m,n) \qquad D_i(m,n) = D(m,n) \\ & \forall i \in Route^f(m,n) \quad \forall m = 1,2,\ldots,N \quad \forall n = 1,2,\ldots,N. \end{aligned} \tag{4.8}$$

To satisfy all traffic requests, the following traffic request constraint should be considered:

$$\begin{aligned} & r_i \times |P_i| \geq D_i = \sum_{m=1}^{N} \sum_{n=1}^{N} D_i(m,n) \\ & \forall i = 1,2,\ldots,L \quad \forall m = 1,2,\ldots,N \quad \forall n = 1,2,\ldots,N \end{aligned} \tag{4.9}$$

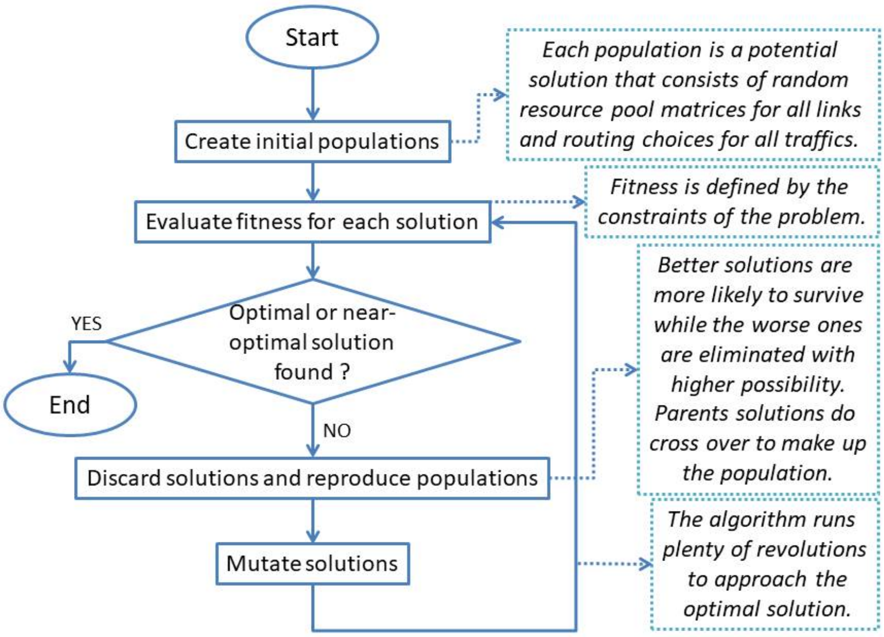

Fig. 4.5 Genetic algorithm flowchart

To this end, an optimisation problem, of joint routing and resource allocation for backhaul networks, has been formulated. In summary, the objective in (4.4) needs optimisation subject to the constraints in (4.5)-(4.9).

4.3.2 Genetic Algorithm (GA) Solution

As the resource allocation variable, P, must be an integer, the optimisation problem identified here is inherently NP-hard [110]. For integer programming problems, GA performs well due to its population-based features, which can avoid local optimal solutions for the most part [111]. Even though no optimal solution is guaranteed, GA can be executed for sufficient time so that it approaches the optimal one. Here we employ GA to obtain a near-optimal solution.

In the following paragraph, we develop a GA, of which the flowchart is shown in Fig. 4.5, according to the problem we have identified, step by step. Then, the complexity of the resultant algorithm is analysed.

Initialisation

The population consists of H individuals, which are different solutions to the problem. Each individual consists of $L+M$ chromosomes, which are $P_i(w,t)$ of L links and routing choices of M traffic requests.

Selection

For each individual, we need to evaluate the solution using our objective function and the constraints (which are represented as 'fitness' in GA). The fitness, which is described below, should be maximised:

$$Fitness = R_S + \gamma \times (C_1 + C_2 + C_3)$$

$$C_1 = \sum_{i=1}^{L} \sum_{\forall j \in A_i} \sum_{w=1}^{W} \sum_{t=1}^{T} min(0, 1 - P_i(w,t) - P_j(w,t))$$

$$C_2 = \sum_{i=1}^{L} \sum_{w=1}^{W} \sum_{t=1}^{T} \left(P_i(w,t) - max(0, 1 - \sum_{\forall j \in A_i} P_j(w,t)) \right)$$

$$C_3 = \sum_{i=1}^{L} min(0, D_i - r_i \times \sum_{w=1}^{W} \sum_{t=1}^{T} P_i(w,t)))$$

where, γ is a penalty parameter intended to balance the fitness value of the constraints. C_1, C_2 and C_3 are derived from link interference, achievable throughput and traffic request constraints respectively.

After the fitnesses are calculated, selection is performed. Better individuals have a greater chance of survival and the worse ones are more likely to be eliminated. In this step, half of the population are selected as survivors.

Crossover

Two survivors are randomly chosen as parents to swap a random number of chromosomes, in corresponding locations, to generate two children. This mating procedure is repeated until the total population reaches H.

Mutation

A small number of individuals are randomly chosen to change the value of one bit in one chromosome. This step is aimed at avoiding locally optimal solutions.

Evolution

Steps 2-4 are repeated I times in order to evolve the solutions generation by generation.

Complexity Analysis

The complexity of the GA is analysed as follows. With a population size variable of H, a link number variable of L and the total traffic requests variable, M, the complexity of the initialisation in step 1 is $\mathcal{O}[H(L+M)]$. Re step 2, the complexity of the fitness calculation is $\mathcal{O}[H(L+M)]$ and of the selection process is $\mathcal{O}(H)$. For step 3, the complexity is $\mathcal{O}(\frac{1}{4}H)$. The complexity of mutation (in step 4) is negligible and can be ignored. The GA will run I evolution rounds in step 5. Thus, the total complexity of the GA is $\mathcal{O}[H(L+M)]+I\times[\mathcal{O}[H(L+M)]+\mathcal{O}(H)+\mathcal{O}(\frac{1}{4}H)]=\mathcal{O}[IH(L+M)]$.

4.4 Digraph-based Joint Routing and Resource Allocation

In reality, the traffic requests in the network may change frequently. In order to adapt to dynamic traffic requests, a relatively fast algorithm is required to solve the problem. A particularly difficult aspect of this, from a computational point of view, is that the system may not be able to redo all the routing and resource allocation assignments whenever a traffic variation occurs. Thus, although the optimal solution is only achievable via a global process, the algorithm should be able to work well on a single traffic request basis. For one single traffic request, we try to make the routing and resource allocation decisions in terms of benefitting the whole network throughput – in order to approach the optimal solution. Because of their low computational complexity, greedy algorithms based on the interference digraph are proposed in this section. The development of these algorithms is described in detail as follows.

4.4.1 Interference Digraph Construction

The network topology can be abstracted into an interference digraph $G=(V,E)$. V is a vertex set whose elements are links in the network, and where $l_i \in V$. The edge set, E, consists of directed edges which represent potential occurrences of interference between links. For instance, edge $e(l_i,l_j)$ indicates that link l_j is in the interference range of l_i. Fig. 4.6 shows an example of the transforming of a topology digraph into an interference digraph.

In graph theory, for any vertex, the number of head ends adjacent to the vertex is defined as the in-degree of the vertex, which is denoted by $deg^{-}(l_i)$. This represents the number of links that may cause interference to link l_i. For instance, as shown in Fig. 4.6, the in-degree of $link12$ $deg^{-}(link12)$ is equal to 1. Likewise, the number of tail ends adjacent to the vertex is defined as the out-degree of the vertex, which in this case indicates the number of links

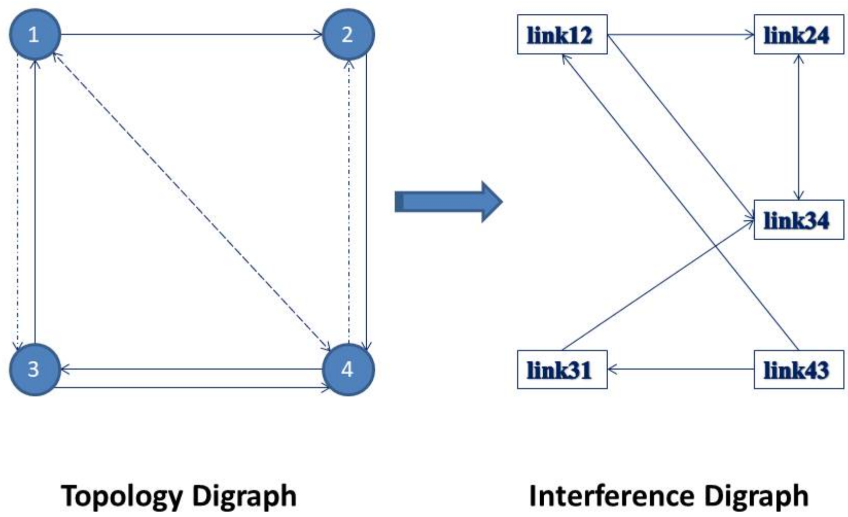

Fig. 4.6 Interference digraph construction

that may receive interference from link l_i; this is denoted by $deg^+(l_i)$. An example is the out-degree of $link12$ in Fig. 4.6 $deg^+(link12) = 2$.

4.4.2 Problem Analysis

In terms of improving the system throughput, the system spectral efficiency of the network, which is related to the degrees of the interference digraph, is a key factor. Besides, to increase the system throughput, it is helpful if the network can contain more traffic, and here load balancing plays a vital role. Detailed analysis will be given in regard to these two aspects, based on the interference digraph across two scenarios.

LFM-Based Backhaul Network Scenario

For backhaul networks operated via low-frequency microwave, the link spectral efficiency among various links can be ignored – as discussed before. Thus, to avoid the complexity of considering different link transmission rates, we treat all of these as the same in this scenario.

In the interference digraph, the degrees of the links characterise, in a clear fashion, the potential interference between links. The out-degree, $deg^+(l_i)$, relates the resource allocation of link l_i to the network throughput, at least to a certain extent. For instance, if $deg^+(l_i) \cdot P_i$ is larger, the available resources for links in interference set B_i is smaller and so is the network throughput. Though the above is not guaranteed to be always true in the global perspective, it is reasonably helpful to approach the local optimal solution by reducing $deg^+(l_i) \cdot P_i$.

From a load balancing point of view, when choosing a path for a traffic request, we also need to consider the available resource blocks for the link l_i which are left; this can be related to the link in-degree $deg^-(l_i)$. For example, the larger $deg^-(l_i)$ is, the larger the number of links that interfere with link l_i. If those links are assigned with resource blocks, these resource blocks are not available for link l_i. To characterise the resource block status of link l_i quantitatively, we introduce matrix $Q_i(w,t)$. If $Q_i(w,t) = 1$, this indicates that link l_i or links in set A_i have already occupied resource block $b(w,t)$, which is thus not available for link l_i. Otherwise, resource block $b(w,t)$ can still be assigned to link l_i.

Based on the above discussion, we define a cost function for one link in the routing process in the LFM-based backhaul network scenario as follows:

$$cost1(l_i) = \lceil \frac{D(m,n)}{r_i} \rceil \cdot deg^+(l_i) + Q_i. \tag{4.10}$$

Thus the cost function of one path is calculated by:

$$Cost1(Route^f(m,n)) = \sum_{\forall l_i \in Route^f(m,n)} cost1(l_i). \tag{4.11}$$

mmW-Based Backhaul Network Scenario

For backhaul networks operated in the millimetre wavelength, the modulation level determines the spectral efficiency of each link, which further impacts the system spectral efficiency of the network. As opposed to the situation with LFM-based backhaul networks, not only are the degrees of the interference digraph and load balancing vital, but also the modulation level for each link plays a key role in improving the system throughput. Detailed analyses follow.

The link spectral efficiency is the net bitrate divided by the bandwidth in hertz, which is measured in $bit/s/Hz$ [108]. This metric can indicate how much information one resource block can carry. For comparing links for loading the same traffic, the link with the larger link spectral efficiency needs less resource blocks to transmit the same amount of data. Thus, the use of the link with the higher modulation level, M_i, saves more spectral resources. On the premise that total spectral resources are limited, assigning more resource blocks to links with larger link spectral efficiency leads to higher system throughput.

When taking interference into consideration, links with larger link spectral efficiencies leave more available resource blocks for the links in interference set B_i. From another perspective, after routing and resource allocation, $|P_i|$ resource blocks are assigned to link

l_i, which means that these resource blocks cannot be used by any link in interference set B_i. As the link spectral efficiency may vary for those links, the cost for interference avoidance is $|P_i| \cdot \sum_{\forall j \in B_i} r_j$. One advantage of millimetre wavelengths over low-frequency microwave is that the interference set, B_i, is generally smaller due to the narrow antenna beam.

Load balancing is also important in this scenario, because if one link is fully loaded, some other routes become unavailable. In this case, new traffic demand requiring those routes cannot be processed by the network, which brings down the system throughput. Thus, matrix $Q_i(w,t)$ is adopted here to balance the load for different links.

Based on the above discussion, we define a cost function for one link in the routing process in the mmW-based backhaul network scenario:

$$cost2(l_i) = \lceil \frac{D(m,n)}{r_i} \rceil \cdot \sum_{\forall j \in B_i} r_j + Q_i - \omega \cdot M_i. \tag{4.12}$$

where ω is a penalty parameter to balance the value of aforementioned costs.

Thus the cost function of one path is calculated by:

$$Cost2(Route^f(m,n)) = \sum_{\forall l_i \in Route^f(m,n)} cost2(l_i). \tag{4.13}$$

4.4.3 Proposed Algorithm

Taking advantage of the interference digraph analysis, cost functions for routing have been developed for the two scenarios given in the last section. A greedy algorithm that jointly considers both routing and resource allocation will be given in this section.

During the routing process, we have to make sure that all the links in the path are able to load the traffic, which means that the available resource blocks of each link in the path can satisfy the traffic demand. We develop $Algorithm$ 6 to check if $Route^f(m,n)$ is available.

Considering the requirement for low computational complexity, a digraph-based greedy algorithm (DBGA) is proposed in $Algorithm$ 7, of which the details are given as follows.

This algorithm deals with a single traffic request. For one traffic request, we first evaluate the available paths with the cost function defined in the previous section. Then the path with minimum cost is selected. Note that $Cost$ is written in the algorithm to denote the $Cost1$ in the LFM-based backhaul network scenario and the $Cost2$ in the mmW-based backhaul network scenario. Next, the traffic request is assigned to each link in the chosen path. Lastly, considering avoidance of interference, available resource blocks are assigned to each link. When more than one traffic requests occur, the one with larger rate demand is a priority to process.

Algorithm 6: Availability Check of Feasible Path

Input:
- f^{th} feasible path: $Route^f(m,n)$
- Number of total sub-channels: W
- Number of total time slots: T
- Resource block status $\forall l_i \in Route^f(m,n)$: $Q_i(w,t)$
- Transmission rate per resource block $\forall l_i \in Route^f(m,n)$: r_i

Output:
- Availability of $Route^f(m,n)$

if $\forall l_i \in Route^f(m,n), \quad r_i \cdot (WT - |Q_i|) > D(m,n)$ **then**
  $Route^f(m,n)$ is available.
else
  $Route^f(m,n)$ is not available.

The computational complexity of the proposed algorithm is analysed as follows. In the algorithm, there are two main variables, which are the number of feasible paths, F, and the number of links, L. The number of links in one feasible path is also a variable, and we believe that the number of links in any feasible path is less than L. First, the complexity of the feasible path availability check and cost calculation function is $\mathscr{O}[FL]$. The selection of the minimum cost path is of complexity $\mathscr{O}[F]$. Then, assigning resource blocks to one link requires $\mathscr{O}(1)$, and the update of resource block statuses for set B_i is of complexity $\mathscr{O}[L]$. Thus the total complexity of the algorithm is $\mathscr{O}[FL] + \mathscr{O}[F] + L \cdot (\mathscr{O}(1) + \mathscr{O}[L]) = \mathscr{O}[L(F+L)]$.

4.5 Performance Evaluation

In this section, we will evaluate the proposed algorithm in relation to the LFM-based backhaul network scenario and the mmW-based backhaul network scenario. First, the setup for the simulation will be described. Then, the simulation results will be presented and analysed.

4.5.1 Simulation Setup

The backhaul network is randomly generated in an area of 1000 $m \times$ 1000 m with 20 nodes in total. The transmission and interference ranges are set as 250 m and 400 m, respectively. For the LMF-based antennas, we assume that the angle of radiation is 360°. For the mmW-based antennas, the radiation angle is set as 20° per link direction.

For the resource block settings, we divide the spectral resources into 50 sub-channels. Suppose that one time slot is 5 ms, which results in 20 time slots in one second. In the

Algorithm 7: Digraph-Based Greedy Algorithm for LMF-Based Backhaul Networks

Input:
- Interference graph: $G(V,E)$
- Traffic request: $D(m,n)$
- Feasible paths: $Route(m,n)$
- Number of total sub-channels: W
- Number of total time slots: T
- Interference sets $\forall l_i \in V$: A_i and B_i
- Resource block status $\forall l_i \in V$: $Q_i(w,t)$
- Transmission rate per resource block $\forall l_i \in V$: r_i
- Out degree $\forall l_i \in V$: $deg^+(l_i)$

Output:
- Routing result for traffic request $D(m,n)$: $route(m,n)$
- Resource allocation result $\forall l_i \in route(m,n)$: P_i
- Updated resource block status $\forall l_i \in V$: $Q_i(w,t)$

```
for f = 1 to F do
    Check availability of Route^f(m,n)
    if Route^f(m,n) is available then
        Calculate Cost(Route^f(m,n))
    end
end
Choose the path with minimum Cost as route(m,n)
forall the i ∈ route(m,n) do
    D_i = D(m,n)
    for w = 1 to W do
        for t = 1 to T do
            if D_i > 0  &&  Q_i(w,t) == 0 then
                P_i(w,t) = 1
                Q_i(w,t) = 1
                D_i = D_i − r_i
                forall the l_j ∈ B_i do
                    Q_j(w,t) = 1
                end
            end
        end
    end
end
```

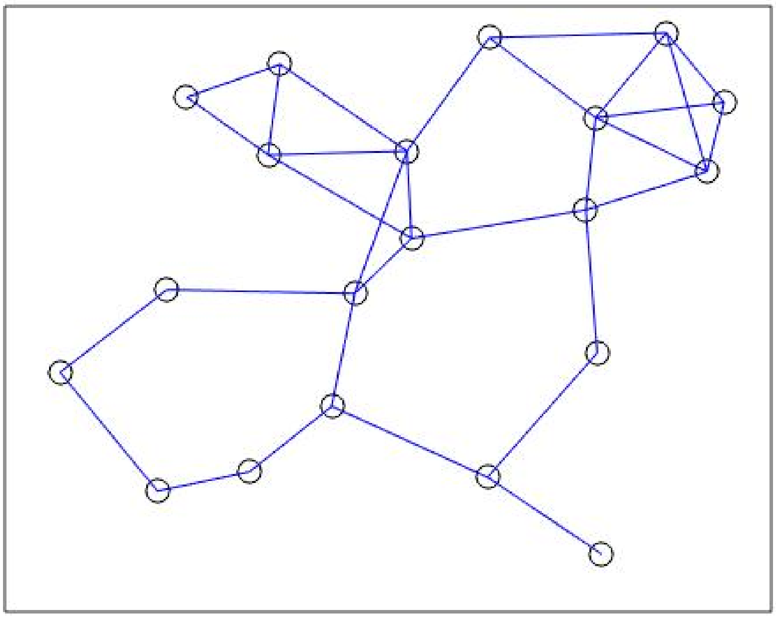

Fig. 4.7 Network topology

LMF-based backhaul network scenario, the link transmission rate is not a variable, and that has a strong impact on the results. So we assume that all links achieve a 0.1 Mb/s rate with one resource block assigned. For the mmW-based backhaul network, we assume that 4 modulation levels are applied to the adaptive modulation technique. So the link transmission rate can be 0.05, 0.1, 0.15 and 0.2 Mb/s per resource block respectively, according to the modulation level.

Four services are considered in the simulations. The request rates of the four services are set to be 0.5, 1, 1.5, 2 Mb/s, respectively. The service-oriented virtual sub-network is constructed by selecting random connected nodes. The traffic requests are randomly generated between two nodes in the corresponding sub-network. Feasible paths are produced using a depth-first search algorithm.

4.5.2 Performance Analysis

The network throughput performance with increasing traffic requests in one topology-fixed network is our primary concern. Fig. 4.7 shows the topology of the network for this simulation. Typical star, chain, tree and ring topologies can all be found within this overall topology, which makes it suitable for our performance evaluation. Two commonly used

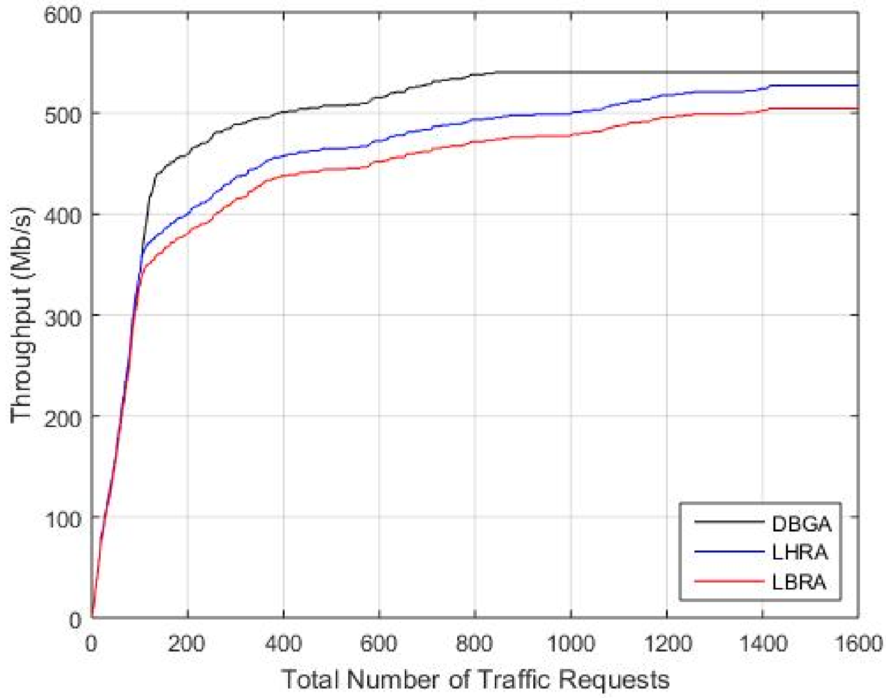

Fig. 4.8 System throughput performance in LFM-based backhaul network

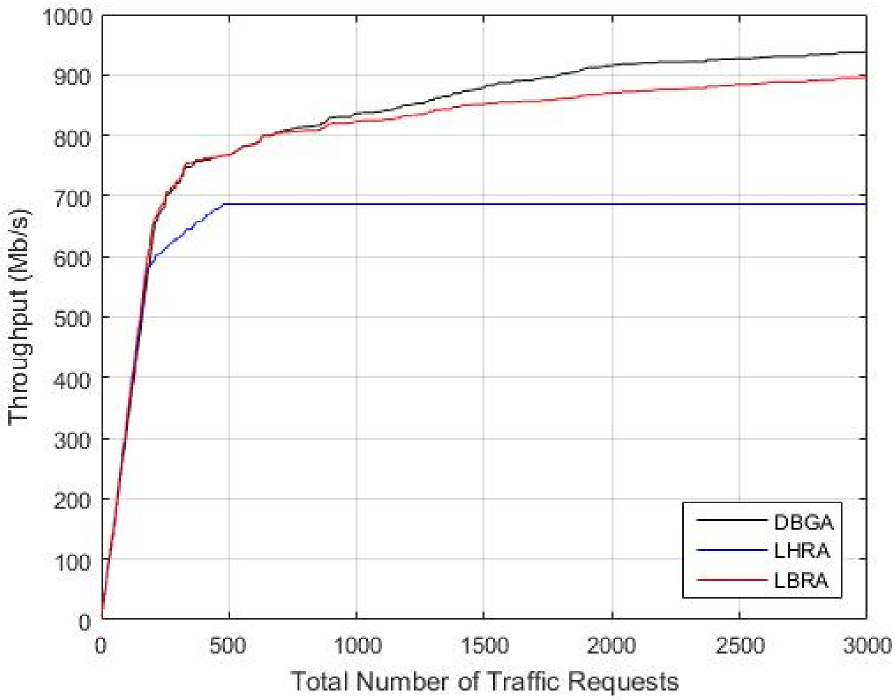

Fig. 4.9 System throughput performance in mmW-based backhaul network

and effective algorithms, which are the least-hop routing algorithm (LHRA) and the load-balancing routing algorithm (LBRA), are compared with the proposed algorithm . To check the network throughput performance of these algorithms, we gradually add new traffic requests to the network. Simulation results for the LFM-based backhaul network and the mmW-based backhaul network are shown in Fig. 4.8 and Fig. 4.9 respectively.

In both figures, the throughput trends are quite similar. In the beginning, all algorithms work well to route the traffic requests. Then the lines separate at around 350 for LFM-based network and 600 for mmW-based network when some items of traffic are failed, and not successfully routed. At last all lines become steady as the network cannot load more traffic. The line for DBGA ends up higher than those for the other algorithms. For LFM-based network, the proposed algorithm outperforms LHRA and LBRA by 10 and 30 Mb/s respectively. For mmW-based network, the proposed algorithm outperforms LHRA and LBRA by 250 and 30 Mb/s respectively In conclusion, our proposed algorithm improves the ability to route traffic requests and increases the network throughput.

Comparing the two figures, it is clear that the system throughput ends up higher in the mmW-based backhaul network scenario. Besides the impact of the larger link transmission rates, another reason is that there is less interference in a mmW-backhaul network, which leads to higher system spectral efficiency. Another interesting difference is that LHRA performs badly in relation to the mmW-based backhaul network. This is mainly caused by the neglect of link transmission rate.

Then we ran a genetic algorithm on the same network setup as in the previous simulations. The result showed that a system throughput of 550.5 Mb/s can be achieved in an LFM-based backhaul network and a throughput of 984 Mb/s can be achieved in a mmW-based backhaul network. The system throughput attained by the use of the proposed algorithm is only slightly smaller than that attained by the GA. To conclude, our algorithm achieves a fair network throughput compared to the near-optimal benchmark.

Lastly, to verify the generality of our algorithm across different network topologies, we ran another 25 sets of simulations with randomly generated networks for both scenarios. We show the average network throughput of the algorithms in Fig. 4.10 and Fig. 4.11. For LFM-based network, the proposed algorithm outperforms LHRA and LBRA by 10 and 25 Mb/s respectively. For mmW-based network, the proposed algorithm outperforms LHRA and LBRA by 210 and 35 Mb/s respectively These results also show that our proposed algorithm is better in terms of increasing network throughput.

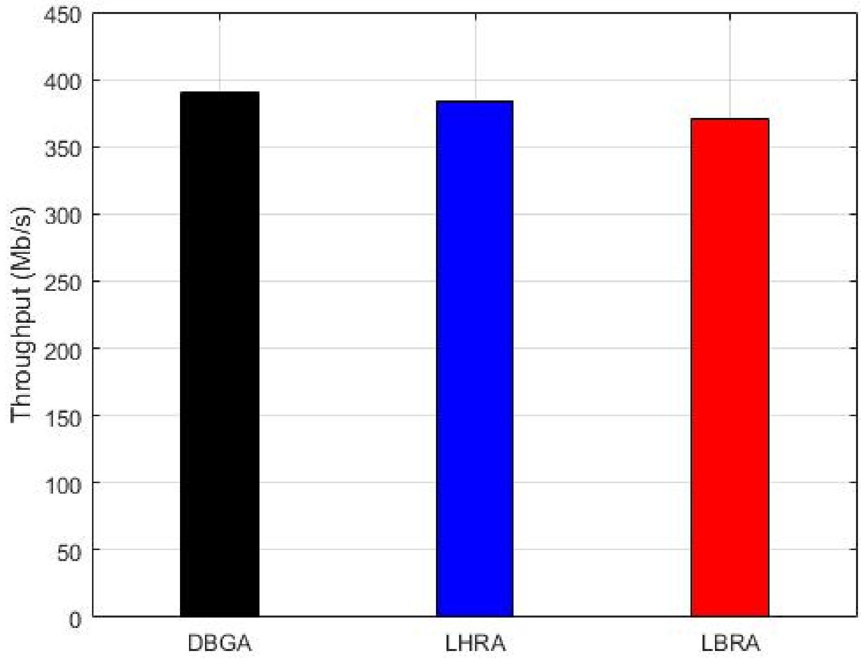

Fig. 4.10 Average system throughput performance in LFM-based backhaul networks

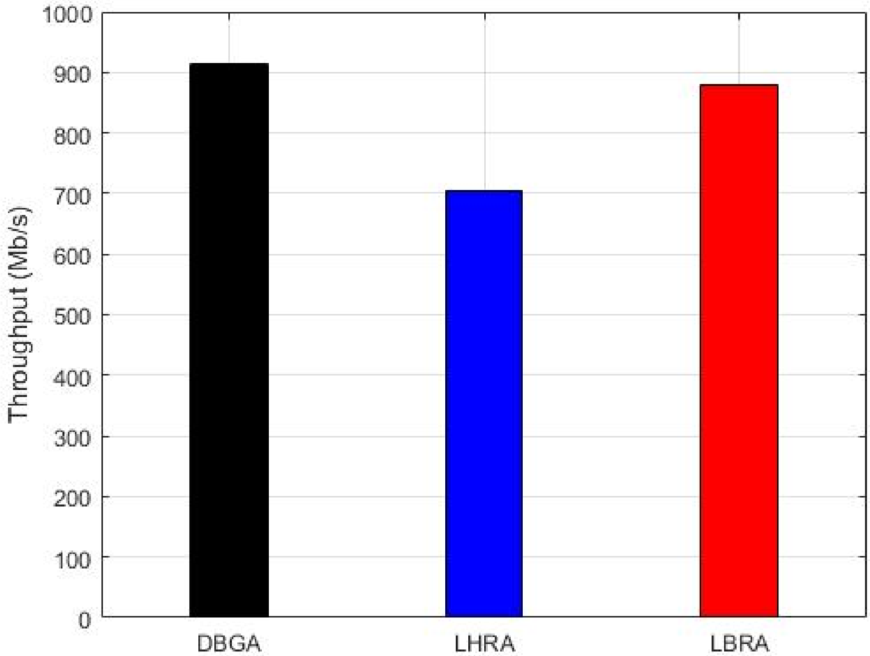

Fig. 4.11 Average system throughput performance in mmW-based backhaul networks

4.6 Conclusions

In this chapter, the joint routing and resource allocation problem in OFDMA-based SDBN is investigated. First, a system model was developed for the routing and resource allocation modules in SDBN. Two scenarios are considered in this work: LFM-based backhaul networks and mmW-based backhaul networks. Detailed mathematical assumptions are made in order to characterise both scenarios. Then, based on the proposed SDBN system model, the joint routing and resource allocation problem is formulated as a system throughput optimisation problem. A genetic algorithm is employed to obtain a near-optimal solution as a benchmark. Next, to reduce the computational complexity, a digraph-based greedy algorithm is proposed – through analysing the relationship between degree of vertices in the constructed interference digraph and system throughput. At last, simulation results have shown that the proposed algorithm outperforms existing algorithms in terms of improving traffic load bearing abilities and increasing system throughput in both scenarios; this also verifies the cost functions that we proposed based on the interference digraph.

Chapter 5

Quality of Service Provisioning in Software-Defined Backhaul Networks

Overview

With traffic load dramatically increasing in backhaul networks, the QoS may be severely impaired due to traffic delays and the limitation of insufficient spectral resources. Furthermore, service differentiation has increased the difficulty of providing QoS – considering diverse QoS requirements. In this chapter, we investigate the QoS provisioning problem in software-defined backhaul network. More specifically, delay and bandwidth are the two main QoS aspects to consider in backhaul networks. First, a delay-aware routing algorithm is proposed to guarantee that the end-to-end delay in traffic meets the delay requirement of the corresponding service. Then, QoS-aware bandwidth allocation problem is modelled as an optimisation problem. Genetic algorithm is employed to obtain a near-optimal solution and a greedy algorithm is developed to suit the highly dynamic traffic demand. Numerical results have shown that the proposed algorithms provide QoS support in terms of delay and throughput.

5.1 Introduction

Due to the popularisation of user equipment and the IoT, recent years have witnessed a dramatic increase in data traffic [1], which has presented great challenges to the mobile networks. With explosive growth in traffic load over wireless networks and the network densification [112], network performance may severely reduce due to the limitation of insufficient spectral resources. In this situation, traffic demand may not be fully supported and the QoS becomes a crucial problem. Furthermore, service differentiation has made the problem even more complicated with all kinds of equipment connected to the mobile network and various service demands from users.

There have been many studies which have addressed the QoS provisioning problem. In [113], a subcarrier and power allocation algorithm based on water-filling was proposed in order to improve fairness in downlink OFDMA systems. In [114], a radio resource management scheme was proposed to provide QoS support in multi-cell heterogeneous OFDMA networks. In [115], QoS was supported by considering channel assignment and traffic routing. In [116], channel allocation, multipath routing, link scheduling and radio assignment were jointly considered in order to satisfy the QoS requirements.

However, the current state-of-the-art QoS support algorithms are not applicable to OFDMA-based backhaul networks due to the different network architectures (e.g., stable mesh topology) and also the complex interference situation. In addition, benefitting from SDNs [5–7] and network function virtualisation [117], QoS provisioning can be virtualised as an independent software-defined component; this enables delicate control.

In this paper, we investigate the QoS provisioning problem in OFDMA-based software-defined backhaul networks. The two main QoS aspects in backhaul networks, delay and bandwidth, are analysed in particular. A delay-aware routing algorithm is proposed in order to guarantee various end-to-end delay requirements. Queueing sequential order adjustment is introduced to improve the routing-successful rate. The QoS-aware bandwidth allocation problem is formulated as a resource block allocation optimisation problem. A genetic algorithm is employed to approach the optimal solution of this NP-hard problem, and a greedy algorithm is proposed to reduce the computational complexity. The numerical results validated the performance of the proposed algorithms.

The remainder of the paper is organised as follows. In Section 5.2, we introduce the QoS-related system model. The delay-aware routing algorithm is proposed in Section 5.3. The QoS-aware bandwidth allocation problem is formulated and solved in Section 5.4. Numerical results obtained from simulations and analyses are presented in Section 5.5. Finally, we conclude this work in Section 5.6.

5.2 System Model

The basic system model that is related to QoS, as shown in Fig. 5.1, is described in this section. Please refer to Section 3.2.1 and 4.2.1 for more details about software-defined backhaul networks.

5.2.1 Software-defined Backhaul Network

A software-defined backhaul network with N nodes and L directional links is considered in this work. Each node is equipped with two OFDMA radios to support simultaneous transmission and reception.

The network is modelled as a directional communication graph $G(V,A)$, where V is the set of nodes and A is the set of directed communication links. For link $l(i,j) \in A$, we define node $i \in V$ as the transmitting node and node $j \in V$ as the receiving node.

A protocol model [96] is applied as the interference model in this work. Two conditions have to be met for a successful transmission from node i to node j: 1) node j is in the transmission range of node i; and 2) any other node which is in interference range of node j is not transmitting the same resource blocks as link $l(i,j)$. The transmission range of each node is R while the interference range is R'.

We assume that the whole spectrum is divided into W sub-channels and that there are T time slots within each one second. Each sub-channel in one time slot is called a resource block. The OFDMA radio can transmit and receive data on any combination of the resource blocks. In this paper, spectral efficiency is irrelevant to our research. We simply assume that the link rate is $P\ Mb/s$ when the link is using the whole spectrum. Thus, if a link is assigned with one resource block per second, its throughput is calculated as $\frac{P}{W \cdot T}\ Mb/s$. For convenience, if not specifically mentioned, the calculations are based on one second time and the traffic demand unit is defined as the amount of data that one resource block can carry.

The total number of services is denoted as G in this work. One particular service is denoted as service g. The quality of service requirements vary across the kinds of service. For example, the download service needs a large bandwidth, but does not ask for low latency. A voice over IP (VoIP) service sets a high demand on latency, but requires little bandwidth. The detailed definitions regarding quality of service will be given in the next section. In this work, the traffic demand of a service, g, is assumed to be a data rate requirement, which is denoted as $D\ Mb/s$, of which the maximum is denoted by D_{max}.

5.2.2 Quality of Service

Generally, the quality of service in wireless networks considers bandwidth, delay, delay variation (jitter), and packet loss parameters [118]. The nodes in a backhaul network are fixed in most cases; this makes the channel state relatively stable. Thus, the issues of quality-of-service in backhaul network are different from those in the access network serving UEs. Packet loss parameters are the last which must be considered in backhaul networks. As the main bottleneck of wireless backhaul networks is that the spectral resource is limited,

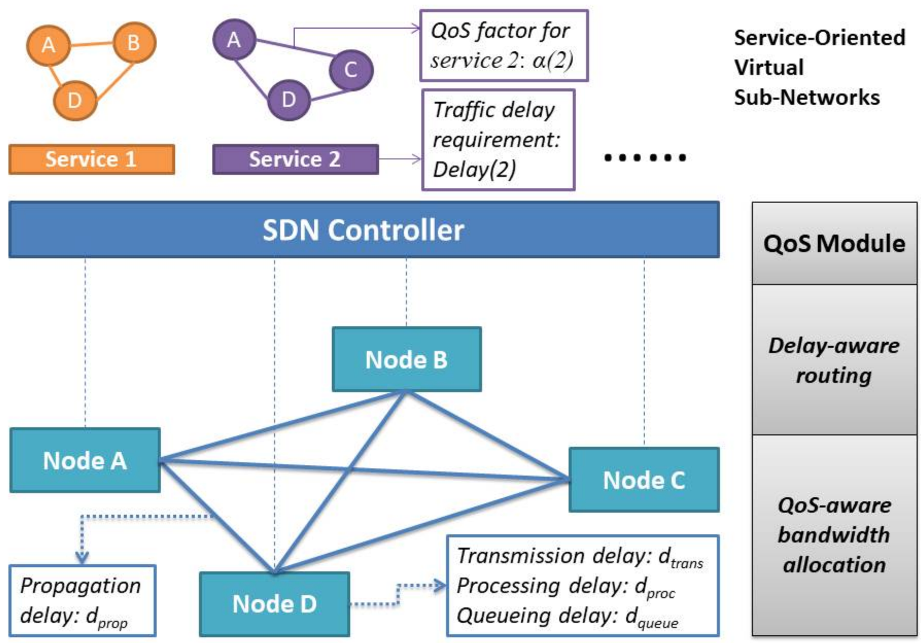

Fig. 5.1 QoS module of software-defined backhaul network

the bandwidth is vital to provide quality of service. Also, to operate various services, the minimum delay for each kind of services should be satisfied. Thus, bandwidth and delay are the two issues which will be considered in terms of providing quality of service in software-defined backhaul networks.

Delay

Delay consists of four elements: transmission delay, propagation delay, processing delay and queueing delay.

The transmission delay is the amount of time required to send all the data into transmission. As we assume that the traffic demand of one particular service is its data rate requirement, the transmission delay for the service, which is denoted by d_{trans}, depends on the packet size. In this work, we assume that for each service, the packet size is a constant value.

In wireless communications, the propagation delay is calculated as $d_{prop} = \frac{Length}{v}$, where $Length$ is the length of the link and v is the signal propagation speed.

In multi-hop transmissions, time is consumed to process and forward the data in relay nodes no matter what kind of relay scheme is applied. The processing delay is denoted as d_{proc} for each relay.

In addition, a relay node can only process a limited number of data packets at a time. If packets arrive faster than a node can process them, the overloaded packets are put into a queue. The queueing delay for traffic m in node i is denoted by $d_{queue}(m,i)$. The queueing delay is decided by the sequential order of the traffic, m, and the congestion level of the node, i. We assume that there is a default sequential order of traffic which is based on the time sequence.

Thus, the total delay for one item of traffic can be represented as:

$$d = d_{trans} + d_{prop} + d_{proc} + d_{queue}. \tag{5.1}$$

For each particular service, g, the traffic delay requirement is treated as a constraint, denoted by $Delay(g)$. The total delay, d, on the selected route, should not exceed the traffic delay requirement, $Delay(g)$.

Bandwidth

Bandwidth is the difference between the upper and lower frequencies in a continuous band of frequencies. In an OFDMA system, bandwidth is refered to as RBs. For a given link with fixed transmission parameters, the throughput of the link is related to the bandwidth of the link. Increasing the bandwidth leads to larger throughput of the link, which means that the link can convey more traffic. Iff the spectral resources of the network are sufficient , the bandwidth that is assigned to each link can achieve the throughput that is required by all the services going through that link.. In this case, the quality of service is not a problem from bandwidth point of view. Thus, we consider the scenario where full traffic demand cannot be satisfied because of limited spectral resources.

We assume that the service subscriber will enjoy the best experience of the service when the full traffic demand $D\ Mb/s$ is achieved. When the traffic data rate drops, the quality of service will also decline. If the data rate decreases beyond a certain point, the service cannot be continued. Thus, we suppose that there is a minimum traffic demand capability which provides quality assured services. To consider the diverse QoS requirements of different services, a QoS factor, $\alpha \in [0,1]$, is employed to characterise the minimum traffic demand of different service types [116]. Thus, the bandwidth requirement of a certain service type is calculated as $\alpha \cdot D$ Mb/s. It is worth noting that the QoS factor for a real-time service (e.g. VoIP) should be larger than that for a best effort service (e.g. a download service).

5.3 Delay-aware Routing

In this section, because it is an important aspect of QoS, the delay inherent in the routing problem is analysed. Then, a delay-aware routing algorithm is proposed based on this discussion.

5.3.1 Problem Analysis

When performing routing for incoming traffic , it is necessary to guarantee that the delay of the traffic does not exceed the delay requirement of the service – which provide acceptable QoS. As introduced in the last section, the end-to-end delay can be estimated with $Equation\ (5.1)$. Thus, the SDN controller can select appropriate routes based on this delay estimation.

As the scenario wherein full traffic demand cannot necessarily be guaranteed is considered in this work, spectral resource availability is not taken into consideration in the routing procedure. However, in order to relieve the stress for the bandwidth allocation, load balancing should be performed during the routing procedure. Besides, an advantage of load balancing is that it can potentially decrease the congestion level at certain nodes, which reduces the queueing delay of traffic.

The SDN controller can make simple decisions regarding allocating available routes that meet the delay requirement of the service in question. However, the scenario that no routes can satisfy the delay requirement should be considered. In this case, the only delay parameter that can be reduced is the queueing delay. Thus, a queueing sequential order adjustment mechanism is proposed. The adjustment is performed by switching the order of the incoming traffic with previous traffic in the queue of the relay node. However, the adjustment should be based on the premise that other routed traffic is not affected such that it would have its delay requirement exceeded. Thus, a matrix, $Surplus$, is proposed to record the gap between the real delay and the delay requirements of the different kinds of traffic. If $Surplus(m) > 0$, this means that the traffic, m, can be placed further back in the queue and that this sequence-switching is permissible.

5.3.2 Proposed Algorithm

Based on the above discussion, a flow chart of a delay-aware routing algorithm is given in Fig. 5.2 and this delay-aware routing algorithm can be decomposed into the following components.

As described in Chapter 4, in the configuration phase of the backhaul network, the nodes and links can be processed, using a depth-first-search algorithm [107], to produce feasible

Algorithm 8: Delay Check of Feasible Routes

Input:
• Feasible path set: $Route$
• Transmission delay of service g: $d_{trans}(g)$
• Propagation delay of link $l(i,j)$: $d_{prop}(i,j)$
• Processing delay of node i: $d_{proc}(i)$
• Queueing delay of traffic m in node i: $d_{queue}(m,i)$
• Delay requirement of service g: $Delay(g)$
Output:
• Delay of $Route^f$: $d(f)$
• Filtered feasible route set: $Route$ • Route with least delay: $Route_{ld}$ • Delay of route $Route_{ld}$: d_{ld}
Initialisation:
• $d_{ld} = \infty$

forall the $Route^f \in Route$ **do**
  forall the *link* $l(i,j)$ *in* $Route^f$ **do**
    $d_{prop} = \sum d_{prop}(i,j)$
  end
  forall the *Transmitting node i in* $Route^f$ **do**
    $d_{proc} = \sum d_{proc}(i)$
    $d_{queue} = \sum d_{queue}(m,i)$
  end
  $d(f) = |Route^f| \cdot d_{trans} + d_{prop} + d_{proc} + d_{queue}$
  if $d(f) < Delay(n)$ **then**
    $Route^f$ meets the delay requirement.
  else
    $Route^f$ does not meet the delay requirement.
    Filter $Route^f$ out.
  if $d(f) < d$ **then**
    $Route_{ld} = Route^f$
    $d_{ld} = d(f)$
end

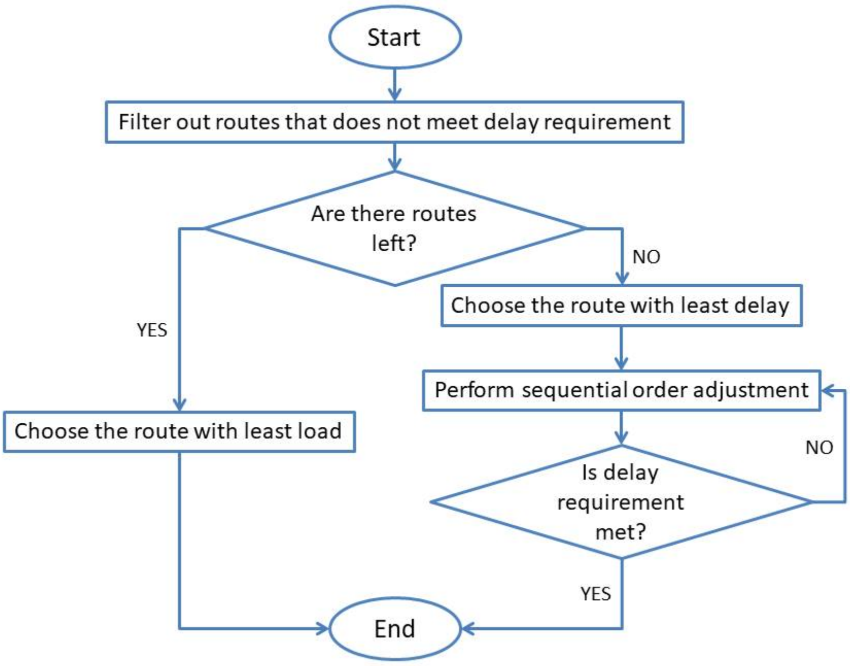

Fig. 5.2 Delay-aware routing flow chart

Algorithm 9: Route Selection Based on Current Load

Input:
- Filtered feasible path set: $Route$
- Delay of $Route^f$: $d(f)$
- Current routed traffic demand on link $l(i,j)$: $load(i,j)$

Output:
- Routing result: $Route(m)$
- Delay surplus: $Surplus(m)$

Initialisation:
- $Load = \infty$

forall the *$Route^f$ in Filtered feasible path set Route* **do**
    forall the *link $l(i,j)$ in $Route^f$* **do**
        $Load^f = \sum load(i,j)$
    end
    if $Load^f < Load$ **then**
        $Route(m) = Route^f$
        $Surplus(m) = d(f)$
        $Load = Load^f$
end

Algorithm 10: Queueing Sequential Order Adjustment

Input:
- Route with least delay: $Route_{ld}$
- Delay of route $Route_{ld}$: d_{ld}
- Delay surplus for traffics: $Surplus$
- Delay gap between two queueing traffic: d_{gap}

Output:
- Updated delay surplus for traffic x: $Surplus(x)$
- Updated delay for the target traffic: d_{ld}

for *traffic x that is affected by the potential adjustment* **do**
    if $Surplus(x) - d_{gap} > 0$ **then**
        $Surplus(x) = Surplus(x) - d_{gap}$
        $d_{ld} = d_{ld} - d_{gap}$
        Move the sequential order of the target traffic in node *i* by one.
    else
        Jump to the next relay node.
end

Algorithm 11: Delay-Aware Routing

---

for *Incoming traffic* **do**
    Run Algorithm 8 to filter feasible routes.
    if *There are feasible routes that meet the delay requirement,* **then**
        Run Algorithm 9 to select the route for the target traffic.
    else
        $Route = Route_{ld}$
        while $d_{ld} > Delay(g)$ **do**
            Run Algorithm 10 at the relay nodes by the order of traffic flow.
            if *no possible adjustment can be made.* **then**
                Routing failed.
                End Loop
        end
end

---

routes between any two nodes. All feasible routes from node i to node j are recorded in set $Route(i, j)$, the size of which is $F(i, j)$. The f^{th} feasible route is denoted as $Route^f(i, j)$, whose elements are link indices.

To guarantee that the service is provided with acceptably limited delays, the total delay of the selected route should be less than the delay requirement of the service. Thus in the first step, among feasible routes, the routes that do not meet the delay requirement should be filtered out. The delay check of feasible routes is shown in $Algorithm$ 8.

If there are several routes available, the one with the least load based on current traffic demand on the links must be selected – for load balancing. The route selection algorithm is shown in $Algorithm$ 9.

If there is no path that meets the delay requirement, the sequential order of the service has to be adjusted. The route that has the least total delay is chosen as the difficulty of the adjustment will be the least. The queueing sequential order adjustment algorithm is shown in $Algorithm$ 10.

Based on $Algorithm$ 8, $Algorithm$ 9, $Algorithm$ 10 and the earlier discussions, the full delay-aware routing algorithm is shown in $Algorithm$ 11. The algorithm deals with new incoming traffic. As the size of the feasible route set, F, the number of links, L, the number of nodes, N and the size of the queue, U, are variables, the complexity of $Algorithm$ 11 is calculated as $\mathscr{O}(FL) + \mathscr{O}(NU) = \mathscr{O}(FL + NU)$.

5.4 QoS-aware Bandwidth Allocation

After all traffic is routed and traffic demands are assigned to each link, the bandwidth allocation can be performed, aiming at QoS provisioning on a single link basis. In this section, the bandwidth allocation problem is formulated into a resource allocation optimisation problem. Two solutions, which are via genetic algorithm and via a greedy algorithm, are proposed.

5.4.1 Problem Formulation

As the bandwidth of OFDMA systems are adjusted by selecting different numbers of resource blocks, the QoS-aware bandwidth allocation problem can be converted into a resource block allocation optimisation problem. We develop this problem as follows.

As the spectral resource is limited, the traffic demands cannot be all satisfied. Thus, we set our objective as maximising the traffic load that is transmitted in one second, which is equivalent to maximising the total number of resource blocks consumed by the network (denoted by Y). Suppose that one resource block in time slot t and sub-channel s assigned to service g in link $l(i,j)$ is denoted as $X(i,j,g,t,s)$. Then the objective function becomes:

$$Maximise\ Y = \sum_{l(i,j)\in A} \sum_{g=1}^{G} \sum_{t=1}^{T} \sum_{s=1}^{W} X(i,j,g,t,s). \tag{5.2}$$

Meanwhile, we have the following resource allocation constraint:

$$\begin{aligned} & X(i,j,g,t,s) \in \{0,1\} \\ & \sum_{g=1}^{G} X(i,j,g,t,s) \leq 1 \\ & \forall l(i,j) \in A \quad \forall t = 1,2,\ldots,T \\ & \forall s = 1,2,\ldots,W \quad \forall g = 1,2,\ldots,G. \end{aligned} \tag{5.3}$$

The link interference constraint is needed to avoid interference between links:

$$\begin{aligned} & \sum_{g=1}^{G} X(i,j,g,t,s) + \sum_{g=1}^{G} X(p,q,g,t,s) \leq 1 \\ & \forall l(i,j) \in A \quad \forall l(p,q) \in I_{l(i,j)} \quad \forall t = 1,2,\ldots,T \\ & \forall s = 1,2,\ldots,W \quad \forall g = 1,2,\ldots,G \end{aligned} \tag{5.4}$$

where $I_{l(i,j)}$ is the interference set of link $l(i,j)$.

To support QoS, the bandwidth requirement constraint is defined as follows:

$$\begin{aligned} & \sum_{t=1}^{T} \sum_{s=1}^{W} X(i,j,g,t,s) \geq \alpha(n) \cdot D(i,j,g) \\ & \forall l(i,j) \in A \quad \forall t = 1,2,\ldots,T \quad \forall s = 1,2,\ldots,W \\ & \forall g = 1,2,\ldots,G \quad \alpha(g) \in [0,1] \quad D(i,j,g) \in [0, D_{max}] \end{aligned} \tag{5.5}$$

where $\alpha(g)$ represents the QoS factor of service g and $D(i,j,g)$ denotes the traffic demand of service g in link $l(i,j)$.

Considering the limited spectral resources, the assigned resource blocks should not exceed the traffic demand. Hence, we have the following limited spectral resource constraint:

$$\begin{aligned}&\sum_{t=1}^{T}\sum_{s=1}^{W} X(i,j,g,t,s) \leq D(i,j,g)\\ &\forall l(i,j) \in A \quad \forall t = 1,2,\ldots,T \quad \forall s = 1,2,\ldots,W\\ &\forall g = 1,2,\ldots,G \quad D(i,j,g) \in [0, D_{max}]\end{aligned} \tag{5.6}$$

To this end, an optimisation problem of resource block allocation has been formulated. In summary, the objective in (5.2) needs optimisation subject to constraints in (5.3-5.6).

5.4.2 Solutions of the Optimisation Problem

In this section, two solutions will be introduced to solve the optimisation problem. First, a genetic algorithm is employed to obtain a near-optimal solution. Second, and to find a solution with less computational complexity, a greedy algorithm is proposed which is appropriate to dynamic traffic.

Genetic Algorithm

As the resource allocation variable, X, must be an integer, the optimisation problem which has been identified is inherently NP-hard [110]. For integer programming problems, a genetic algorithm will often perform well due to its population-based features, which can avoid local optimal solutions for the most part. Although no optimal solution is guaranteed, a genetic algorithm may usually be executed until such time as an optimal solution is approached. Here we employ a GA in order to obtain a near-optimal solution.

In the following paragraph, we develop a GA based on the formulated problem – step by step.

Initialisation:

The population consists of H individuals, which are different solutions to the problem. Each individual consists of L chromosomes, which are $X(i,j,g,t,s)$ of all links.

Selection:

For each individual, we need to evaluate the solution using our objective function and the constraints (these are termed fitness in GA). The fitness, which is described below, should be maximised:

$$Fitness = Y + P \times (C_1 + C_2 + C_3)$$

$$\begin{aligned}C_1 = &\sum_{l(i,j)\in A}\ \sum_{l(p,q)\in I_{l(i,j)}}\ \sum_{g=1}^{G}\sum_{t=1}^{T}\sum_{s=1}^{W} min\\ &\times (0, 1 - X(i,j,g,t,s) - X(p,q,g,t,s))\end{aligned}$$

$$C_2 = \sum_{l(i,j)\in A} \sum_{g=1}^{G} min \times (0, \sum_{t=1}^{T} \sum_{s=1}^{W} X(i,j,g,t,s) - \alpha(g) \cdot D(i,j,g))$$

$$C_3 = \sum_{l(i,j)\in A} \sum_{g=1}^{G} min(0, D(i,j,g) - \sum_{t=1}^{T} \sum_{s=1}^{W} X(i,j,g,t,s))$$

where P is a penalty parameter to balance the fitness value of the constraints. C_1, C_2 and C_3 are derived from the link interference constraint, the QoS requirement constraint and the limited spectral resource constraint, respectively.

After the fitnesses are calculated, selection is performed. Better individuals have a greater chance of survival and worse ones are more likely to be eliminated. In this step, half of the population are selected as survivors.

Crossover:

Two survivors are randomly chosen as parents to swap a random number of chromosomes in order to generate two children. This mating procedure is repeated until the total population reaches H.

Mutation:

A small number of individuals are randomly chosen to have the value of one bit in one of their chromosomes changed. This step is aimed at avoiding locally optimal solutions.

Evolution:

Steps 2-4 are repeated I times to evolve the solutions generation by generation.

The computational complexity of this GA can be analysed as follows. With a population size variable, H, and a link number variable L, the complexity of the initialisation in step 1 is $\mathcal{O}(HL)$. In step 2, the complexity of the fitness calculation is $\mathcal{O}(HL)$ and that of selection is $\mathcal{O}(H)$. For step 3, the complexity is $\mathcal{O}(\frac{1}{4}H)$. The complexity of the mutation in step 4 is negligible and can be ignored. GA will run I evolution rounds in step 5. Thus, the total complexity of the GA is $\mathcal{O}(HL) + I \times [\mathcal{O}(HL) + \mathcal{O}(H) + \mathcal{O}(\frac{1}{4}H)] = \mathcal{O}(IHL)$

Greedy Algorithm

In software-defined backhaul networks, the traffic demand at different links can change frequently. In order to adapt to these dynamic traffic demands, a relatively fast algorithm is needed to solve the problem. Because of its low computational complexity, a greedy algorithm is proposed for this, as shown in *Algorithm* 12, and the details are as follows.

For the purposes of QoS support, the traffic demand is divided into $m = 1, 2$ parts in the first four lines. Next, the algorithm mainly runs in two loops. First, when $m = 1$, the algorithm assigns resource blocks for the first part of the traffic demand to meet the QoS requirements. In the second loop when $m = 2$, the algorithm processes part two of the traffic

Algorithm 12: GR-QABA

```
Input:
• Communication graph G(V,A)
• Interference status table of link l(i,j): Φ(i,j,t,s)
• Traffic demand of service g on link l(i,j): D(i,j,g)
• QoS factor of serviceg: α(g)
• Number of total time slots: T
• Number of total sub-channels: W
• Number of total services: G
Output:
• Resource block allocation results for service g in link l(i,j): X(i,j,g,t,s)
Initialisation:
• Φ(i,j,t,s) = 0

for g = 1 to G do
    D(i,j,1,g) = α(g) · D(i,j,g)
    D(i,j,2,g) = D(i,j,g) − D(i,j,1,g)
end
for m = 1,2 do
    forall the l(i,j) ∈ A do
        for t = 1 to T do
            for s = 1 to W do
                for g = 1 to G do
                    if D(i,j,m,g) > 0&&Φ(i,j,t,s) == 0 then
                        X(i,j,g,t,s) = 1
                        D(i,j,m,g) = D(i,j,m,g) − 1
                    end
                end
            end
        end
        forall the l(p,q) ∈ A do
            Update Φ(i,j,t,s)
        end
    end
end
```

demand as much as possible with limited resource blocks.

For the resource block assignment step of the algorithm, the link interference constraint, which is captured by the interference status table $\Phi(i,j,t,s)$, and traffic load are both considered. More specifically, $\Phi(i,j,t,s) = 1$ means that the resource block of time slot t and sub-channel s is already occupied by an interfering link of $l(i,j)$. The assignment procedure stops when no traffic load is left or no resource block is available. After the assignment of one link, the interference status table is updated to be further used in subsequent assignments.

Lastly, the complexity of $GR-QABA$ may be analysed as follows. In QACWA, the number of links, L, is a variable and the number of resource blocks for each link is constant. So the complexity of assigning resource blocks in regard to one link is $\mathcal{O}(1)$. Updating $\Phi(i,j,t,s)$ of each link is essentially the same case. Thus, the total computational complexity is $L \cdot (\mathcal{O}(1) + \mathcal{O}(1)) = \mathcal{O}(L)$, which is much lower than that of the genetic algorithm.

5.5 Performance Evaluation

In this section, the performance of the proposed algorithms will be evaluated. First, as the function of delay-aware routing is to guarantee that the delays in traffic meet the delay requirement of the services in question, the services that do not demand low latency can work well with or without the delay-aware feature. Thus, delay-aware routing will be simulated in a delay-critical scenario and the performance will be compared with the other two classic routing algorithms. Then, the QoS-aware bandwidth allocation will be simulated in a resource-limited scenario. The proposed GR-QABA will be compared with the benchmark genetic algorithm first and then the performance of the GR-QABA algorithm will be evaluated in terms of traffic demand compared with a state of the art resource allocation algorithm.

5.5.1 Delay-aware Routing

In this section, the simulation parameters used will be introduced and the simulation results relating to the delay-aware routing algorithm (DARA) will be compared with those of the least-hop routing algorithm (LHRA); also, the load-balancing routing algorithm (LBRA) will be presented.

Network topologies are randomly generated in an area of 1000 m $\times$1000 m. Each topology has 20 nodes which are randomly located. The transmission range is set as 250 m. Fig. 5.3 shows one example of a random network topology used in the simulations.

To simplify the simulation, just one kind of service is employed in testing the routing algorithms. We assume that the traffic between two random nodes arrives in a Poisson

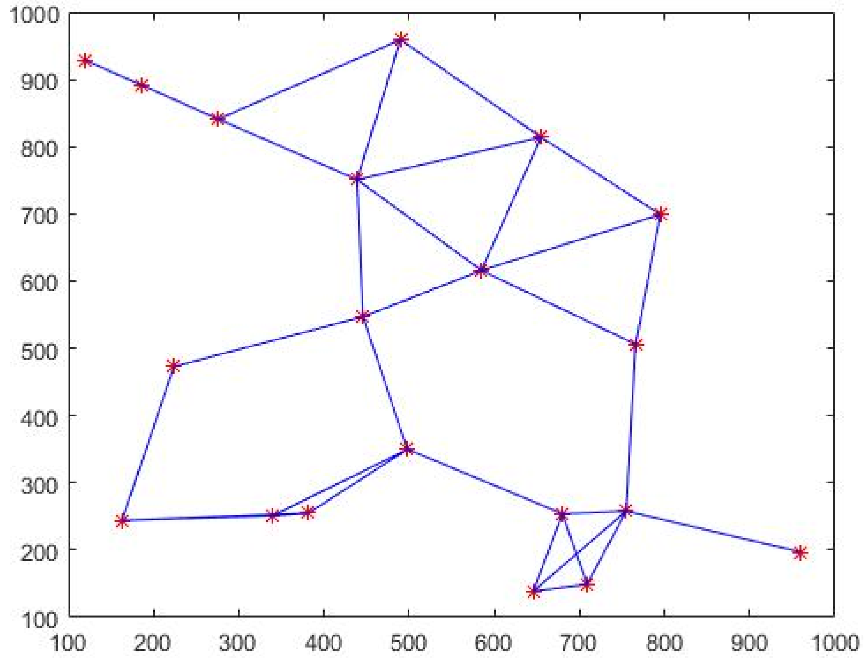

Fig. 5.3 Random network topology

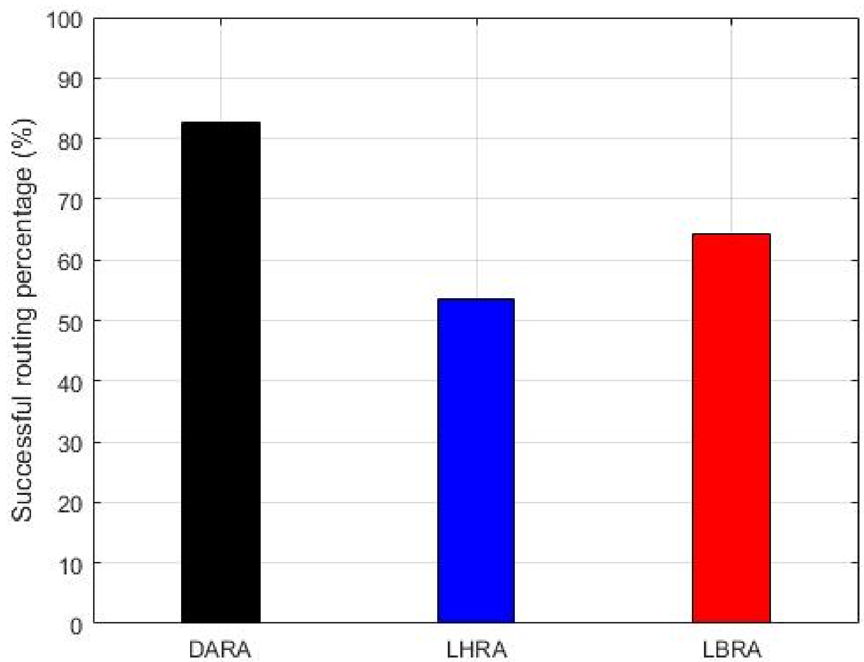

Fig. 5.4 Statistical result of successful routings

process with one request per timeslot on average and each request has a duration that follows an exponential distribution with a mean of ten timeslots. The traffic demand is set as 1 Mb/s. The node can process one traffic flow per time slot and the buffer size is ten. We assume that the transmission delay and the propagation delay take two timeslots and 0.2 timeslots, respectively. We set the delay requirement as 20 timeslots.

100 network topologies were randomly generated. For each network topology, simulations were run for 5000 timeslots. If the traffic delay was within the delay requirement, the corresponding routing was marked as successful. Fig. 5.4 shows the statistical results from the simulations. As the delay requirement is critical, successful routing by LHRA was very low, which was around 53%. The performance of LBRA was better, which was around 64% as it can avoid congestion nodes to some degree. The proposed delay-aware routing algorithm not only keeps the advantages of LBRA but also introduces the queueing sequential order adjustment mechanism, which leads to a high successful routing rate at around 82%. Thus, the delay-aware routing algorithm outperforms the other two algorithms significantly.

5.5.2 QoS-aware Bandwidth Allocation

In this section, we will report on simulations which were performed in a resource-limited scenario to see if QoS could be guaranteed. First, the simulation parameters are presented. Then, the performance of GR-QABA and the benchmark genetic algorithm will be compared. Finally, the performance of GR-QABA, tested in various traffic scenarios, is presented.

Simulation Setup

As reported for the previous section, the network topologies were randomly generated in an area of 1000 m$\times$1000 m. Each topology had 20 nodes. The transmission and interference range are set as 250 m and 500 m respectively. With these parameters, a relatively crowded OFDMA-based network is simulated; this set-up can help in analysing the QoS properties of the proposed algorithms. It is worth noting that the number of links is also random in every simulation.

Suppose that the whole spectral resource is 40 MHz and the link rate is 100 Mb/s when using the whole spectrum. We divide the whole spectrum into 50 sub-channels. Suppose that one time slot is 5 ms and hence there are 20 time slots in one second.

Four services are considered in these simulations. The QoS factors, α, of the services are 0.8, 0.6, 0.3, 0.1, respectively. The traffic demand at each link is generated with delay-aware routing.

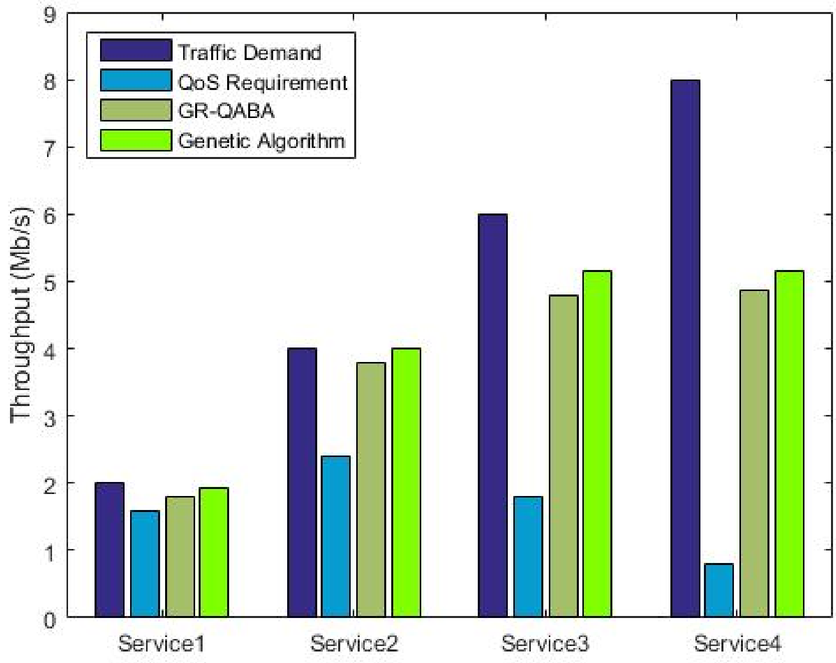

Fig. 5.5 Performance comparison between GR-QABA and Genetic Algorithm

Comparison between genetic algorithm and GR-QABA

Due to the long computing time required for the genetic algorithm, we first compare the results of the genetic algorithm and our proposed greedy algorithm (GR-QABA) across ten simulations. For the genetic algorithm, the population size is set at 40 and the evolution goes 1000 rounds; this almost guarantees an optimal result. The traffic demands of the four services are set to 2, 4, 6 and 8 Mb/s respectively. Fig. 5.5 shows the average results from the genetic algorithm and GR-QABA as related to traffic demand and QoS requirements. As can be seen, both algorithms can guarantee an appropriate QoS for all services. The performance of the proposed greedy algorithm is close to that of genetic algorithm with a difference of 0.4 Mb/s at most; this latter can be regarded as having near-optimal performance. Better yet, the low computational complexity of the greedy algorithm can make up for the minor loss in performance.

GR-QABA Performance Analysis

In this section, we compare our proposed greedy algorithm GR-QABA with GR-SRORA in [119]; both aim at throughput improvement in OFDMA-based networks. We run every

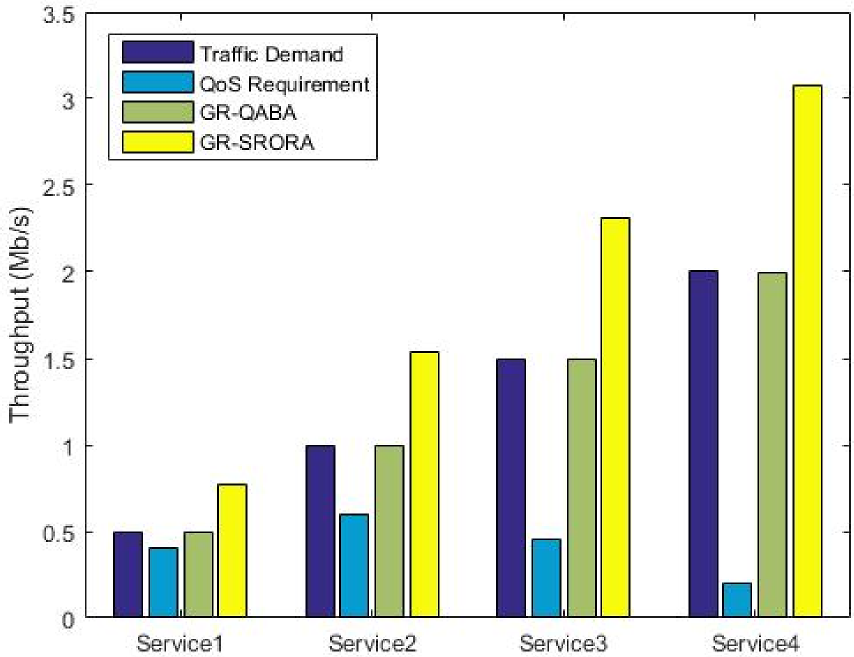

Fig. 5.6 Performance when traffic demand is {0.5, 1, 1.5, 2}Mb/s

simulation 10000 times to widen the network situations examined and obtain an average result which avoids any singularities.

First, we set the traffic demand of the four services to 0.5, 1, 1.5 and 2 Mb/s in order to simulate the scenario that the spectral resource is sufficient for the traffic demand. The results are compared with the traffic demand and QoS requirement given in Fig. 5.6. As can be seen , our algorithm can only satisfy the traffic demand, while the GR-SRORA algorithm can improve the network throughput significantly.

After this we ran three sets of simulations to see how the algorithms perform in the limited-resource scenarios with increasing traffic demand. The traffic demands of the four services were set to be {1, 2, 3, 4}, {2, 4, 6, 8} and {3, 6, 9, 12} Mb/s for the three sets respectively. In Fig. 5.7, it can be seen that both algorithms can achieve the QoS requirement. However, the throughput of GR-QABA is greater than that facilitated by GR-SRORA for the first three services as GR-QABA satisfies the QoS requirement part of traffic demand first. For the last service, GR-SRORA gained higher throughput than GR-QABA as the QoS requirement of service 4 is low. With the increased traffic demand as presented in Fig. 5.8, GR-SRORA cannot satisfy the QoS requirement while GR-QABA can still support the QoS – and the throughput performance is obviously better for the first two services. For the last two services, GR-SRORA can still satisfy the QoS requirements as the traffic demands

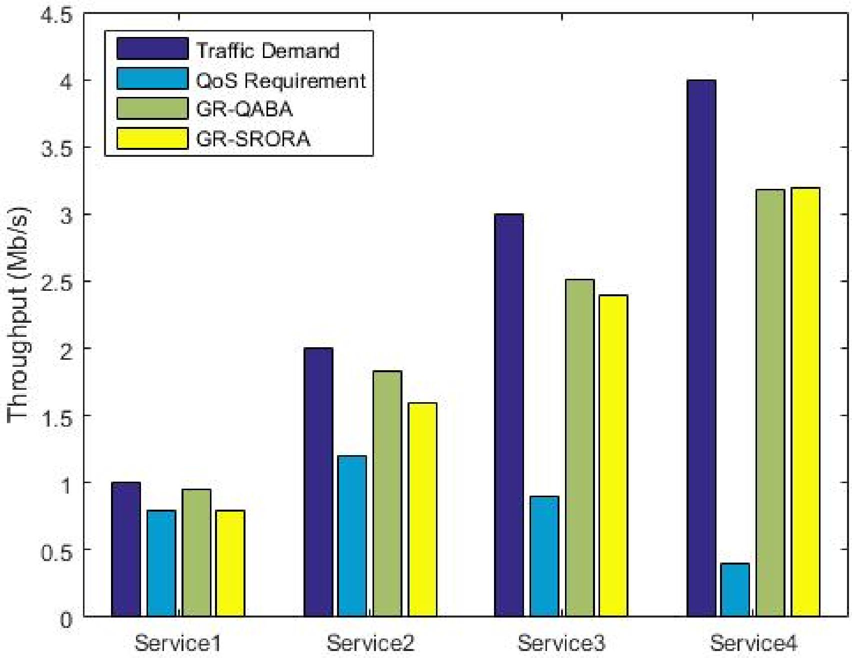

Fig. 5.7 Performance when traffic demand is {1, 2, 3, 4}Mb/s

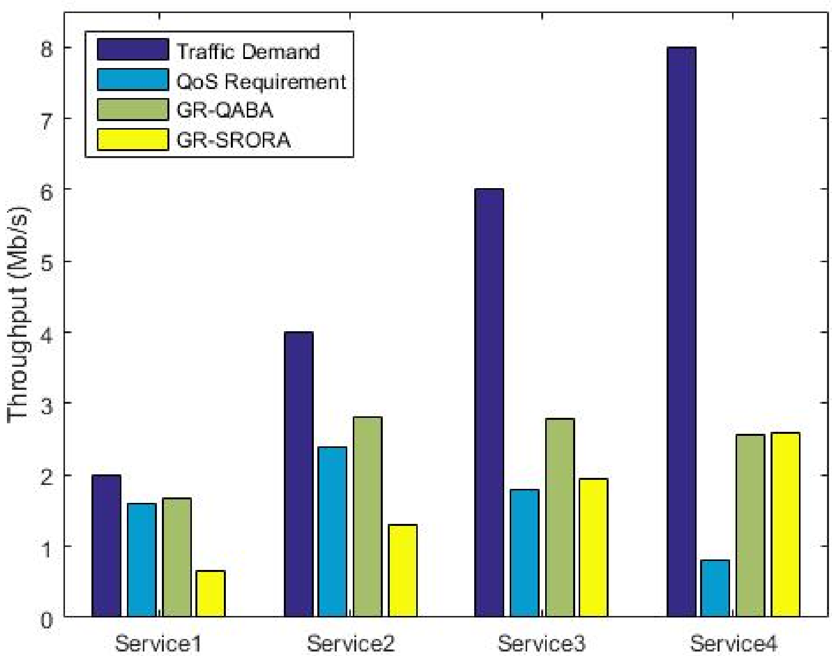

Fig. 5.8 Performance when traffic demand is {2, 4, 6, 8}Mb/s

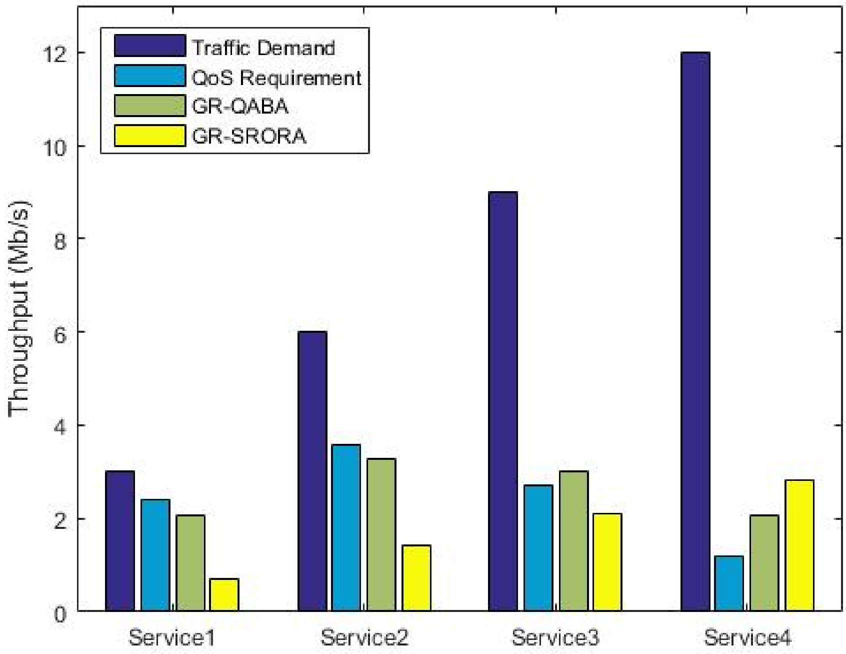

Fig. 5.9 Performance when traffic demand is {3, 6, 9, 12}Mb/s

are high while the QoS requirements are low for these two services. In Fig. 5.9, the traffic demand is so large that neither algorithm can meet the QoS requirements for the first two services. However, GR-QABA performs better than GR-SRORA in terms of QoS support. For service 3, GR-SRORA can not satisfy the QoS requirement while GR-QABA can. For service 4, both algorithms can still satisfy the QoS requirement as the QoS requirement is low. From the results of the above simulations, we can conclude that GR-QABA is suitable for OFDMA-based backhaul networks in resource-limited scenarios.

5.6 Conclusions

In this chapter, the QoS provisioning problem in OFDMA-based software-defined backhaul networks is investigated. Two aspects of QoS are considered, which are delay and bandwidth. To guarantee that data traffic is transmitted within the delay requirement of the service, a delay-aware routing algorithm is proposed. Considering the scenario where spectral resources are insufficient, an optimisation problem reflecting QoS-aware bandwidth allocation is proposed. As a benchmark, a GA is employed to obtain a near-optimal solution. However, a greedy algorithm is proposed in order to reduce the computational complexity. Numerical results show that the proposed algorithms can provide QoS support when dealing with diverse QoS requirements.

Chapter 6

Exploiting Adaptive Modulation in E-band Software-defined Backhaul Network

Overview

To address the bottleneck for traffic scheduling in SDBN, the mmW band, especially the E-band, promises to offer high capacities for 5G backhaul networks with low atmospheric attenuation. As a key issue of E-band propagation, rain attenuation has to be carefully taken into account. Unfortunately, the weather is not controllable in the real world, and backhaul networks must be able to adapt to the various rain events. To address this problem, adaptive modulation technology has been introduced into E-band software-defined backhaul networks for this study. Rain-rate data from rain gauge sensors and an advanced channel model are employed to improve link transmission capacity. The impact of sensing errors and channel model errors on the reliability of the transmission is analysed through numerical results.

6.1 Introduction

As the main bottleneck for traffic scheduling in SDBN is the increasing demand for high capacity backhaul networks, the E-band frequency appears to be a potential candidate for offering quality-guaranteed services addressing this problem [72]. However, with the fact that E-band frequency suffers severely from rain attenuation and backhaul networks must satisfy the traffic demand of access networks, an E-band backhaul network design that is adaptive to rain events and traffic demand may solve the above problems and improve system performances.

In an E-band backhaul solution, fixed LOS links are the main components of the backhaul network. However, the channel conditions vary due to many factors (but mainly rain events). To exploit the full potential of the links, an adaptive modulation technique can be applied to enhance the spectral efficiency under different channel conditions [109]. The key idea of adaptive modulation is to choose the most efficient modulation scheme based on the current channel conditions. Hence, channel estimation and feedback for the control loop are necessary for the adaptive modulation technique.

Most conventional adaptive modulation systems identify the channel state by the channel state information (CSI) . In access networks, CSI plays an important role in improving system performance, especially in coherent MIMO systems [120]. However, the CSI scheme needs a large training sequence to deliver accurate estimations. The delay in CSI feedback may reduce the performance of the system. In fixed E-band backhaul links with fixed nodes in stable environments, the channel states do not change significantly compared to those of channels between UE and BSs. When no rain events take place, the variations in channel conditions are not large enough to impact the choice of modulation schemes. As channel conditions are closely related to the heaviness of the rain, CSI is only useful when rain events are impacting the link. In this case, system resources are wasted when no rain occurs.

With the development of the IoT and Big Data, rain-rate data can be obtained by rain sensors in real-time and shared through the IoT [121]. With an accurate channel model, the proposal that an adaptive modulation scheme could use rain-rate data for channel estimation in a control loop for E-band backhaul links has been made in this work. Taking advantage of the software-defined backhaul network infrastructure [5], the SDN controller is in charge of collecting rain-rate data for all links and for making decisions regarding the modulation scheme that should be applied to each link. When the modulation scheme of one link needs adjustment, the SDN-controller gives orders to the nodes of the link through the control panel.

This chapter is organised as follows: In Section 6.2, some brief statements concerning the proposed adaptive modulation system will be made. Details of the simulation for evaluating the performance of the system will be shown in Section 6.4. In Section 6.5, the results will be analysed. Conclusions will be drawn in Section 6.6.

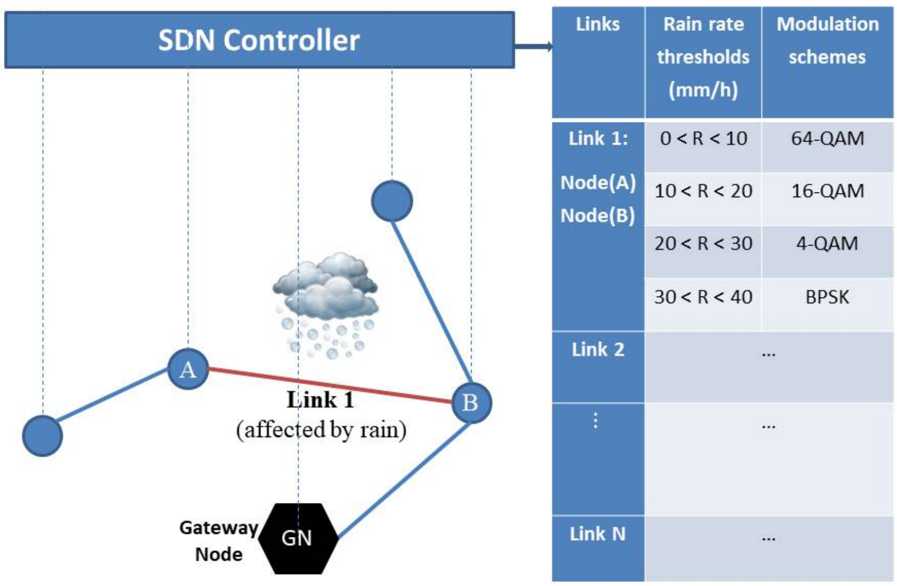

Fig. 6.1 System model

6.2 Adaptive Modulation System for E-band Software-defined Backhual Networks

As can be seen from Fig. 6.1, in the proposed adaptive modulation system, the SDN-controller is responsible for the modulation scheme decisions for all the backhaul links in a particular area, and the data for all the links must be stored and computed in the SDN-controller.

In this section, the development of the thresholds for different modulation schemes will be described. Then, to make sure the adaptive modulation system can adapt to real-world rain events in real-time, the required update frequency of the rain data will be analysed.

6.2.1 Thresholds Development

The modulation schemes and the corresponding rain-rate thresholds should be calculated and recorded when the link is set-up. The right side of Fig. 6.1 shows an example of the data which must be stored in the SDN-controller. The data processing steps are as follows:

First, in order to achieve a certain BER in a link, different SNRs are required when using different modulation schemes. To express the relationship between the BER and the SNR of an M-ary Square QAM , the following equation can be referenced [122]:

$$P_b = \frac{1}{\log_2 \sqrt{M}} \sum_{k=1}^{\log_2 \sqrt{M}} \frac{1}{\sqrt{M}} \sum_{i=0}^{\left(1-2^{-k}\right)\sqrt{M}-1} \{(-1)^{\lfloor \frac{i \cdot 2^{k-1}}{\sqrt{M}} \rfloor} \cdot \left(2^{k-1} - \lfloor \frac{i \cdot 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \rfloor \right) \cdot erfc\lfloor (2i+1)\sqrt{\frac{3\log_2 M \cdot SNR}{2(M-1)}} \rfloor \} \tag{6.1}$$

where M is the order of the QAM and P_b is the BER. In this way, when a link aiming at a certain BER target is demanded, the SNR threshold of each M-ary Square QAM modulation scheme can be obtained.

Next, to derive the relationship between the SNR and the rain rate, a link level analyse is needed:

The received power, P_{RX}, can be obtained by:

$$P_{RX} = P_{TX} + G_{TX} + G_{RX} - Att_{pl} - Att_{rain} \tag{6.2}$$

where P_{TX} denotes the transmitted power, G_{TX} and G_{RX} represent the transmitter and receiver antenna gain respectively, Att_{pl} stands for the free space path loss and Att_{rain} denotes the attenuation caused by rain.

The free space path loss, Att_{pl}, can be obtained using following equation [70]:

$$Att_{pl} = 92.4 + 20\log f_{GHz} + 20\log d_{km} \tag{6.3}$$

where f_{GHz} is the operating frequency and d_{km} is the link length in kilometres.

In ITU-R model [74], Att_{rain} is calculated by the following equation:

$$Att_{rain} = k \cdot R^{\alpha} \tag{6.4}$$

where the coefficients k and α are determined by the operating frequency f_{GHz}, and R is the rain rate.

Thus, the SNR can be obtained using the above equations and some simple measurements made in the link set-up stage to estimate the noise power.

$$SNR = P_{RX}/P_{Noise} \tag{6.5}$$

With the equation between BER and SNR and the approximated relationship between SNR and the rain rate, the final BER equation expressed using the rain rate can be formed as below:

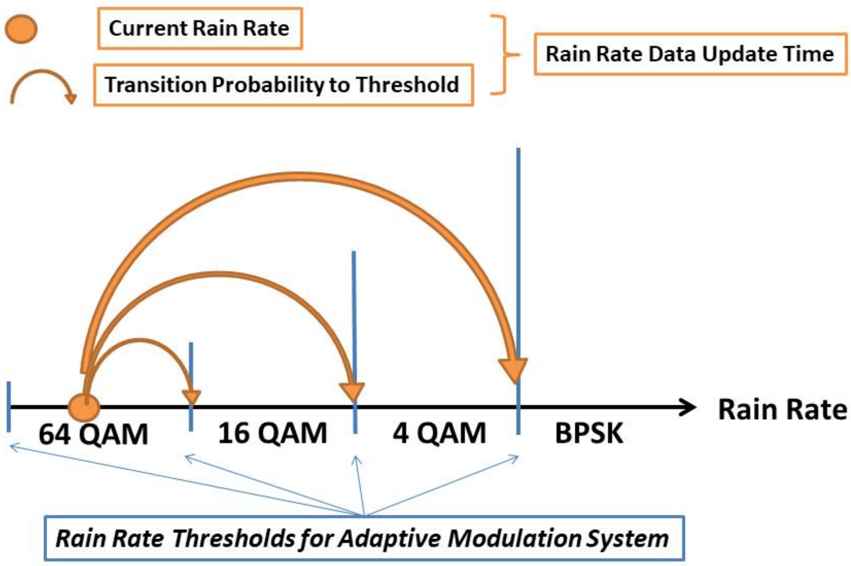

Fig. 6.2 Rain-rate data update time decision

$$P_b = \frac{1}{\log_2 \sqrt{M}} \sum_{k=1}^{\log_2 \sqrt{M}} \frac{1}{\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \{(-1)^{\lfloor \frac{i \cdot 2^{k-1}}{\sqrt{M}} \rfloor} \cdot \left(2^{k-1} - \lfloor \frac{i \cdot 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \rfloor\right) \cdot erfc\lfloor (2i+1) \sqrt{\frac{3 \log_2 M \cdot \frac{(P_{TX}+G_{TX}+G_{RX}-92.4-20\log f_{GHz}-20\log d_{km}-k \cdot R_\alpha)}{P_{Noise}}}{2(M-1)}} \rfloor \} \tag{6.6}$$

Thus, we can process the data with the above equation to obtain the rain-rate threshold for each M-ary QAM scheme to achieve the appropriate BER for each link. An example of the result of this processing is shown in Fig. 6.1.

6.2.2 Update Frequency of Rain Rate Data

When designing the control algorithm for the adaptive modulation system, it was necessary to take into account the dynamic and statistical characteristics of the channel conditions [124]. To ensure the system adapts to real-time channel conditions, the control algorithm of the adaptive modulation has to guarantee that the rain-rate data is updated in time. Thus, the update frequency of the rain-rate data is crucial for the system. In most conventional adaptive modulation systems, the update frequency of the channel state information is fixed. However, if the current state is well within the threshold which is used to adjust the modulation schemes, a frequent update seems redundant. Taking advantage of the characteristics of rain events, the update frequency of rain-rate data is developed as follows.

The rain process can be characterised as a sequence of individual rain events with distinct rain rates. Thus, a low order Markov process can be applied to simulate this random memoryless process [125]. Hence in this work, by employing a first order Markov chain of rain data, the design of the update frequency for the rain-rate data can be proposed. In detail, the current rain rate can be used to determine the next rain-rate data update time with the aid of Markov chain as shown in Fig. 6.2.

The Markov chain transition matrix derived from local rain data should be constructed first. We define the rain-rate state as r and the sampling of the rain process at any particular time, t_n $(n = 1,2,3\dots)$ as X_n, where X_n can be any rain-rate state $r_1, r_2, \dots, r_N$. Then, the probability of $X_n = r_i$ is defined as:

$$p_i(n) = P\{X_n = r_i\} \tag{6.7}$$

The state transition probability of state r_i to state r_j at time t_n is defined as the conditional probability:

$$p_{ij} = P\left\{\frac{X_n = r_j}{X_{n-1} = r_i}\right\} \tag{6.8}$$

where $\sum_{j=1}^{N} p_{ij} = 1$.

With the historic rain-rate data, we generate statistics concerning rain-rate states, r, from time $t_(n-1)$ to time t_n. Finally, we calculate the transition probability p_{ij}, using these statistics. The transition matrix can be formed as follows:

$$M = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \tag{6.9}$$

Based on the transition matrix, the transition probabilities governing transitions from current rain-rate states to the identified rain-rate thresholds for different modulation schemes can be obtained. To guarantee that the data update is performed as needed, the larger the transition probability is, the smaller the update time interval should be. Thus, the maximum of the transition possibilities obtained is chosen as p and the update time interval, T, is determined by:

$$T = \gamma \lfloor \frac{1}{p} \rfloor \tag{6.10}$$

where γ is a factor for adjusting the value based for practical applications – as the slope characteristics of rain in different regions are different.

6.3 Proposed Algorithm for the Proposed Adaptive Modulation System

With the adaptive modulation system developed earlier, the SDN-controller can manage the modulation schemes of each link with the controlling algorithms shown in *Algorithm* 13 and *Algorithm* 14.

Algorithm 13: Adaptive Modulation System with Rain-Rate Data

Input:
- Link l_i of the backhaul network: $l_i \in L$
- Upper threshold of n_{th} modulation scheme for link l_i: $UpperThr_i(n)$, $n \in N$
- Lower threshold of n_{th} modulation scheme for link l_i: $LowerThr_i(n)$, $n \in N$
- Real-time rain rate for link l_i: R_i
- Current modulation scheme for link l_i: MS_i

Output:
- Modulation scheme decision for link l_i: Mod_i

Initialisation:
- $Mod_i = 0$

for $n = 1$ *to* N **do**
  if $LowerThr_i(n) \leqslant R_i < UpperThr_i(n)$ **then**
    $Mod_i = n$
end
if $MS_i \neq Mod_i$ **then**
  Send modulation scheme index Mod_i to link l_i through control panel;
else
  Control panel for adaptive modulation stays silent.

Algorithm 13 deals with single links in the backhaul network. For each link, the algorithm finds out the corresponding modulation scheme for the real-time rain rate. If the modulation decision is the same as the current modulation scheme of the link, the control panel stays silent. Otherwise, the SDN-controller sends the modulation scheme index to the link.

In *Algorithm* 13, the number of modulation schemes, N, is a variable. So the computational complexity of the search for the corresponding thresholds for the collected rain-rate data related to one link is $\mathcal{O}(N)$. The decision for the control panel takes $\mathcal{O}(1)$. Thus, the total complexity of $Algorithm$ 9 is $\mathcal{O}(N)+\mathcal{O}(1)=\mathcal{O}(N)$.

Algorithm 14: Rain-Rate Data Update Algorithm

Input:
- Link l_i of the backhaul network: $l_i \in L$
- Lower threshold of n_{th} modulation scheme for link l_i: $LowerThr_i(n)$, $n \in N$
- Real-time rain rate for link l_i: R_i
- Transition matrix: M
- Current time: T_c

Output:
- Rain-rate data update time for link l_i: T_u

Initialisation:
- $p = 0$

for $n = 1$ *to* N **do**
  Denoting the transition probability from state R_i to $LowerThr_i(n)$ as $p(n)$
  if $p(n) > p$ **then**
    $p = p(n)$
end
Rain-rate data update time: $T_u = T_c + \gamma \lfloor \frac{1}{p} \rfloor$
Update rain-rate data for link l_i at T_u

$Algorithm$ 14 also deals with a single link. The algorithm compares the transition probabilities from a current rain-rate state to the lower threshold for all modulation schemes. Then, the largest one is substituted into Equation 4.10. The system acquires the rain-rate data from the rain sensor after the calculated time interval.

In $Algorithm$ 14, the number of modulation schemes, N, is a variable. So the computational complexity of comparing N transition probabilities is $\mathcal{O}(N)$. The calculation of the rain-rate data update time takes $\mathcal{O}(1)$. Thus, the total complexity of $Algorithm$ 9 is $\mathcal{O}(N)+\mathcal{O}(1)=\mathcal{O}(N)$.

The number of links, N, in the backhaul network is also a variable. When dealing with all links in the system, the algorithms are of $\mathcal{O}(NL)$ time complexity.

6.4 Simulation Setup

In this section, the simulation parameters for the proposed adaptive modulation system are presented.

Transmission Frequency	$f_{GHz} = 76GHz$
ITU-R Rain Attenuation Model Parameters	$k = 1.11, \alpha = 0.71$
Link Length	$d_{km} = 3km$
Transmitter Transmission Power	$P_{TX} = -8dB$
Noise Power	$P_{noise} = -100dB$
Transmitter and Receiver Antenna Gain	$G_{TX} = G_{RX} = 45dBi$

Table 6.1 Parameters for link simulation

Rain-Rate Thresholds (mm/h)	Modulation Schemes
$0 \leqslant R < 8.3$	$1024 - QAM$
$8.3 \leqslant R < 12.5$	$256 - QAM$
$12.5 \leqslant R < 16.8$	$64 - QAM$
$16.8 \leqslant R < 21.1$	$16 - QAM$
$21.1 \leqslant R < 25.1$	$4 - QAM$
$R \geqslant 25.1$	$BPSK$

Table 6.2 Rain rate thresholds for adaptive modulation schemes

6.4.1 E-band Backhaul Link Setup

First, an E-band backhaul transmission link using commercial device parameters is simulated. We assume that the link was set up in Sheffield, which is in the F zone on the rain map [123]. To achieve 99.99% availability, the parameters in Table 6.1 are used.

6.4.2 Adaptive Modulation System Setup

The target BER is set at 10^{-5}. As introduced in Section 6.2, the rain-rate thresholds for different QAMs can be obtained in Table 6.2.

6.4.3 Error Simulation

All the possible inaccuracies, including rain sensor error, propagation model error and other factors impacting the link are summarised as ‘error’ in this simulation. The error is simulated by random numbers following a Gaussian distribution with zero mean. The error is added to the rain rate in the following simulations. Four sets of errors, for which the standard deviations are set to be 0.5, 1, 1.5 and 2.5 respectively, are selected with which to perform the simulation. As the standard deviation of the Gaussian distribution is the main parameter that

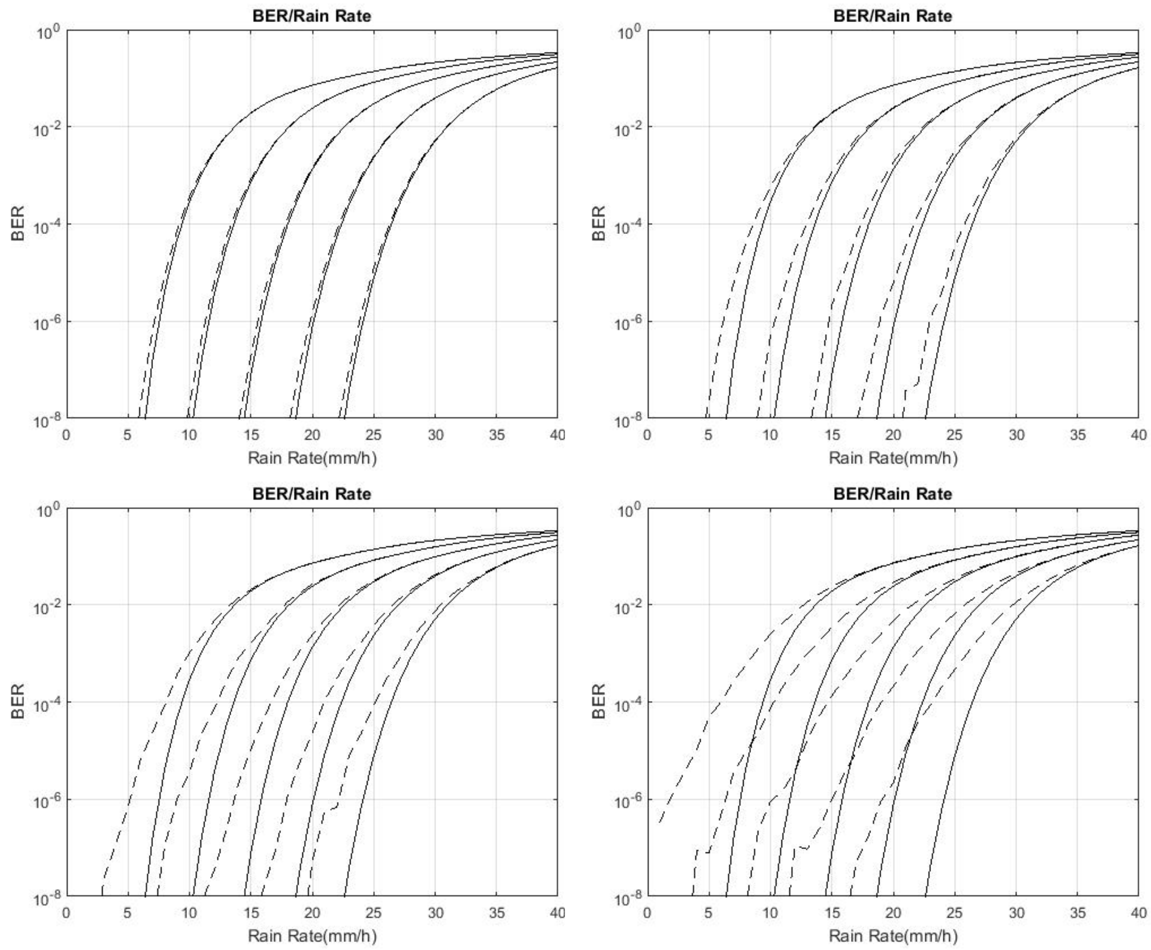

Fig. 6.3 BER performance with error level = 0.5, 1, 1.5 and 2.5 respectively
(5 solid lines from left to right represent ideal performances of 4-QAM, 16-QAM, 64-QAM, 256-QAM and 1024QAM; 5 dashed lines represent performance with errors correspondingly.)

determines the degree of error, it is denoted as ‘error level’ for convenience in the following sections.

6.5 Performance Evaluation

Error may reduce the performance of the system. The target BER may not be achieved when the system is suffering from error.

As can be seen in Fig. 6.3, the BER performance decreases when the error level increases. As the system is designed to achieve a BER of 10^{-5}, the ideal performance is exactly 10^{-5} at the rain-rate thresholds of the system. The BER performance is examined at the rain-rate thresholds when error exists. When the error level is 0.5, the BER is around 2×10^{-5} for

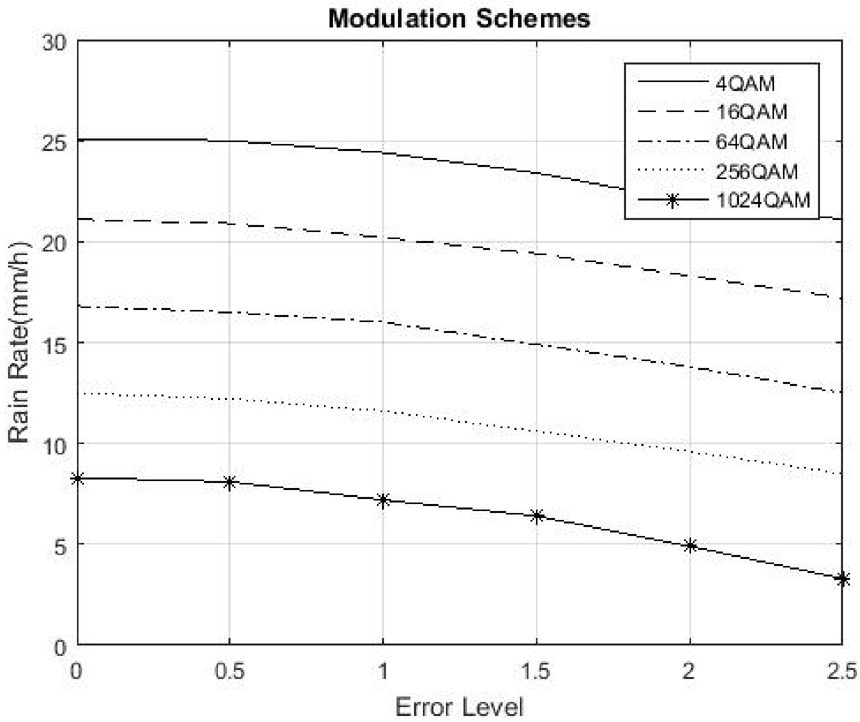

Fig. 6.4 Rain rate thresholds (upper bound) in different error levels

each modulation. As the error level increases to 1, the BER drops to around 8×10^{-5}. Rain sensors in the commercial market may well be subject to this kind of error level. When error is taken into account in relation to the current propagation models, the error level may reach 1.5, in which case the BER may be around only 10^{-4}. When the system is suffering from unexpected situations, e.g. the error level is 2.5, the BER can be as low as 10^{-3}.

To compensate for the reduced BER performance which results from error, the thresholds of the different modulation schemes should be adjusted. In other words, enough of a margin for error should be allowed for when designing the system; this may lead to capacity decrement.

Fig. 6.4 shows the adjusted rain-rate thresholds corresponding to different error levels. The higher the error level is, the larger the margin is, and hence the lower the thresholds are.

The capacity performances corresponding to rain rates are shown in Fig. 6.5. The error level pulls down the system capacity by decreasing the modulation switching thresholds. This suggests that the capacity of the link is only impacted when rain events are occurring. In order to check the impact level, the average capacity of the system is calculated by multiplying the capacity with the corresponding rain-rate probability [58].

As shown in Fig. 6.6, the average capacity drops insignificantly with the increasing of error level. The reason is that the capacity in low rain rates dominates the average capacity

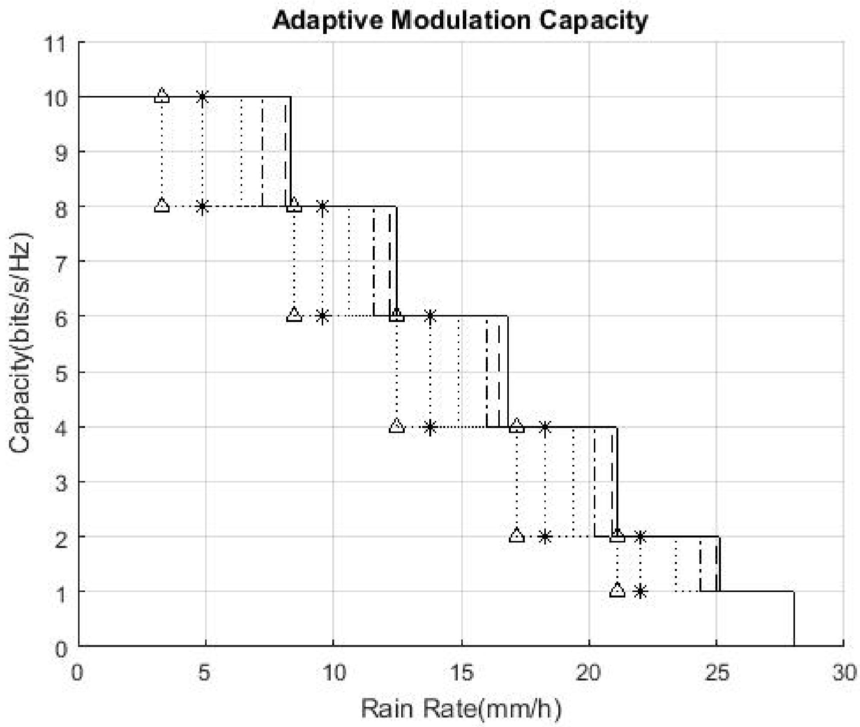

Fig. 6.5 Link capacity versus rain rate with different error levels

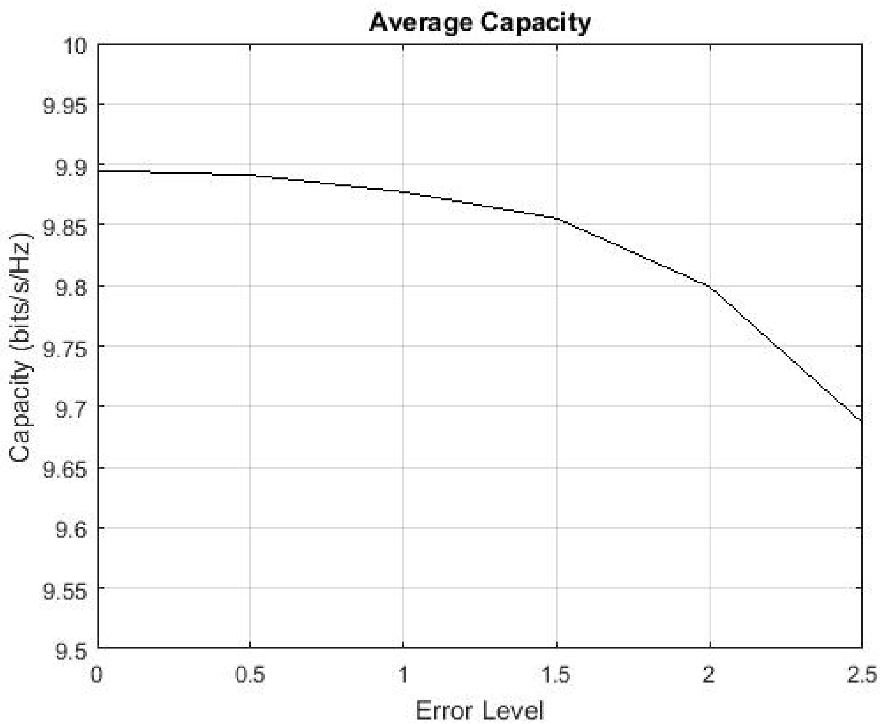

Fig. 6.6 Average link capacity versus error levels

due to the high probability of low rain rates. This suggests that adding a margin for error when dealing with error levels when designing adaptive modulation systems may not impact the system performance in a significant way in moderate rain zones.

6.6 Conclusions

In this work, a conceptual adaptive modulation system based on rain-rate data has been proposed for E-band software-defined backhaul networks. A centralised computing and controlling algorithm was also proposed. By evaluating the BER and capacity performances of the proposed system, the conclusion has been drawn that the proposed system has good potential in terms of improving spectrum efficiency for E-band SDBNs. Numerical results show that accurate channel model and rain sensor data help the E-band SDBN to enhance its transmission capacity.

Due to lack of measurement data, this work is limited to simulations only. Further comparison with measurement data could be undertaken to verify this proposed system. Thus far, the proposed adaptive modulation system only considers link level performance. More work could be focused on the design of adaptive multi-hop E-band backhaul systems based on this work.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this thesis, traffic scheduling was mainly investigated from three aspects. First, virtual network embedding problem was investigated to achieve flexible and dynamic control over various services in the same wireless backhaul infrastructure. Then, joint routing and resource allocation approach was employed to provide interference management and improve the network throughput. At last, the QoS provisioning problem considering delay and throughput metrics was investigated in the scenario where spectral resources can not fulfil the traffic demands.

Focusing on the software-defined backhaul network, the VNE problem was investigated to provide a solution to embed various virtual networks onto the same substrate infrastructure. The SDBN model was proposed to detail the features of backhaul networks. Based on function, nodes were categorised into three types: serving node, gateway node and caching node. Substrate network model was presented based on the role of backhaul networks. The coverage of serving nodes and the caching capacity of caching nodes were modelled as the resources of the nodes. For the model of link resources, resource pool matrix was proposed to capture the interference states of links based on OFDMA systems. Then, the virtual network request was correspondingly described as location recommendation and caching capacity demand for nodes, and resource block request for links. For the serving node mapping procedure, two scenarios were discussed for the serving node mapping where multiple nodes can provide the required coverage. Solutions were given considering backhaul resource consumption and QoS. Virtual gateway node mapping was performed based on the consideration of reducing the cost of backhaul traffic. A caching gain metric was proposed to solve the virtual caching node mapping problem. By evaluating the caching nodes in terms of their abilities to reduce backhaul traffics, the substrate caching node with largest caching gain is mapped in this work. Two weights were defined for the link mapping procedure, which were the spectrum weight that considered system spectral efficiency and the resource weight that reflected the available resource state of substrate links. Based on above discussions, the VNE-SDBN algorithm was proposed to solve the specific VNE problem for SDBNs. Substrate node clustering algorithm was proposed to virtualise the network structure and next hop group algorithm was employed to embed the virtual links. Through extensive simulations, the proposed caching gain metric in relation to backhaul traffic reduction was verified. The real-time and accumulative revenues and VNR acceptance ratio of the proposed algorithm were compared with the other two VNE algorithms. The results showed that our algorithm outperforms the other two algorithms in the wireless backhaul scenario. In this work, caching capacity and location were first introduced to VNE problem. The VNE problem tailored for SDBNs was solved taking advantage of the features of the network architecture.

By employing a joint approach, the traffic scheduling that aimed at improving the network throughput in SDBNs was investigated in terms of routing and resource allocation. Two network scenarios were considered, which were the backhaul networks operated in LFM and mmW. The feasible paths between any two nodes were proposed to be generated and stored in SDN-controllers. Interference model was described and resource pool was proposed to manage the resource assignment and interference management. Link transmission rate was modelled on resource block level. System throughput was given in equations in two forms. The joint routing and resource allocation problem was developed as an optimisation problem. Genetic algorithm was employed to obtain an near-optimal solution. To realize online dynamic scheduling, a low complexity algorithm was proposed based on the interference digraph. The relationship between the indegree and outdegree of the vertices and the network throughput was analysed. Based on the discussion, two cost functions were defined for LFM-based and mmW-based backhaul networks, respectively. The proposed greedy algorithm worked on single traffic basis in two steps. First, the routing procedure was performed by selecting the available path with minimal cost. Then, the resources were assigned according to resource pool to avoid interference to existing traffics. Extensive simulations were conducted in various network topologies and the results showed that the proposed algorithm outperformed the other two classic routing algorithms combining basic resource management schemes in OFDMA-based networks. The performance of the proposed algorithm approached the near-optimal genetic algorithm benchmark and the complexity of the algorithm could make up for the performance drawback.

Based on the SDBN models that were proposed in previous works, QoS provisioning was investigated through routing and resource allocation. In the QoS module, delay and bandwidth were taken into consideration. Different services had specific requirements on delay and throughput. To model the diverse requirements, the delay requirement and QoS factor were employed to characterise the various services. Delay of one traffic in the backhaul network was estimated by the end-to-end delay model, which comprised of transmission, propagation, processing and queueing delays. The delay requirement of one service was regarded as a constraint, which can not be exceeded. Through routing procedure, paths that satisfied the constraint were regarded as candidates. If no path could meet the delay requirement, traffic delay was reduced through queueing sequential order adjustment. Based on above discussion, delay-aware routing algorithm was proposed. For the bandwidth assignment section, the resource limited scenario was investigated. QoS factor that characterised the minimal traffic demand of services was employed and the QoS-aware bandwidth allocation problem was developed as a resource block allocation optimisation problem. Genetic algorithm was developed as a benchmark. A greedy algorithm that aimed to provide online optimisation ability was proposed with the core idea that the assigned resource blocks should satisfy the minimum demand first. In the simulation set-up, delay-critical service was simulated to evaluate the performance of the delay-aware routing. Statistical results showed that the successful routing ratio was greatly improved compared with the other two routing algorithms. The performance of QoS-bandwidth allocation algorithm was tested in resource-limited scenarios. Numerical results showed that the proposed algorithm could satisfy various service QoS requirements with limits.

At last, an adaptive modulation system was proposed for E-band backhaul network to counter the rain attenuation problem. Based on the stable channel condition feature in backhaul network and accurate channel model, rain rate data was proposed to monitor the channel state. A centralised computing and controlling algorithm was also proposed. By evaluating the BER and capacity performances of the proposed system, the conclusion has been drawn that the proposed system had good potential in terms of improving spectrum efficiency for E-band SDBNs. Numerical results showed that accurate channel model and rain sensor data could help the E-band SDBN to enhance its transmission capacity.

7.2 Future Work

In this thesis, traffic scheduling for software-defined backhaul network has been investigated to satisfy various service requirements. Virtual networks have been employed to achieve separate control over various services. Joint routing and resource allocation algorithm has been developed to improve the system throughput. QoS has been provided in terms of delay and throughput for delay-critical services and insufficient spectral resources scenarios. Some of the potential future topics based on these works are summarised as follows.

As can be seen in Chapter 3, the resources that are assigned to one virtual network are regarded as exclusive to the specific virtual network. However, the virtual network may not fully utilize all the resources in certain periods. To further improve the network performance, dynamic resource sharing across various virtual networks may be investigated. Besides, the caching node mapping procedure is based on the assumption that only one caching node is required in the virtual network request. With the development of caching technique, multiple caching locations have been proven to further benefit the backhaul network. Thus, the caching node mapping may be refined by combining the multiple caching locations with effective caching strategies.

As can be seen in Chapter 4, the routing procedure only considers one path from the source to destination. In fact, multi-path has advantages in improving traffic survivability and enhancing load balancing for the network. Thus, path splitting is worth further investigation in the software-defined backhaul network. Besides, routing and resource allocation in the network operated in mmW is investigated with the assumption that adaptive modulation technique is applied. However, the channel condition may vary due to weather. Thus, once the channel condition changes, the SDN controller has to decide whether a reroute should be performed. A monitoring system should be proposed to update the traffic information.

As can be seen in Chapter 5, only delay and bandwidth are taken into account to provide QoS through routing and resource allocation. However, QoS is concerned with many other aspects such as jitter, link reliability and so on. Other approaches like transmission power control and interference cancellation can also impact the QoS. Thus, there is more room for research in QoS provisioning.

As can be seen in Chapter 6, adaptive modulation technique is employed to improve the network spectral efficiency and counter the rain attenuation problem for E-band. There are many other attenuation mitigation techniques (e.g. frequency diversity). Further researches can be conducted by combining these techniques with the backhaul network architecture.

In addition to the extensions based on this thesis, by applying latest technologies to software-defined backhaul network, performances can be further improved. For example, apply big data in SDBN: by analysing network traffic data, the traffic scheduling can be proactive and further improve the network performance. Another example is applying machine learning in the SDN controller: the SDN controller can improve the traffic scheduling by itself through studying patterns of existing algorithms and developing new ones that are more effective for the real-time traffic management. We believe that it is a trend for the network to be proactive and smart.

References

[1] Cisco, “Global mobile data traffic forecast update, 2011-2016,” White Paper, Feb. 2012, pp. 1-29.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash “Internet of things: A survey on enabling technologies, protocols, and applications,” in *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347-2376, 2015.

[3] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, “OFDMA femtocells: A roadmap on interference avoidance,” in *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 41-48, 2009.

[4] NGMN Alliance, “Small cell backhaul requirements”, June 2012, available at *https://www.ngmn.org/uploads/media/NGMN_Whitepaper_Small_Cell_Backhaul_Requirements.pdf*.

[5] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, “A survey of software-defined networking: Past, present, and future of programmable networks,” in *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1617–1634, 3rd Quart., 2014.

[6] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, “ Software-defined networking: A comprehensive survey,” in *Proc. IEEE*, vol. 103, no. 1, pp. 14-76, Jan. 2015.

[7] W. Xia, Y. Wen, C. H. Foh, D. Niyato and H. Xie, “ A survey on software-defined networking,” in *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 27-51, 1st Quart., 2015.

[8] M. Jarschel, T. Zinner, T. Hoßfeld, P. Tran-Gia, and W. Kellerer, “Interfaces, attributes, and use Cases: A compass for SDN,” in *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 210–217, 2014.

[9] Y. Li, and M. Chen, “Software-defined network function virtualization: A survey,” in *IEEE Access*, vol. 3, pp. 2542-2553, 2015.

[10] I. T. Haque, and N. Abu-Ghazaleh, “Wireless software defined networking: A survey and taxonomy,” in *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2713-2737, 2016.

[11] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, “Software defined optical networks (SDONs): A comprehensive survey,” in *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2738-2786, 2016.

[12] H. I. Kobo, A. M. Abu-Mahfouz, and G. P. Hancke, “A survey on software-defined wireless sensor networks: Challenges and design requirements,” in *IEEE Access*, vol. 5, pp. 1872-1899, 2017.

[13] J. H. Cox, J. Chung, S. Donovan, J. Ivey, R. J. Clark, G. Riley, and H. L. Owen, “Advancing software-defined networks: A survey,” in *IEEE Access*, vol. PP, no. 99, pp. 1-1, 2017.

[14] A. Mendiola, J. Astorga, E. Jacob, and M. Higuero, “A survey on the contributions of software-defined networking to traffic engineering,” in *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 918-953, 2017.

[15] S. Bhardwaj, L. Jain, and S. Jain, “Cloud computing: A study of infrastructure as a service (iaas),” in *Int J. Eng. and Information Technology*, vol. 2, pp. 60-63, 2010.

[16] N. M. M. K. Chowdhury, “Network virtualization: State of the art and research challenges,” in *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 20-26, Jul. 2009.

[17] K. Tutschku, T. Zinner, A. Nakao, and P. Tran-Gia, “Network virtualization: Implementation steps towards the future internet,” in *Proc. Workshop on Overlay and Network Virtualization at KiVS*, Kassel, Germany, Mar. 2009.

[18] A. Berl, A. Fischer, and H. de Meer, “Using system virtualization to create virtualized networks,” in *Proc. WowKiVS 2009*, Kassel, Germany, Mar. 2009, vol. 17, EASST.

[19] A. Haider, R. Potter, and A. Nakao, “Challenges in resource allocation in network virtualization,” in *20th ITC Specialist Seminar*, vol. 18, 2009, p. 20.

[20] D. G. Andersen, “Theoretical approaches to node assignment,” Dec. 2002, unpublished Manuscript.

[21] J. Kleinberg, "Aproximation algorithms for disjoint paths problems," Ph.D. dissertation, Massachusetts Institute of Technology, 1996.

[22] S. G. Kolliopoulos and C. Stein, "Improved approximation algorithms for unsplittable flow problems," in *Proc. 38th Annual Symposium on Foundations of Computer Science*, pp. 426-436, 1997.

[23] A. Fischer, J. F. Botero, M. T. Beck, H. de Meer, and X. Hesselbach, "Virtual network embedding: A survey," in *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1888–1906, 4th Quart., 2013.

[24] H. Cao, L. Yang, Z. Liu, and M. Wu, "Exact solutions of VNE: A survey," in *China Communications*, vol. 13, no. 6, pp. 48-62, June 2016.

[25] P. Zhang, H. Yao, and Y. Liu, "Virtual network embedding based on the degree and clustering coefficient information," in *IEEE Access*, vol. 4, pp. 8572 - 8580, 2016.

[26] Z. Heng and Y. Guo, "An exact virtual network embedding algorithm based on integer linear programming for virtual network request with location constraint," in *China Communications*, vol. 13, no. 8, pp. 177-183, Aug. 2016.

[27] Y. Mao, Y. Guo, H. Hu, Z. Wang, and T. Ma, "Sharing based virtual network embedding algorithm with dynamic resource block generation," in *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2126-2129, 2015.

[28] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Efficient virtual network embedding with backtrack avoidance for dynamic wireless networks," in *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2669 - 2683, 2016.

[29] G. Chochlidakis and V. Friderikos, "Mobility aware virtual network embedding," in *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1343-1356, 2017.

[30] O. Soualah, N. Aitsaadi, and I. Fajjari, "A Novel Reactive Survivable Virtual Network Embedding Scheme Based on Game Theory," in *IEEE Transactions on Network and Service Management*, vol. 14, no. 3, pp. 569-585, 2017.

[31] X. Gong, Z. Ning, L. Guo, X. Wei, and Q. Song, "Location-recommendation-aware virtual network embedding in energy-efficient optical-wireless hybrid networks supporting 5G models," in *IEEE Access*, vol. 4, pp. 3065-3075, 2016.

[32] R. Mijumbi, J. Serrat, J. Gorricho, and R. Boutaba, “A path generation approach to embedding of virtual networks,” in *IEEE Transactions on Network and Service Management*, vol. 12, no. 3, pp. 334-348, 2015.

[33] A. Fischer, and H. d. Meer, “Generating virtual network embedding problems with guaranteed solutions,” in *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 504-517, 2016.

[34] L. Gong, H. Jiang, Y. Wang, and Z. Zhu, “Novel location-constrained virtual network embedding LC-VNE algorithms towards integrated node and link mapping,” in *IEEE/ACM Trans. Netw.*, vol. 24, no. 6, pp. 3648-3661, 2016.

[35] C. Yu, W. Hou, Y. Guan, Y. Zong, and P. Guo, “Virtual 5G network embedding in a heterogeneous and multi-domain network infrastructure,” in *China Communications*, vol. 13, no. 10, pp. 29-43, 2016.

[36] F. Esposito, D. D. Paola, and I. Matta, “On distributed virtual network embedding with guarantees,” in *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 569-582, 2016.

[37] I. F. Akyildiz, and X. Wang, “A survey on wireless mesh networks,” in *IEEE Commun. Mag.*, Vol. 43, no. 9, pp. S23-S30, 2005.

[38] R. A. Guerin, A. Orda, and D. Williams, “QoS routing mechanisms and OSPF extensions,” in *Proc. IEEE GLOBECOM 1997*, Vol. 3, pp. 1903-1908, 1997.

[39] Z. Wang and J. Crowcroft, “Quality-of-service routing for supporting multimedia applications,” in *IEEE J. Sel. Areas Commun.*, Vol. 14, no. 7, pp. 1228-1234, 1996.

[40] T. Le, and M. Gerla, “A load balanced social-tie routing strategy for DTNs based on queue length control,” in *Proc. MILCOM 2015*, pp. 383-387, 2015.

[41] J. Yan, M. Zhou, and Z. Ding, “Recent advances in energy-efficient routing protocols for wireless sensor networks: A review,” in *IEEE Access*, Vol. 4, pp. 5673-5686, 2016.

[42] Z. Li, and Y. Wu, “Smooth mobility and link reliability-based optimized link state routing scheme for MANETs,” in *IEEE Commun. Lett.*, Vol. 21, no. 7, pp. 1529-1532, 2017.

[43] C. Hong, A. Pang, and J. C. Wu, “QoS routing and scheduling in TDMA based wireless mesh backhaul networks,” in *Proc. IEEE WCNC 2007*, pp. 3232-3237, 2007.

[44] M. P. Anastasopoulos, P. M. Arapoglou, R. Kannan, and P. G. Cottis, “Adaptive routing strategies in IEEE 802.16 multi-hop wireless backhaul networks based on evolutionary game theory,” in *IEEE J. Sel. Areas Commun.*, Vol. 26, no. 7, pp. 1218-1225, 2008.

[45] Y. Xu, S. Wan, J. Tang, and R. S. Wolff, “Interference aware routing and scheduling in wireless backhaul networks with smart antennas,” in *2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 1-9, 2009.

[46] D. Hui; J. Axnas, “Joint routing and resource allocation for wireless self-backhaul in an indoor ultra-dense network,” in *Proc. IEEE PIMRC 2013*, pp. 3080-3088, 2013.

[47] H. Miao, and M. Faerber, “Self-organized multi-hop millimeter-wave backhaul network: Beam alignment and dynamic routing,” in *Proc. EuCNC 2015*, pp. 275-279, 2015.

[48] J. Lessmann, Y. Cheng, S. Pfeiffer, and X. Costa-Perez, “On the benefits of time-varying routing in realistic mobile backhaul networks,” in *Proc. IEEE ICCW 2015*, pp. 56-61, 2015.

[49] E. Pateromichelakis, M. Shariat, A. U. Quddus, and R. Tafazolli, “Joint routing and scheduling in dense small cell networks using 60 GHz backhaul,” in *Proc. IEEE ICCW 2015*, pp. 2732-2737, 2015.

[50] H. Li, D. Hu, and S. Ci, “iCacheOS: In-RAN caches orchestration strategy through content joint wireless and backhaul routing in small-cell networks,” in *Proc. IEEE GLOBECOM 2015*, pp. 1-7, 2015.

[51] E. Zola, A. J. Kassler, and W. Kim, “Joint user association and energy aware routing for green small cell mmWave backhaul networks,” in *Proc. IEEE WCNC 2017*, pp. 1-6, 2017.

[52] H. Yin and S. Alamouti, “OFDMA: A broadband wireless access technology,” in *IEEE Sarnoff Symposium*, pp. 1-4, 2006.

[53] Q. Zhao, and D. Grace, “OFDMA: A broadband wireless access technology,” in *Proc. IEEE ICCS 2012*, pp. 423-427, 2012.

[54] L. Yen, Y. Dai, and K. Chi, “Resource allocation for multi-channel multi-radio wireless backhaul networks: A game-theoretic approach,” in *Proc. IEEE WCNC 2013*, pp. 481-486, 2013.

[55] Z. Shi, Y. Wang, L. Huang, and T. Wang, "Dynamic resource allocation in mmWave unified access and backhaul network," in *Proc. IEEE PIMRC 2015*, pp. 2260-2264, 2015.

[56] P. Hao, X. Yan, J. Li, Y. R. Li, and H. Wu, "Flexible resource allocation in 5G ultra dense network with self-backhaul," in *Proc. IEEE Globecom Workshops 2015*, pp. 1-6, 2015.

[57] C. Hua, Y. Luo, and H. Liu, "Wireless backhaul resource allocation and user-centric clustering in ultra-dense wireless networks," in *IET Communications*, Vol. 10, no. 15, pp. 1858-1864, 2016.

[58] V. M. Vishnevsky, A. A. Larionov, R. E. Ivanov, and M. Dudin, "Applying graph-theoretic approach for time-frequency resource allocation in 5G MmWave backhaul network," in *2016 Advances in Wireless and Optical Communications*, pp. 221-224, 2016.

[59] I. Loumiotis, P. Kosmides, E. Adamopoulou, K. Demestichas, and M. Theologou, "Dynamic allocation of backhaul resources in converged wireless-optical networks," in *IEEE J. Sel. Areas Commun.*, Vol. 35, no. 2, pp. 280-287, 2017.

[60] W. Xia, J. Zhang, S. Jin, C. Wen, F. Gao, and H. Zhu, "Large system analysis of resource allocation in heterogeneous networks with wireless backhaul," in *IEEE Trans. Commun.*, Vol. PP, no. 99, pp. 1-1, 2017.

[61] ITU Recommendation E.800, "Definitions of terms related to quality of service," 2008.

[62] NGMN Alliance, "Guidelines for LTE backhaul traffic estimation," July 2011, available at $http://www.ngmn.org/uploads/media/NGMN\_Whitepaper\_Guideline\_for\_LTE\_Backhaul\_Traffic\_Estimation.pdf$.

[63] J. W. Guck, A. V. Bemten, M. Reisslein, and W. Kellerer, "Unicast QoS routing algorithms for SDN: A comprehensive survey and performance evaluation," in *IEEE Commun. Surveys Tuts.*, Vol. PP, no. 99, pp. 1-1, 2017.

[64] S. Lee, G. Narlikar, M. Pal, G. Wilfong, and L. Zhang, "Admission control for multihop wireless backhaul networks with QoS support," in *Proc. IEEE WCNC 2006*, Vol. 1, pp. 92-97, 2006.

[65] T. H. Szymanski, "Provisioning backhaul traffic flows in TDMA/OFDMA infrastructure wireless mesh networks with near-perfect QoS," in *2010 IEEE Sarnoff Symposium*, pp. 1-7, 2010.

[66] M. Kretschmer, C. Niephaus, D. Henkel, and G. Ghinea, “QoS-aware wireless backhaul network for rural areas with support for broadcast services in practice,” in *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*, pp. 758-764, 2011.

[67] P. Lin, R. Cheng, and L. Liao, “Performance analysis of two-level QoS scheduler for wireless backhaul networks,” in *IEEE Trans. Veh. Technol.*, Vol. 61, no. 3, pp. 1361-1371, 2012.

[68] L. J. Chaves, I. C. Garcia, and E. R. M. Madeira, “OpenFlow-based mechanisms for QoS in LTE backhaul networks,” in *Proc. IEEE ISCC 2016*, pp. 423-427, 2016.

[69] C. Christodoulou, and G. Ellinas, “Priority scheduling algorithms for QoS support in WDM PON-based mobile backhaul networks,” in *Proc. 2017 ICT*, pp. 1-5, 2017.

[70] H. T. Friis, “A Note on a Simple Transmission Formula,” in *Proceedings of the IRE*, vol.34, no. 5, pp. 254–256, 1946.

[71] FCC Bulletin 70, “Millimeter wave propagation: Spectrum management implications,” July 1997.

[72] P. Wang et al., “Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks,” in *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 168–178, Jan. 2015.

[73] ITU-R P.676-10, “Attenuation by atmospheric gases,” 2013.

[74] ITU-R P.838-3, “Specific attenuation model for rain for use in prediction methods,” 2005.

[75] ITU-R P.840-6, “Attenuation due to clouds and fog,” 2013.

[76] F. Boccardi et al., “Five disruptive technology directions for 5G,” in *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[77] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” in *IEEE Commun. Mag.*, Vol. 49, No. 6, pp. 101-107, Jun. 2011.

[78] J. Wells, “Faster than fiber: the future of multi-Gb/s wireless,” in *IEEE Microwave Mag.*, vol. 10, no. 3, pp. 104–12, May 2009.

[79] D. Bojic et al., "Advanced wireless and optical technologies for small-cell mobile backhaul with dynamic software-defined management," in *IEEE Commun. Mag.*, vol. 51, no. 9, Sept. 2013, pp. 86–93.

[80] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!" in *IEEE Access*, vol. 1, pp. 335-349, 2013.

[81] C. Colombo and M. Cirigliano, "Next-generation access network: A wireless network using E-band radio frequency (71-86GHz) to provide wideband connectivity," in *Bell Labs Technical Journal*, vol. 16, no. 1, pp. 187-206, 2011.

[82] T. Tjelta and T. O. Breivik, "Measured attenuation data and predictions for a gigabit radio link in the 80 GHz band," in *Proc. EuCAP-2009*, Berlin, pp. 657-661, Apr. 2009.

[83] T. S. Rappaport et al., "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," in *IEEE Trans. Antennas Propag.*, Vol. 61, No. 4, pp. 1850-1859, April 2013.

[84] S. Nie et al., "73 GHz millimetre wave indoor measurements for wireless and backhaul communications," in *Proc. IEEE PIMRC*, London, pp. 2429-2433, Sep. 2013.

[85] G. R. MacCartney and T. S. Rappaport, "73 GHz millimeter wave propagation measurements for outdoor urban mobile and backhaul communications in New York city," in *Proc. IEEE ICC*, Sydney, pp. 4862-4867, Jun. 2014.

[86] H. Mehrpouyan et al., "Improving bandwidth efficiency in E-band communication systems," in *IEEE Commun. Mag.*, vol. 52, no. 3, pp. 121–128, Mar. 2014.

[87] S. Rangan, T. S. Rappaport and E. Erkip, "Millimetre-wave cellular wireless networks: Potentials and challenges," in *Proc. IEEE*, vol. 102, no. 3, pp. 366-385, Feb. 2014.

[88] X. Ge, H. Cheng, M. Guizani and T. Han, "5G wireless backhaul networks: Challenges and research advances," in *IEEE Netw.*, vol. 28, no. 6, pp. 6-11, Dec. 2014.

[89] J. Bartelt et al., "Heterogeneous backhaul for cloud-based mobile networks," in *Proc. IEEE VTC Fall*, Las Vegas, pp. 1-5, Sep. 2013.

[90] R.J. Weiler et al., "Enabling 5G backhaul and access with millimetre-waves," in *Proc. EuCNC*, Bologna, pp. 1-5, Jun. 2014.

[91] J.A. Zhang et al., “Low latency integrated point-to-multipoint and E-band point-to-point backhaul for mobile small cells,” in *Proc. IEEE ICC*, Sydney, pp. 592-597, Jun. 2014.

[92] M. Zhang, H. Luo and H. Zhang, “A survey of caching mechanisms in information-centric networking,” in *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1473-1499, 2015.

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

[94] Q. Duan, N. Ansari and M. Toy, “Software-defined network virtualization: an architectural framework for integrating SDN and NFV for service provisioning in future networks,” in *IEEE Netw.*, vol. 30, no. 5, pp. 10-16, Sep./Oct. 2016.

[95] M. Kamel, W. Hamouda and A. Youssef, “Ultra-dense networks: A survey,” in *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522-2545, 2016.

[96] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” in *Wireless Netw.*, vol. 11, no. 4, pp. 471-487, Jul. 2005.

[97] K.-m. Park and C.-k. Kim, “A framework for virtual network embedding in wireless networks,” in *Proc. 4th International Conference on Future Internet Technologies*, ser. CFI ’09. New York, NY, USA: ACM, 2009, pp. 5-7.

[98] C. H. Papadimitriou, “Computational Complexity,” Reading, Mass.: Addison Wesley Longman, 1994. xv, 523 s. ISBN 0-201-53082-1.

[99] L. Gong, Y. Wen, Z. Zhu, and T. Lee, “Toward profit-seeking virtual network embedding algorithm via global resource capacity,” in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, Apr./May 2014, pp. 1-9.

[100] S. Zhang, Z. Qian, S. Guo, and S. Lu, “FELL: A flexible virtual network embedding algorithm with guaranteed load balancing,” in *Proc. IEEE ICC 2011*, Kyoto, Japan, Jun. 2011, pp. 1-5.

[101] R. Wang, D. Butnariu, and J. Rexford, “ OpenFlow-based server load balancing gone wild,” in *Proc. 11th USENIX Conf. Hot-ICE Netw. Serv.*,p. 12, 2011.

[102] P. Pisa et al., “ OpenFlow and Xen-based virtual network migration,” *Communications: Wireless in Developing Countries and Networks of the Future.*, Berlin, Germany: Springer-Verlag, pp. 170–181, 2010.

[103] M. F. Ramdhani, S. N. Hertiana and B. Dirgantara, “ Multipath routing with load balancing and admission control in Software-Defined Networking (SDN),” in *Proc. 4th ICoICT*, pp. 1-6, 2016.

[104] S. Q. Zhang, Q. Zhang, H. Bannazadeh and A. Leon-Garcia, “ Routing algorithms for network function virtualization enabled multicast topology on SDN,” in *IEEE Transactions on Network and Service Management*, pp. 580-594, Dec. 2015.

[105] T. Feng, J. Bi and K. Wang, “ Allocation and scheduling of network resource for multiple control applications in SDN,” in *China Communications*, vol. 12, no. 6, pp. 85-95, Jun. 2015.

[106] Q. Duan, N. Ansari and M. Toy, “Software-defined network virtualization: an architectural framework for integrating SDN and NFV for service provisioning in future networks,” in *IEEE Network*, vol. 30, no. 5, pp. 10-16, Sep./Oct. 2016.

[107] H. S. Stone and P. Sipala, “The average complexity of depth-first search with backtracking and cutoff,” in *IBM Journal of Research and Development*, vol. 30, no. 3, pp. 242-258, May 1986.

[108] S. M. Scott, “The fundamentals of communication,” French Forest, N.S.W.: Prentice Hall-Sprint Print, 1998.

[109] S. T. Chung and A. J. Goldsmith, “Degrees of freedom in adaptive modulation: a unified view,” in *IEEE Trans.on Comm.*, vol. 49, no. 9, pp. 1561-1571, 2001.

[110] C. H. Papadimitrious and K. Steiglitz, “Combinatorial optimization: algorithms and complexity,” in *Mineola, NY: Dover*, ISBN 0486402584, Jan. 1998.

[111] X. Yang, “Nature-inspired metaheuristic algorithms,” Frome: Luniver Press, 2008.

[112] X. Ge, S. Tu, G. Mao, C.-X. Wang and T. Han, “5G ultra-dense cellular networks,” in *IEEE Trans. Wireless Commun.*, Vol. 23, No. 1, Feb. 2016.

[113] T. D. Nguyen and Y. Han, “A proportional fairness algorithm with QoS provision in downlink OFDMA systems,” in *IEEE Commun. Lett.*, vol. 10, issue 11, pp. 760-762, Nov. 2006.

[114] Y. Liu, L. Cuthbert, X. Yang and Y. Wang, "QoS-aware radio resource allocation for multi-cell OFDMA network," in *Proc. IEEE ICCS*, Nov. 2012.

[115] J. Tang, G. Xue, and W. Zhang, "Interference-aware topology control and QoS routing in multi-channel wireless mesh networks," in *Proc. ACM MobiHoc'05*, Urbana-Champaign, IL, May 2005.

[116] Y. Long, H. Li, M. Pan, Y. Fang and Tan F. Wong, "A fair QoS-aware resource-allocation scheme for multiradio multichannel networks," in *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3349-3358, 2013.

[117] Y. Li and M, Chen, "Software-defined network function virtualization: A survey," in *IEEE Access*, vol. 3, pp. 2542-2553, 2015.

[118] ITU-R E.800, "Definitions of terms related to quality of service," 2009.

[119] X. Wang, P. Huang, J. Xie and M. Li, "OFDMA-based channel-width adaptation in wireless mesh networks," in *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 4039-4052, 2014.

[120] K. Ghavami and M. Naraghi-Pour, "MIMO detection with imperfect channel state information using expectation propagation," in *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8129-8138, 2017.

[121] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," in *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22-32, 2014.

[122] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," in *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074-1080, Jul. 2002.

[123] ITU-R P.837-6, "Characteristics of precipitation for propagation modelling," 2012.

[124] B. C. Grémont, A. P. Gallois, M. Filip, and S. D. Bate, "Comparative analysis and performance of two predictive fade detection schemes for Ka band fade countermeasures," in *IEEE J. Select. Areas Comm.*, vol. 17, pp. 180-192, Feb. 1999.

[125] L. Csurgai-Horvath and J. Bito, "Testing the markov property of rain fading on millimetre band terrestrial radio link," in *Proc. EuCAP 2009*, pp. 2901-2905, Berlin, Mar. 2009.