On the optimal operation of wireless networks

by

Yu Jie

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Computer Engineering

Program of Study Committee:

Ahmed E. Kamal, Major Professor
Daji Qiao
Lu Ruan
Arun K. Somani
Zhengdao Wang

Iowa State University

Ames, Iowa

2017

DEDICATION

To my family, my beloved husband Jinchun, and my dear friends for their unconditional love and constant support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my appreciations to those who helped me throughout my study in Iowa State University.

I would like to thank my PhD advisor, Dr. Ahmed E. Kamal for his guidance throughout these past five years. Dr. Kamal is a person with patience, encouragement and respect to his students. He greatly helps me in both technical and personal skill development. Without his guidance and support, I could not complete this thesis and enjoy my research.

I would also like to thank Dr. Daji Qiao, Dr. Zhengdao Wang, Dr. Arun Somani, and Dr Lu Ruan for their feedback and comments about my research works, and with their inspiration, I am able to finish my research work.

I want to thank my lab colleagues, Mirzad Mohandespour, Min Yoon, Mohamed Selim, Abdullah Almasoud, Abdulkadir Celik, Ahmad Alsharoa and Abdullah Alqasir. I enjoy the discussion with them about the research work and life in iastate. The discussion with them makes my research work much more fun. Their help and support encourages me to overcome difficulties.

I also want to thank my dear friends in iastate. Their company enrichs my life in US. I want to thank my parents, with their love, I can face difficulties in life and enjoy my work. Finally, I want to thank my husband, with his love, encouragement, patience and company, I am able to finish this thesis and be brave.

ABSTRACT

With the ever increasing mobile traffic in wireless networks, radio frequency spectrum is becoming limited and overcrowded. To address the radio frequency spectrum scarcity problem, researchers proposed advanced radio technology-Cognitive Radio to make use of the uncommonly used and under-utilized licensed bands to improve overall spectrum efficiency. Mobile service providers also deploy small base stations on the streets, into shopping center and users' households in order to improve spectrum efficiency per area. In this thesis, we study cooperation schemes in cognitive radio networks as well as heterogeneous networks to reuse the existing radio frequency spectrum intelligently and improve network throughput and spectrum efficiency, reduce network power consumption and provide network failure protection capability.

In the first work of the thesis, we study a multicast routing problem in Cognitive Ratio Networks (CRNs). In this work, we propose a new network modeling method, where we model CRNs using a Multi-rate Multilayer Hyper-Graph (MMHG). Given a multicast session of the MMHG, our goal is to find the multicast routing trees that minimize the worst case end-to-end delay, maximize the multicast rate and minimize the number of transmission links used in the multicast tree. We apply two metaheuristic algorithms (Multi-Objective Ant Colony System optimization algorithm (MOACS) and Archived Multi-Objective Simulated Annealing Optimization Algorithm (AMOSA)) in solving the problem. We also study the scheduling problem of multicast routing trees obtained from the MMHG model.

In the second work of the thesis, we study the cell outage compensation function of the self-healing mechanism using network cooperation scheme. In a heterogeneous network environment with densely deployed Femto Base Stations (FBSs), we propose a network cooperation scheme for FBSs using Coordinated Multi-Point (CoMP) transmission and reception with joint processing technique. Different clustering methods are studied to improve the performance of the network cooperation scheme.

In the final work of the thesis, we study the user cooperative multi-path routing solution for wireless Users Equipment (UEs)' streaming application using auction theory. We assume that UEs use multi-path transport layer service, and establish two paths for streaming events, one path goes through its cellular link, another path is established using a Wi-Fi connection with a neighbor UE. We study user coordinated multi-path routing solution with two different energy cost functions (LCF and EAC) and design user cooperative real-time optimization and failure protection operations for the streaming application. To stimulate UEs to participate into the user cooperation operation, we design a credit system enabled with auction mechanism.

Simulation results in this thesis show that optimal cooperation operations among network devices to reuse the existing spectrum wisely are able to improve network performance considerably. Our proposed network modeling approach in CRN helps reduce the complicated multicast routing problem to a simple graph problem, and the proposed algorithms can find most of the optimal multicast routing trees in a short amount of time. In the second and third works, our proposed network cooperation and user cooperation approaches are shown to provide better UEs' throughput compared to non-cooperation schemes. The network cooperation approach using CoMP provides failure compensation capability by preventing the system sum rate loss from having the same speed of radio resource loss, and this is done without using additional radio resources and will not have a significant adverse effect on the performance of other UEs. The user cooperation approach shows great advantage in improving service rate, improving streaming event success rate and reducing energy consumption compared to non-cooperation solution.

CHAPTER 1. OVERVIEW

1.1 Introduction

In modern cellular networks, mobile traffic is exponentially growing because of the broad use of smartphones, tablets together with the 'data hungry' applications, such as wireless high definition video application, location navigation service, online gaming etc. Meanwhile, the radio frequency spectrum is becoming overcrowded with the ever-increasing traffic demand. New advanced traffic-consuming applications create new challenges to the current network service.

To meet the growing demand on wireless network service, wireless technologies have experienced several generations and each generation is faster, more secure and reliable compared to previous generations. Starting from the first generation (1G) of mobile technology, only simple phone calls are supported. With the second generation (2G), mobile device can support simple text messaging. When the third generation (3G) is deployed, the mobile device becomes real smart. It can browse web, read email, download video and share pictures with other smart phones. With the forth generation (4G) and Long Term Revolution (4G-LTE), the wireless service rate is improved to at least 100 Megabits per second and up to 1 Gigabit per second. Now the fifth generation is under-development, and it supports operations in millimetre wave spectrum to achieve peak download speed up to 35 gigabits per second. With the scaled bandwidth support, 5G is designed to provide higher capacity than current 4G, allowing a higher density of mobile broadband users and supporting device-to-device, ultra reliable and massive machine communications [1].

Aside from expanding the usable frequency spectrum for wireless transmission, researchers also study new ideas to improve the current radio frequency spectrum efficiency. Cognitive radio

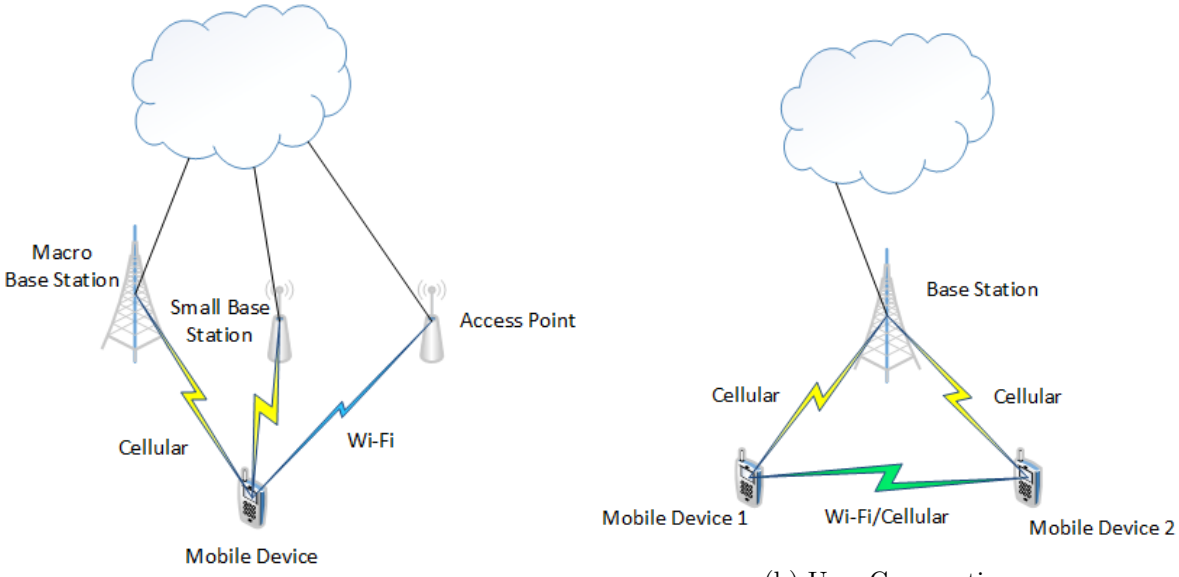
is an upgraded radio receiver and radio communication technology to utilize radio spectrum more efficiently. A cognitive radio is able to monitor the spectrum and detect the available frequencies, and access the available frequencies efficiently. The concept of cognitive radio was first proposed by Joseph Mitola III in 1998 and published in an article in 1999 [2].

One of the main goals of cognitive radio is to increase the spectrum utilization of the underutilized frequencies. Current frequency allocation is regulated by governments in most countries. Spectrum was assigned through administrative licensing, and was divided into different frequency bands where each has a specific application. Because of the fixed allocation, some frequency bands are under-utilized. A typical example is the TV White Space [3], which is the unused broadcasting frequencies left by television networks for buffering purposes. With cognitive radio, unlicensed Secondary Users (SU) can dynamically access under-utilized licensed spectrum bands, meanwhile controlling the interference towards the licensed Primary Users (PU).

Aside from cognitive radio, cooperation among network devices is another promising solution to improve spectrum efficiency, especially in the densely deployed small cell networks. Cooperation among network devices can increase bit rate, improve service reliability, and satisfy users' growing quality of service demand. Cooperation among network devices can be classified into two different cooperation schemes based on the cooperative network device type: network cooperation and user cooperation. Network cooperation is a cooperation scheme among base stations or other service provider deployed wireless devices, while user cooperation is a cooperation scheme among user owned mobile devices.

Fig. 1.1 shows an example of network cooperation and user cooperation. Network cooperation [4, 5] allows cooperation among base stations or relay points to provide service to UEs using signal coding and beam forming techniques. This approach can reduce intercell interference greatly and increase service rate, what's more, it can provide UEs with network failure protection capability. There are a number of studies that applied network cooperation approach to improve network performance [6–13]. The disadvantage of network cooperation approach is that not all wireless networks have coverage over the mobile devices. For example, Wi-Fi networks have disjoint coverage, and most mobile devices only have one subscriber

identity module (SIM) card and can only connect to one wireless network at one time. User cooperation [14–17] allows mobile devices in the vicinity to cooperate providing mobile users with stable quality of service. It's advantages over network cooperation approach is that: 1) power consumption of the UE is balanced by nearby UEs that relay packets for it; 2) if the UE can only connect to cellular network or Wi-Fi network, it can still enhance its bandwidth by connecting to relay UEs using its other available wireless interfaces [18]. User cooperation also has some major problems that need to be addressed: 1) most users are self interested, and are not willing to cooperate without any incentives; 2) information is not secure passing along different relay users; 3) without properly designed user cooperation scheme, some opportunistic users can take advantage of the user cooperation system.



(a) Network Cooperation

(b) User Cooperation

Figure 1.1: Cooperation schemes in wireless networks

1.2 Objectives of Research

The work of this thesis is motivated by the desire to explore advanced techniques to optimize various performance objectives in modern wireless networks.

We first explore techniques to improve multicast communication of unlicensed users in a cognitive radio network. Multicast, or one to many communication, is one of the important

service models. A growing number of applications rely on multicast for service support. For example, group-oriented mobile commerce, real-time event broadcasting, distance education and intelligent transportation system. For multicast application, the available unlicensed bandwidth or their licensed bandwidth may not be sufficient. With cognitive radio, the dynamic access of the other licensed spectrums can improve the accessible bandwidth but create new challenges. In this thesis, we employ a cooperative multi-hop system for unlicensed multicast user to dynamically access licensed user's frequency bands. The designed multicast routing scheme optimizes the throughput, end-to-end delay and number of hops used in the multicast communication.

Next, we explore network cooperation scheme to improve UE's spectrum efficiency in a heterogeneous network (HETNET). In a HETNET, mobile operators deploy small cell base stations (SBSs) overlaid with traditional macro base stations (MBSs) to increase spectrum efficiency, network capacity per area, and meet the users service requirements. SBSs in a dense HETNET suffer severe interference from other SBSs and MBSs, and they also suffer from high probability of failure caused by human interactions, since most of the SBSs are installed to their shops and households by users and are easily accessible to humans. We explore network cooperation schemes to reduce interferences among SBSs, increase system throughput and improve system's capability in providing service compensation to UEs when failures happen.

Finally, we explore user cooperation scheme to improve UE's spectrum efficiency and reduce UE's energy consumption in a HETNET. User cooperation scheme has the advantage to increase transmission throughput, reduce total energy consumption and balance the energy consumption among UEs. Moreover, user cooperation can help providing network failure protection. In the study, UEs are assumed to support multi-path transport layer protocol and device to device communication is enabled for UEs to establish user cooperation transmission. We explore user cooperation solutions to establish multi-path cooperative communications, improve communication capacity and reduce system energy consumption. We also study the failure protection capability of the user cooperation scheme. What's more, we design a credit system enabled with auction mechanism to incentivize UEs to participate into the user cooperation process.

1.3 Contributions

Our contributions of this thesis can be summarized as follows:

1. In the first contribution, we propose a new network modeling approach (Multi-rate Multi-layer Hyper-Graph (MMHG)) to model CRNs, and simplify the multicast routing problem in CRNs in terms of finding routing trees on the graph. We propose solutions to find the multicast routing trees that optimize multiple objectives of the multicast sessions.

2. In the second contribution, we propose schemes that use the Coordinated MultiPoint transmission and reception (CoMP) technique to provide small cell outage compensation capability in heterogeneous networks. We form a resource allocation problem to optimize the CoMP operation that maximizes the system throughput and compensates the system throughput in the case of failures.

3. In the third and final contribution, we study a mobile device relay scenario that is different from those in the related user cooperation schemes proposed in the literature. We study user cooperation operation in a single relay multi-path transmission scenario to optimize the transmission performance and provide transmission failure protection capability for streaming application. We design an auction mechanism that incentivizes UEs to provide relay service for other UEs. We also design UEs' real-time optimization and failure protection operation to provide better quality of service.

1.4 Dissertation Organization

The dissertation is organized as follows: Chapter 2 provides the literature review; Multi-objective multicast routing optimization in cognitive radio networks is introduced in Chapter 3; A CoMP-based self healing solution for heterogeneous femtocell networks is introduced in Chapter 4; User cooperative multi-path routing solution for streaming application using auction theory is introduced in Chapter 5. Finally, conclusion and future research are summarized in Chapter 6.

CHAPTER 2. LITERATURE REVIEW

2.1 Multicast Routing in Cognitive Radio Networks

There has been a lot of studies on multicast routing algorithms. As early as 1990s [19], multicast routing protocol for the internet has been proposed and a various number of multicast protocols are proposed and studied after that. This thesis introduces four of them: DVMRP, MOSPF, PIM and CBT.

- Distance Vector Multicast Routing Protocol (DVMRP): this protocol is defined in RFC 1075 [20]. DVMRP uses reverse path forwarding technique, that is when a router receives a packet, it floods the packet out of all paths except the one that leads back to the packet's source. DVMRP will periodically reflood in order to reach any new hosts that want to receive a particular group. However, it causes significant scaling problems.

- Multicast extension to Open Shortest Path First (MOSPF) [21]: it is an extension to the OSPF unicast routing protocol, and it works by including multicast information in OSPF link state advertisements. It builds a shortest-path tree rooted at the source node, and the trees are built on demand and the results are cached for use by subsequent packets. MOSPF is best suited for environments that have relatively few source/group pairs active at any given time.

- Protocol Independent Multicast (PIM) [22]: it is a family of multicast routing protocols for internet protocol networks. This protocol discovers network topology using routing information provided by other routing protocols. It has sparse mode (SM) and dense mode (DM). PIM-SM is commonly used in IPTV systems for routing multicast streams, and PIM-DM uses reverse path forwarding and is very like DVMRP.

- Cored-Based Trees (CBT) [23]: a proposal for making IP multicast scalable by constructing a shared tree of routers. It is a center-based tree protocol and the routing tree comprises of multiple "cores" routers. CBT is very efficient in terms of the amount of state routers need to keep and it scales better than flood-and-prune protocols.

These protocols have their advantages in traditional internet, but they do not perform well in a dynamic changing wireless network environment. Studies [24, 25] list challenges for supporting multicast service in mobile ad hoc networks (MANETs) and summarize multicast routing protocols designed for MANETs. In MANETs, network topology dynamically changes due to random node mobility. Hence the routing algorithms need to be more robust. Study [24] classifies the multicast protocols of MANETs as tree and mesh-based approaches. Tree-based multicast protocols can be further classified as source-tree based and shared-tree based. For source-tree based protocols, there are different protocols based on different objectives, for example, minimal hop based protocol (MAODV [26]), minimal link based (BEMRP [27]), stability-based (ABAM [28]) and zone-based (MZRP [29]). For shared-tree based protocols, there are cluster-based protocol (ST-WIM [30]), session-specific protocol (AMRIS [31]) and IP-Multicast based (AMRoute [32]). Mesh-based multicast routing protocols can be classified as source-initiated (ODMRP [33]) and receiver initiated (CAMP [34]) protocols.

CRN is different from traditional wireless network. In CRN, channels available to different users are different, and the channel availability is changing over time. CRN posts new challenges to multicast service. There is a number of works that study multicast problems in CRNs, and they explore the cooperative communication and scheduling techniques to maximize a number of different objectives: throughput maximization [35–37], spectrum efficiency [36, 38] and end-to-end delay minimization [39]. Reference [35] developed an optimal power control policy on base station (BS) and an efficient cooperative communication schedule among SUs so that aggregate throughput on all SUs is maximized. In their proposed framework, base station multicasts data to a subset of SUs first by carefully tuning the power. Concurrently, SUs opportunistically perform cooperative transmissions using locally idle primary channels. They adopt network coding to reduce transmission overhead and perform error control and recovery. Reference [36] studied a spectrum pooling opportunistic spectrum access approach. That is

spectrum ranges from different spectrum owners are merged into a common pool, from which the secondary users may temporarily rent spectrum resources during idle periods of licensed users. The authors of this paper study efficient resource allocation scheme in a OFDM-based CRN with common spectrum pool. A non-linear program was formulated to find the optimal spectrum and power allocation scheme on BS as well as SUs to maximize the total sum rate of SUs. Reference [39] considered multicast scheduling in a cell formed by a Mesh Router and its served Mesh Clients. To reduce the effect of the channel heterogeneity property on the multicast throughput in cognitive radio wireless mesh networks, and minimize multicast end to end delay, they proposed an assistance strategy that allows multicast receivers to assist the source in delivering the data, and allows the transmission of coded packets so that multicast receivers belonging to different multicast groups can decode and extract their data concurrently. Reference [38] studied a multicast communication problem in a multi-hop ad hoc cognitive radio network. To support a set of multicast session with a given bit rate requirement for each multicast session, they formulated the multicast routing problem as a mixed linear program with a multiple set of communication constraints, and an objective of minimizing the required network-wide resources. Reference [37] proposed a distributed and on-demand multicast routing protocol COCAST, where all nodes in the multicast group cooperatively detect other signals, and exchange information through piggybacked Join Query and Join Reply messages using a predefined common channel. Nodes select their channels based on exchanged information.

Different from their work, in this thesis, we propose a new network modeling approach to model the CRN, and we study the cooperation among SUs and multi-hop transmission scheduling problem to find the multicast routing trees that optimize multiple objectives of the multicast sessions.

2.2 Self Healing System

Self healing system is one of the major components of self-organizing network. Self-organizing network [40] is a network model that helps the network operator to reduce operational expenses (OPEX) through automation of some network planning, configuration and optimisation processes. Self-organizing network functionalities are commonly divided into three

main sub-functional groups [41]: self-configuration, self-optimization, and self-healing. Self-configuration functions enable new BSs to automatically be configured and integrated into the network. While self-optimization functions optimize the cell site control parameters according to different observations of network operation scenarios. When some nodes in the network become inoperative, self-healing mechanisms will automatically perform fault detection, fault diagnose, fault compensation and recovery algorithms, aiming to reduce the impacts from the failure.

Third Generation Partnership Project (3GPP) has been working on technical specifications on self-healing requirements since the release of LTE-A (a full version of the document, released in 2012, can be found in [42]). According to the specification and other survey papers [41, 43], self-healing system is summarized into four components as shown in Fig. 2.1: 1) fault detection, 2) fault diagnose 3) fault response and 4) fault recovery.

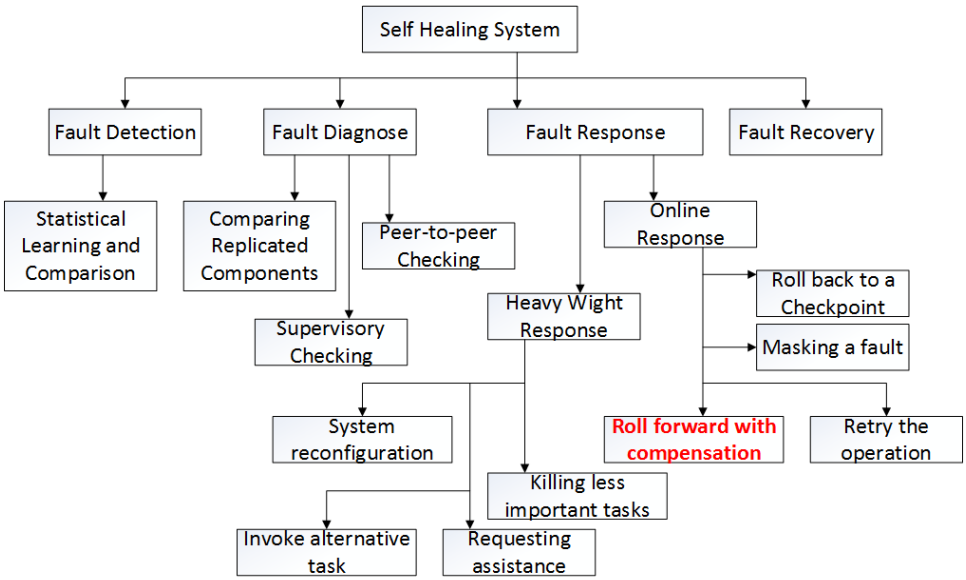


Figure 2.1: Self healing system components

In the research works of fault detection, Most of the proposed anomaly detection methods applied elaborate statistical models as well as deep statistical knowledge to better separate faulty situations from unusual but healthy cases. For fault diagnose, traditionally it is a manual activity based on the knowledge of experts. However, complete automation of the diagnosis process is a long-term result of an evolutionary process that starts with high dependency on

experts' knowledge to the future vision that human operators supervise the automated process while leaving the details to the diagnosis system [41]. Fault response is performed by the system during the time period starting when the network failure happens to the point when the network back to normal operation, and it is performed to compensate the network performance degradation. Depending on the real-time requirement, the fault response is classified into heavy weight response and online response. In heavy weight response scenario, network try to compensate the service by performing system reconfiguration, killing less important tasks, invoking alternative task or requesting external assistance. For online response, the system can mask a fault, roll back to a checkpoint previously saved, continue roll forward with compensation schemes, or retry the operation that caused the fault. For the online response that roll forward with cell outage compensation, the commonly used approaches in most of the published related research works are network element power management, antenna tilt, etc [44–48].

This thesis studies cell outage compensation schemes in a self healing system. Different from those related research works, we apply CoMP technique to provide cell outage compensation capability. We form a resource allocation problem to optimize the CoMP operation that maximizes the system throughput and compensates the system throughput when base stations suffer from failure.

2.3 Cooperation In Wireless Networks

Wireless networks are constrained by limited network resources (time, frequency and power). Cooperation schemes use network resources wisely to improve network coverage, throughput, spectral efficiency and power efficiency. According to [49], cooperative relay technique has three major type of structures as shown in Fig. 2.2: Coordinated Multi-point Transmission (CoMP), fixed relay and mobile relay.

CoMP is a technique to provide cooperative communications among multiple transmission and reception points in cellular systems. The cooperation points can be BSs, remote radio units of the same BS, remote radio unit of different BSs, or user equipments. 3GPP organization introduced CoMP in Lte-Advanced release 11 [5]. In CoMP, transmit points provide coordinated transmission in the downlink (DL), and receive points provide coordinated reception in the

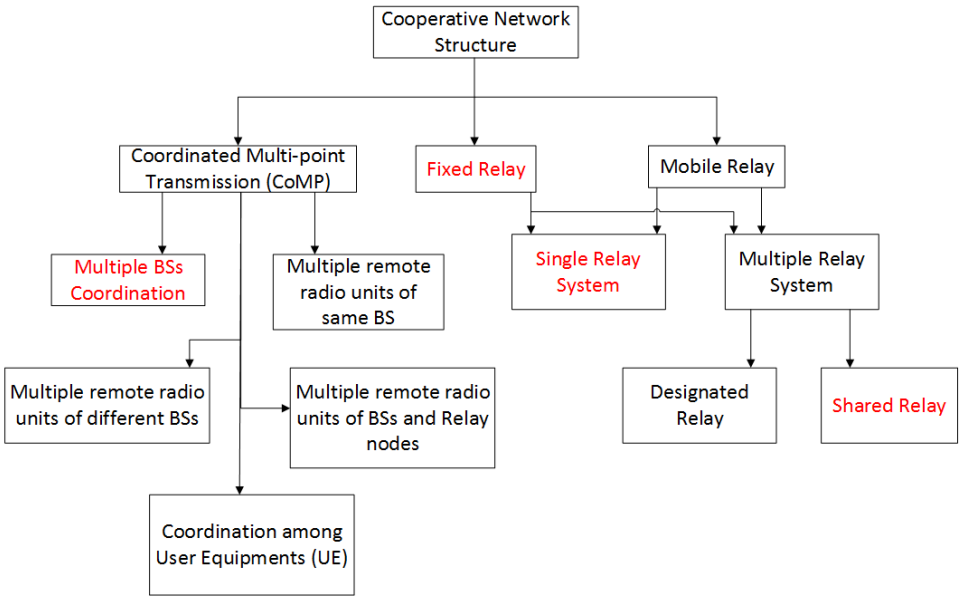


Figure 2.2: Cooperative relay network structure classification

uplink (UL). The transmit and receive antennas can be from multiple antenna site locations, which may or may not belong to the same physical cell.

As summarized by reference papers [50,51], there are three main types of CoMP techniques: 1) coordinated scheduling and coordinated beamforming (CS/CB); 2) joint transmission (JT) and 3) transmission point selection (TPS). Examples of CS/CB and JT are shown in Fig. 2.3. In CS/CB ([6–8]), the coordinated network devices (BS) exchange control information and schedule transmissions from BS to UEs to control the interference that can be generated. Therefore, UEs' data does not need to be transmitted from multiple BSs, and the beam details and scheduling decisions should be coordinated between BSs. In JT ([7,9–13]), data is transmitted to UE simultaneously from a number of different BSs with appropriate beam weight. This can increase signal quality and strength and actively cancel the interference to UEs from the coordinated BSs. TPS is a special form of JT, where transmission of beamformed data for a given UE terminal is performed at a single transmit point at each time instance, while the data is available at multiple coordinated TPs.

The challenge issues of CoMP are the formation of the CoMP operation units (cooperation points clustering and UE scheduling), and resource allocation within the CoMP operation unit

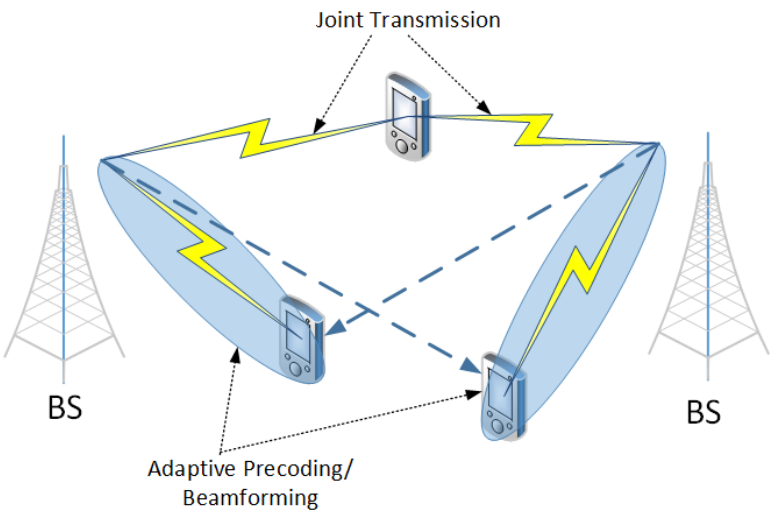
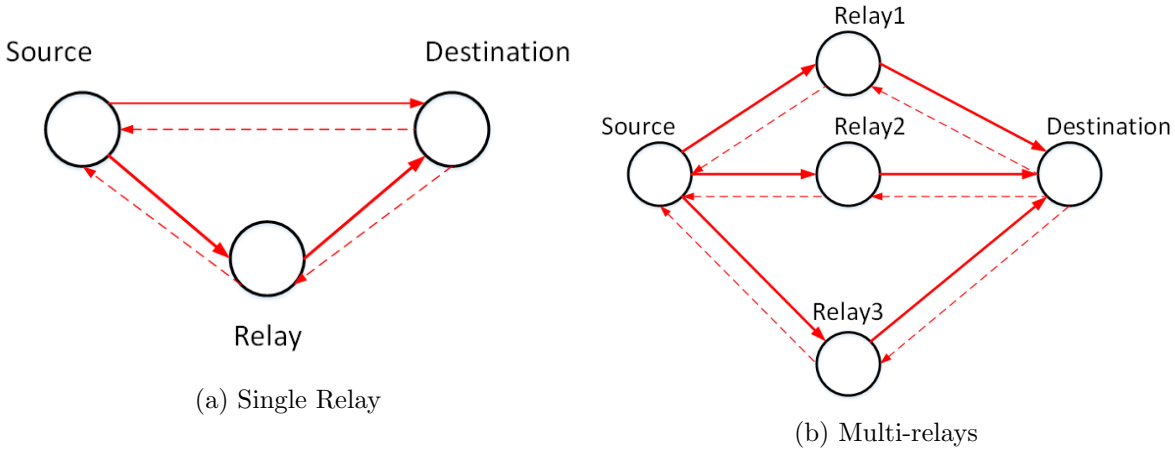


Figure 2.3: CoMP technique

and between CoMP operation units. In this thesis, we apply different clustering algorithms to cluster base stations into non-overlapping CoMP operation units, and form resource allocation problem to maximize the CoMP operation units' throughput.

Fix relays in the cooperation system are base stations or other user equipments. Fix relays' locations remain static, while mobile relays are relays installed in moving buses or trains, and their locations are moving constantly. For both fix relay system and mobile relay system, the cooperation structure can be divided into cooperation using single relay or cooperation using multi-relays. Example of single relay or multi-relays are shown in Fig. 2.4.



(a) Single Relay

(b) Multi-relays

Figure 2.4: Single relay and multi-relays

There are a number of device to device communication technologies that can be used for mobile devices cooperation purpose:

- Bluetooth: most energy efficient, and only supports point-to-point connections. It will be difficult to create large groups of devices and has relatively low speeds.
- Wifi hotspot: one device acts as the access point, and the other as client. It is widely supported and offers high speeds. However, the device acting as access point will consume more energy.
- Wi-Fi Direct [52]: allows true peer-to-peer operation and high speeds.
- LTE Direct [53]: emerging technology that uses the LTE band for energy-efficient device-to-device communication and discovery. LTE Direct is very promising but is not yet available.

In the third work of this thesis, we study a cooperative mobile user relay system, which is a mix of fix relay and mobile relay system. UEs are the relays to each other and each UE can at most employ one other UE as its relay. What's more, UEs are moving in constant time steps. During each time step, UEs remain static. UEs are assumed to be enabled with the technologies above to do device to device communication for relay purpose. The key issues in cooperative mobile user relay system can be divided into the following three points:

- Willingness for nearby mobile devices to forward other UEs' packets. This is because users can not see immediate benefit and reward for their cooperation, and their operations are limited by their battery and radio resource.
- Information security of the data relayed by nearby mobile devices.
- Prevent misbehaving user (selfish, malicious, or hackers) in taking advantage of the user relay system.

Study [54] provides a summary of different research components for user cooperation schemes with different motivations, and it is shown in Fig. 2.5. Good incentive mechanism [55–61] improves user's willingness in providing relays for other users. Enforcement strategies [55–61]

prevent misbehaving users from taking advantage of the system. Good relay selection approaches [14–17, 55–61] let users choose relay users that can benefit themselves the most.

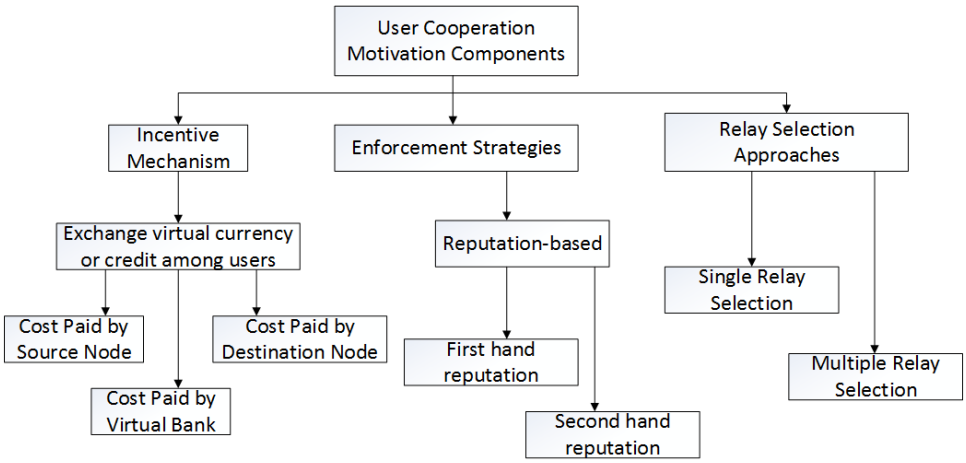


Figure 2.5: User cooperation components with motivation

References [14–17] are some of the recent related research works that study relay selection approaches. Their works are mainly focused on resource allocation, route selection and interference cancelation in cooperative communications, but they assume UEs are willing to provide relay service without any incentives. Reference [55–61] are related works that study incentive mechanisms, enforcement strategies and relay selection approaches jointly to improve performance of device to device communications.

In this thesis, we study a different relay scenario than those related works. We study a single relay multi-path transmission scenario to optimize the transmission performance and provide transmission failure protection capability. To create incentives for users to relay traffic for other users, we design a credit based auction system with iterative matching process. The bid provided by each neighbor user is optimized with respect to the power consumption. We also design a real-time optimization and failure protection operation to improve UEs' quality of service further.

2.4 Multi-path Transport Layer Protocol

In the third work of the thesis, we study user cooperation solution for multi-path streaming application in wireless HETNET. There are three types of applications (app) on mobile devices (real time streaming and interactive multimedia communication app, web browsing app, app download app). The reason we focus on streaming application is that the real time streaming and interactive multimedia communication apps have the highest priority in case of service failure, since they require continuous streaming service to have good QoS. It is much interesting and useful to study operation schemes that can improve performance of streaming application.

There are two major multi-path transport layer protocols in the literature, one is extended from regular TCP and the other one is extended from regular RTP.

Multipath TCP (MPTCP) [62], as the extension to regular TCP, provides the ability to simultaneously use multiple paths between peers. The simultaneous use of multiple paths would improve resource usage within the network, improve user experience through higher throughput and provide resilience to network failure. MPTCP has been implemented and been used by commercial products, such as Siri from Apple [63]. The source code of MPTCP can be found in [64].

MPTCP has some components can be further studied based on user's requirement. From receiver considerations, the design of retransmission strategies and receive buffer sizing, as well as the relationship between retransmission strategies and receive buffer sizing can be studied further. What's more, MPTCP does not mandate any mechanisms for handling retransmissions. Different congestion controllers need to be implemented for MPTCP, each aiming to achieve different properties in the resource pooling/fairness/stability design space, as well as quality of service, reliability and resilience.

Multi-path communication model for Real-time Transport Protocol (RTP) is first proposed by reference paper [65], and the RFC draft can be found on website [66]. MPRTP presents minimal set of required protocol extensions. Reference paper [66] also develops algorithms for scheduling RTP traffic across multiple paths at the sender and a corresponding de-jittering algorithm at the receiver.

MPTCP could be applied to media streaming, but does not consider real-time data and diverse paths, which may lead to worst case delay and thus even longer buffering time. MPRTP is extended from the Real-time Transport Protocol (RTP), and RTP is suitable for applications such as live streaming and broadcast, video on-demand and interactive multimedia communication. RTP uses the associated RTP Control Protocol (RTCP) for monitoring the end to end media delivery and for other control operations.

From a functional perspective, MPRTP [66] must be able to: 1) adapt to bandwidth changes, make use of multiple paths and adapt to their relative capacity changes by redistributing the load, which should be done in a way that avoids oscillation. 2) overcome packet skew. Different paths will likely exhibit different RTTs, mechanisms must be put in place to overcome the resulting skew. 3) choose transmission paths. The choice of suitable transmission path should reflect the demands of the application. From a protocol perspective, RTP must be extended to provide the additional information necessary to perform these functions, yet maintain backwards compatibility.

CHAPTER 3. MULTI-OBJECTIVE MULTICAST ROUTING OPTIMIZATION IN COGNITIVE RADIO NETWORKS

3.1 System Model

In this work, we assume a CRN operating on TV white space channels (television channels 2-51) [67]. Each channel is licensed to a PU. There is a set of SUs in the CRN that can operate on all available channels, but we assume that each SU has one radio only, and can transmit or receive on one channel at a time. Moreover, each SU knows the availability of all channels and the location information of all nodes in the CRN (including all SUs and PUs). The CRN uses an overlay model, such that an SU's transmission can cause interference to an active licensed PU of the same channel if the SU is located within the PU's interference range, and vice versa. We also assume that the network is quasi-static, such that the PUs will not change their activities during our computing process.

With the assumption that all SUs adopt the best modulation and coding scheme, we apply Shannon-Hartley's formula to calculate the maximum achievable rate of a communication session: $C = Wlog_2(1 + \frac{P_t G}{d^\alpha N_0 W})$. C is the maximum achievable rate in $bits/second$, W is the channel bandwidth in Hz, P_t is the transmission power in $Watts$, d in meters is the transmission range of the SU transmitter, G is the antenna gain, N_0 is the noise power spectrum density in W/Hz, and α is the path loss exponent which is in the range from 2 to 4. In our work, we use W as the TV channel bandwidth, 6 MHz, P_t as the maximum transmission power that is allowed by the IEEE 802.11 protocol which is 100 mW, G as 1, and N_0 as the thermal noise power spectrum density -174dbm/Hz and $\alpha = 2$.

The CRN is modeled using a directed HyperGraph. Reference [68] proposed the use of multilayer hypergraph in modeling CRNs. In our work, we use multilayer hypergraph to model

a CRN while introducing a significant extension. In a hypergraph, a hyperedge, also referred to here as a supernode, models a transmission process which consists of a transmitter SU set (TSU) and a receiver SU set (RSU), where they operate on the same channel. TSU only contains one SU, and RSU can have multiple SUs due to the wireless broadcast feature. Let d_r denotes the transmission range of the TSU in the supernode. Then, the RSU is a set of SUs that are located within d_r from the TSU. The transmission rate C within a supernode is defined by Shannon-Hartley's formula, with d set to the transmission range d_r. The same SU of the TSU may form a different supernode that transmits at a higher rate, $C^{'} > C$, and over a shorter distance, $d_r^{'} < d_r$, such that the new RSU is a set of SUs within a maximum distance $d_r^{'}$ from the TSU, and a subset of SUs within a distance d_r from the TSU. The cost of the supernodes with transmission rate C is $\frac{1\,data\,segment}{C}$, where 1 data segment is the size of scheduled transmitted data every transmission cycle, as will be explained below. The hypergraph also has a set of dummy supernodes to represent each transmitting SU and receiving SU. If there are n SUs in the CRN, then n transmitter dummy supernodes (TDsupernode) and n receiver dummy supernodes (RDsupernode) are present in the hypergraph. TDsupernodes and RDsupernodes have no rate, no operating channel, no cost and both their TSU and RSU correspond to the same SU. To distinguish them from dummy supernode, we use Csupernode to refer to communication supernode that is neither TDsupernode nor RDsupernode. TDsupernodes and RDsupernodes will always exist in the hypergraph, but one condition must hold for a Csupernode to be present in a hypergraph: its TSU is not located within the interference range of the active PU licensed on the same channel that it uses.

Fig. 3.1 shows an example of how the supernodes are defined. In this simple network, all SUs are on the same channel. SU 2, 4 and 6 are located within PU A's interference range, and PU A is located within SU 6's interference range. Here we assume that an SU's transmission range $d_{r_1} = 50$ meters with achievable rate $R1$, $d_{r_2} = 100$ meters with rate $R2$, and $d_{r_3} = 150$ meters with rate $R3$, where $R_1 > R_2 > R_3$. Table ?? shows the distances among all SUs of the network shown in Fig. 3.1. As shown in the figure, we can create 3 Csupernodes with TSU as $\{1\}$, with rates $R1$, $R2$ and $R3$, respectively. The numbers between the curly braces are SUs ID. We can also create one TDsupernode as well as one RDsupernode with respect to SU 1 (they

Table 3.1: Table of notation

Notation	Description
APF	Approximated Pareto Front
RDsupernode	Receiver Dummy Supernode
BS	Base Station
RSU	Receiver Secondary User Set
CRN	Cognitive Radio Network
SU	Secondary User
Csupernode	Communication Supernode
TDsupernode	Transmitter Dummy Supernode
MMHG	Multi-rate Multilayer HyperGraph
TSU	Transmitter Secondary User Set
PU	Primary User
URSA	Uniformly Random Search Algorithm

have the same structure but will be distinguished in code implementation). Csupernodes with TSU as {3} or {5} can also be created but they are not shown in Fig. 3.1. Notice that due to interference, Csupernode with TSU as {2}, {4} or {6} cannot be created in the corresponding hypergraph.

Table 3.2: Distances among SUs in Fig. 3.1

Distance(meters)	SU 1	SU 2	SU 3	SU 4	SU 5	SU 6
SU 1	—	50	65	130	180	260
SU 2	50	—	185	55	280	190
SU 3	65	185	—	220	150	320
SU 4	130	55	220	—	330	200
SU 5	180	280	150	330	—	500
SU 6	260	190	320	200	500	—

The following conditions should be satisfied when considering the existence of a link between two supernodes in the hypergraph (we define TSU_i and RSU_i as the TSU and RSU of supernode i):

(1) A link exists from Csupernode i to Csupernode j if $TSU_j \subseteq RSU_i$ and the active PU of the same channel used by Csupernode i is not located within the interference range of the

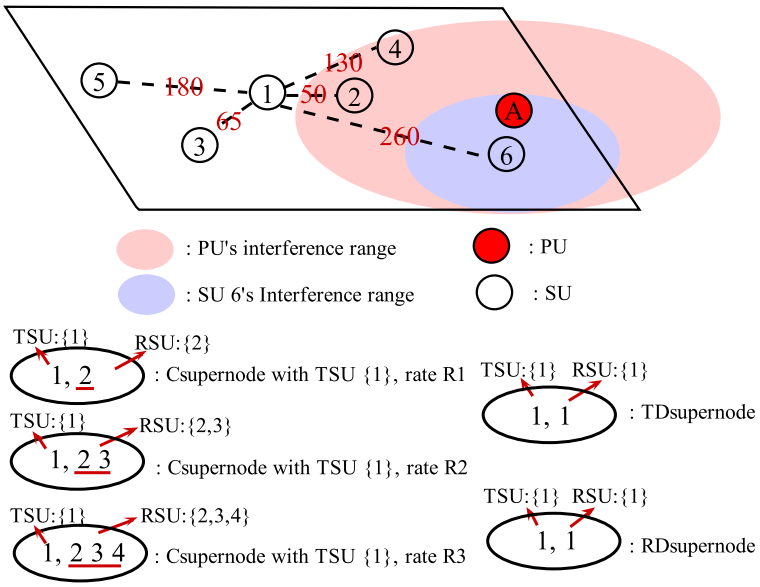


Figure 3.1: Example of supernode construction

SU in TSU_j.

(2) A link exists from TDsupernode i to Csupernode j if TSU_j has the same SU as the SU that TDsupernode i represents.

(3) A link exists from Csupernode i to RDsupernode j if $TSU_j \subseteq RSU_i$ and the active PU of the same channel used by Csupernode i is not located within TSU_j's interference range.

(4) No links exist between TDsupernodes and RDsupernodes.

The cost of a link between two Csupernodes is $k * W * |c_i - c_j|$, where c_i and c_j are the channel numbers used by Csupernode i and Csupernode j, respectively, k is the constant $10ms/10MHz$ [69] and W is the TV channel bandwidth 6 MHz. When $c_i \neq c_j$, the link cost is the switching delay of the same SU in TSU of Csupernode j switching from channel c_i to channel c_j. The costs of links from TDsupernodes to Csupernodes, or from Csupernodes to RDsupernodes are 0 seconds.

Fig. 3.2 shows a multilayer graph representation of a simple CRN of two channels, 6 SUs and 2 active PUs. This CRN is extended from the network shown in Fig. 3.1. The virtual dotted links between SUs show that they can switch between channels. Using the principles we

defined for constructing supernodes and links between supernodes, Fig. 3.3 shows the directed hypergraph mapped from the multilayer graph shown in Fig. 3.2. The six shaded supernodes on the left side of the hypergraph are TDsupernodes, and other six shaded supernodes on the right side of the hypergraph are RDsupernodes. The dashed circles represent Csupernodes operating on channel 1, and the two unshaded undashed circles represent Csupernodes operating on channel 2. Link costs are not shown in Fig. 3.3. We create 3 Csupernodes with TSU as {1}. We also create two Csupernodes with TSU as {3} and one Csupernodes with TSU as {5}. Csupernodes operating on channel 1 with TSU as {2}, {4} or {6} are not present in the hypergraph because SUs 2, 4 and 6 are located within PU A's interference range. On channel 2, we create two Csupernodes with TSU as {4}. Csupernodes with TSU as {1}, {2}, {3} or {5} are not present in the hypergraph, as SUs 1, 2, 3 and 5 are located within PU B's interference range. There are no Csupernodes with TSU as {6} available in the hypergraph since no SUs are located within SU 6's maximum transmission range.

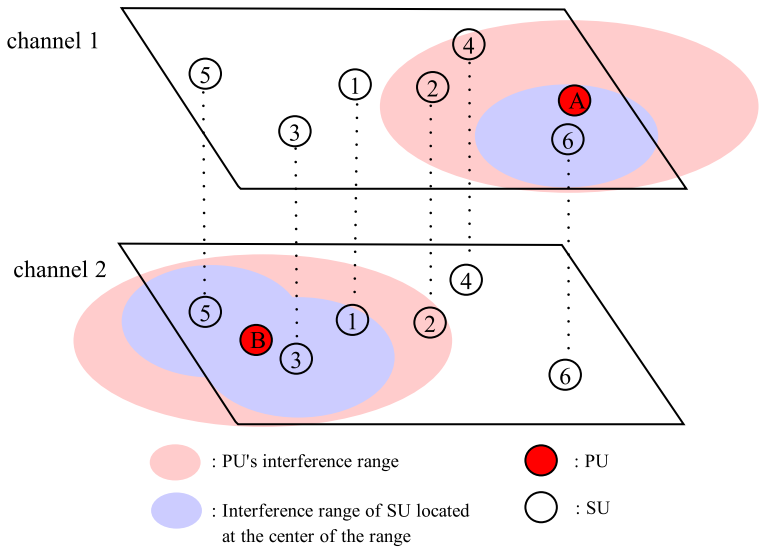


Figure 3.2: A multilayer graph representation

The rest of this work is organized as follows. our research problem is described in Section 3.2. In Section 3.3, we show how to use metaheuristic search algorithms to solve the problem. Simulation results are shown in Section 3.4. Section 3.5 concludes the work.

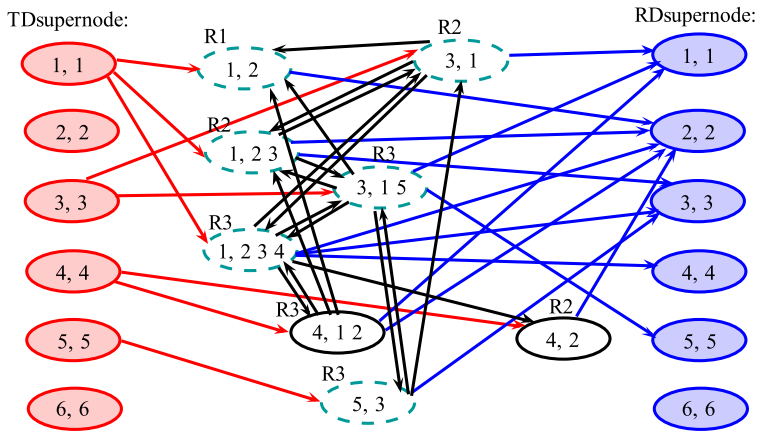


Figure 3.3: The directed hypergraph mapped from the multilayer graph showed in Fig. 3.2

3.2 Problem Description

Given a multicast session (s, Des), where s is a source TDsupernode, Des is a set of destination RDsupernodes, our goal is to find the optimal set of multicast routing trees satisfying three objectives: 1) minimal worst case end-to-end delay (*delay*), 2) maximal data rate (*rate*) and 3) minimal number of links (*numOfLinks*) used in the multicast tree. The problem is a multi-objective optimization problem with decision variables $X_{solu} = (x_1, x_2, x_3, ..., x_n)$, where n is the number of all links in the hypergraph, and $x_k \in \{0, 1\}$ is 1 if link k is selected to build the multicast tree, and 0 if not. With the computed Pareto Front of multicast routing trees, we select the final solution out of the Pareto Front depending on which objective is more important.

3.3 Multi-Objective Optimization Solvers

Evolutionary Algorithms (EA) are usually used for solving multi-objective optimization problems. Kalyanmoy Deb's Non-dominated Sorting Genetic Algorithm-II (NSGA-II) [70] and Eckart Zitzler's Strength Pareto Evolutionary Algorithm-2 (SPEA2) [71] are two classic EAs which are used for this purpose. The basic idea of NSGA-II is that it builds a population of competing individuals, ranks and sorts each individual according to non-domination level,

applies evolutionary operations to create new pool of offsprings, and then combines the parents and offsprings before partitioning the new combined pool into fronts. SPEA2 uses an external archive to contain non-dominated solutions previously found, and non-dominated individuals found through mutation are copied to the external archive at each generation.

Most EAs are suitable for solving optimization problems with continuous decision variables. However, the decision variables in our problem are binary discrete, and are constrained by tree structure. The evolutionary operations to create new pool of offsprings are not efficient for our problem. Therefore, we adopt two other metaheuristic algorithms, which have better features than EAs in solving our problems. The two algorithms are MOACS [72] and AMOSA [73]. For the rest of this section, Subsection 3.3.1 introduces the algorithms for generating a random multicast tree solution. MOACS and AMOSA are described in Subsection 3.3.2 and Subsection 3.3.3, respectively.

3.3.1 Generating A Random Multicast Tree Solution

It is not hard to search for a random multicast tree in a given hypergraph. However, we cannot compute its corresponding *delay* and *rate* unless we find its optimal transmission schedule. The process of generating a random multicast tree solution are as follows:

Step 1, search for a random multicast tree in the hypergraph.

Step 2, modify the tree in order to remove instances of duplicate transmissions. Notice that for Csupernodes operating on the same channel, if Csupernode i and Csupernode j have the same TSU, then $RSU_i \subseteq RSU_j$ or $RSU_j \subseteq RSU_i$. But for Csupernodes operating on different channels, due to different PU activities, two Csupernodes with the same TSU can have RSUs that are not a subset of each other. In essence, two Csupernodes i and j are duplicate transmissions when they have the same TSU and $RSU_i \subseteq RSU_j$ or $RSU_j \subseteq RSU_i$. The multicast tree found by Step 1 sometimes has such duplicate transmissions that are wasteful of bandwidth, and should be merged.

Step 3, find the optimal transmission schedule of the modified tree.

Step 4, compute the *delay*, *rate*, and *numOfLinks* with the given modified tree and transmission schedule.

Step 1 and Step 2 are presented in Subsection 3.3.1.1. Step 3 is explained in Subsection 3.3.1.2. Step 4 is described in Subsection 3.3.1.3.

3.3.1.1 Building a multicast tree

To build a multicast tree in a hypergraph, we apply algorithm BuildTree() from reference [72] with modifications. Algorithm 1 gives the general procedure for finding a multicast tree. We start from the source TDsupernode s, and a non-visited supernode is selected at each step. We assign the probability of selecting next un-visited neighbor node of supernode i using equation 3.1.

$$p_{ij} = \begin{cases} \frac{C_j}{\sum_{\forall g \in N_i} C_g} & \text{if } \ j \in N_i \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

In the equation, N_i is the set of unvisited neighbor supernodes of supernode i, $C_j = 1/(t_{ij} + t_j)$, t_{ij} is the cost of link (i, j), and t_j is the cost of supernode j. This process continues until all the destination RDsupernodes in Des of the multicast session are reached. In MOACS, the equation to compute the probabilities is different from equation 3.1, and will be introduced in later sections.

The repeat loop searches for a random multicast tree. However, the multicast tree will have links that may not lead to the destination supernodes, and we call those branches useless branches. PruneTree() is a function for cutting off useless branches.

The ModifyTree() function is used to do following modifications to the generated multicast tree:

(1) Merge supernodes of the same channel if they are duplicate transmissions.

(2) Merge supernodes of different channels if they are duplicate transmissions.

(3) Check destination RDsupernodes's reachability. In the multicast tree built by BuildTree(), some destination RDsupernodes might not be connected to the Csupernodes (whose RSU contains SU corresponding to the RDsupernode) which have the shortest link distance to the source TDsupernode. In reality, this is not the case. We correct such unrealistic connections.

Fig. 3.4 shows a detailed example of how ModifyTree() works. The example is based on a random generated network of 15 SUs (with ID from 0 - 14), 3 different channels, and 3

Algorithm 1 Procedure of building multicast tree $Xsolu$

Input: graph $H(V_H, E_H)$, multicast session(s, Des)
Output: arc set $Xsolu$ which forms a feasible multicast tree
Init: $Xsolu = \phi$; $D_r = \phi$; $R = \{s\}$
repeat
    randomly select node i of R and build neighbor node set N_i
    if $N_i = \phi$ **then**
        $R = R - \{i\}$
    else
        Select node j from N_i with probability P_{ij}
        $Xsolu = Xsolu \cup (i, j)$
        $R = R \cup j$
        if $j \in Des$ **then**
            $D_r = D_r \cup \{j\}$
        end if
    end if
until $R == \phi$ &&$D_r == Des$
PruneTree($Xsolu$)
ModifyTree($Xsolu$)
PruneTree($Xsolu$)
return $Xsolu$

primary users, one for each channel. The corresponding hypergraph of the network has 100+ supernodes, and 1000+ links. There are three different transmission ranges corresponding to three different transmission rates. Supernode 86 (2, 2) is the source TDsupernode, s, and supernodes 103 (4, 4), 104 (5, 5) and 105 (6, 6) are in destination RDsupernodes set, Des. In Fig. 3.4, 86, 54, 70, 103, 45, 105 and 104 are supernodes. [54: (2, 9 1 4) 36.5 Mbits/s 2] means supernode 54 has TSU as {2} and RSU as {9, 1, 4}. Its transmission rate is 36.5 Mbits/s, and operates on channel 2. From (a) to (b), supernodes 53, 54 are merged into 54, and 44 and 45 are merged into 45. Notice that 53 and 54 operate on the same channel and have the same TSU, and so are 44 and 45. In (b), supernode 35 and 70 have the same TSU, but different channels. But since $RSU_{35} \equiv RSU_{70}$, we merged 35 into 70 as shown in Fig. 3.4(c). Fig. 3.4(c) to Fig. 3.4(d) is the process of checking destination node reachability. Since in Fig. 3.4(c), each destination RDsupernode is connected to the supernode that has the shortest link distance to source TDsupernodes, so Fig. 3.4(d) maintains the same tree structure as in Fig. 3.4(c).

With the final modified multicast tree, we prune the useless branches using pruneTree(). The resulting multicast tree is the final multicast routing tree solution $Xsolu$.

3.3.1.2 Transmission Scheduling

In this subsection, we want to schedule a transmission cycle, which consists of a number of time units that have the same constant duration. There is a set of Csupernodes to be scheduled to transmit at each time unit. Our goal is to compute what Csupernodes of the given multicast tree should be scheduled in which time unit, such that a minimal total number of time units in a transmission cycle is obtained. With the optimal transmission cycle, we compute the time interval needed between each time unit of a cycle to prevent any interference.

In the heuristic algorithm, there are two major phases. Phase 1 uses a weighted graph-coloring heuristic algorithm [74] to compute the transmission cycle of a given multicast tree. Phase 2 formulates a simple linear program (LP) to compute the time interval needed between each time unit of a cycle.

In Phase 1, we use weighted graph-coloring algorithm to compute the transmission cycle. The procedure is shown in Algortithm 2. We assign weight 1 to Csupernodes of the largest rate R_{max} and other Csupernodes are assigned with weight $\lceil R_{max}/R_i \rceil$ where R_i is the rate of the Csupernode i, since the same amount of data transmitted by Csupernode of R_{max} will take more time to be transmitted by Csupernodes with rate lower than R_{max}. Note that we can also use the least common multiple (LCM) of rates of all Csupernodes in the multicast tree, divided by R_i to denote the weight of each Csupernode i, but this may cause the cycle length to be too long to be a feasible transmission schedule.

All Csupernodes of the given multicast tree form all nodes in F_g. For a Csupernode i, we use $IntendDes_i$ to denote the set of SUs that are the intended receivers of TSU_i. The $IntendDes_i$ of TSU_i is the union of TSUs of all supernodes (including RDsupernodes) that have the outgoing link from supernode i in the hypergraph. A link exists from Csupernode i to Csupernode j in F_g when:

(1) Csupernode i and Csupernode j are of the same channel, and TSU_i is located within the interference range of any node in $IntendDes_j$, or $TSU_j \in RSU_i$.

(2) Csupernode i and Csupernode j are on different channels, and $TSU_i \cap TSU_j \neq \emptyset$ (transmitter SU of Csupernode i and Csupernode j are the same) or $TSU_j \subseteq IntendDes_i$ (transmitter SU in Csupernode j is the intended receiver SU of Csupernode i) or $TSU_i \subseteq IntendDes_j$ (transmitter SU in Csupernode i is the intended receiver SU of Csupernode j) or $IntendDes_i \cap IntendDes_j \neq \emptyset$ (Csupernode i and Csupernode j have at least one common intended receiver SUs).

Algorithm 2 weighted graph-coloring algorithm

1: **Input:** Graph $H(V_H, E_H)$,$Xsolu$
2: **Output:** a schedule cycle $XsoluSche$
3: Compute the weight for each node in $Xsolu$ tree
4: Build the conflict graph F_g based on $Xsolu$ and interference model. Assign weight w_i to node v_i in $Xsolu$
5: Construct a new conflict graph $F_g^{'}$ from F_g as follows:
6: for each node v_i with weight w_i, w_i nodes are created,v_i^1,v_i^2,v_i^3,...,$v_i^{w_i}$ and add them to $F_g^{'}$. Add to graph $F_g^{'}$ the edges connecting v_i^a,v_i^b for $1 \leq a < b < w_i$. Add to graph $F_g^{'}$ an edge between v_i and v_j if and only if there is an edge between v_i and v_j in graph F_g.let n be the number of nodes in $F_g^{'}$.
7: **while** $F_g^{'}$ is not empty **do**
8: Find the node v_i with the largest $d_i^{in}(F_g^{'}) - d_i^{out}(F_g^{'})$ (d_i^{in} and d_i^{out} are incoming degree and outgoing of node i) in $F_g^{'}$ and remove this node from $F_g^{'}$ and all its incident edges. Let v_k denote the kth node removed. Process the sequences of nodes v_i from v_n to v_1. Assign each node v_i the earliest time slot not yet assigned to any of its neighbors in $F_g^{'}$.
9: **end while**
10: **return** $XsoluSche$

The following gives an example of the computed transmission cycle of a random generated network of 15 SUs (with ID from 0 - 14), 3 different channels, 3 primary users, one for each channel. The corresponding hypergraph has up to 80+ supernodes and 1000+ links. There are three different transmission rates for all SUs. In the hypergraph, {55} is the source TDsupernode, s, and {72, 73, 74, 75} is the destination RDsupernode set Des. BuildTree() generates a random multicast tree which is shown in Fig. 3.5:

Csupernodes 9 and 30 have rate 36.5 Mbits/s. Csupernodes 38 and 35 have rate 52.2 Mbits/s. So $R_{max} = 52.2$ Mbits/s. The weight for Csupernodes 38 and 35 is 1, and weight for Csupernodes 9 and 30 is $\lceil 52.2/36.5 \rceil$, which is 2. The computed transmission cycle $XsoluSche$ of the multicast tree shown in Fig. 3.5 is:

unit 1: 9, 38

unit 2: 9

unit 3: 30, 35

unit 4: 30

The transmission cycle of the multicast tree has 4 time units. Each time unit has a period of (1 data segment)/(52.2 Mbits/s). Fig. 3.6 shows the transmission procedure in details.

As shown in Fig. 3.6, when SU 2 continues transmitting, the multicast tree follows the same transmission procedure that SU 2 sends out one data segment at time unit 1 of each cycle i and SUs 4 and 5 finish receiving the data segment at time unit 4 of the same cycle i. SUs 6 and 7 finish receiving message at time unit 1 and 3 of cycle $i+1$, respectively. Further more, SU 4, 5, 6, 7 will receive one data segment every cycle starting from cycle 2. So, the maximal data rate the multicast tree can reach is (1 data segment)/(1 cycle time interval). The worst case end-to-end delay is the longest time it takes for one data segment to reach all destination SUs. In Fig. 3.6, the worst case end-to-end delay is the time interval between SU 2 sending out the data segment and SU 7 finishes receiving the data segment, a total of 7 time units. This is the case when we did not consider switching delay interference, and it will be considered in Phase 2.

Notice that the transmission order of SUs in different time units of a transmission cycle computed by Algorithm 2 will affect the worst case end-to-end delay. For example, a transmission cycle of [unit 1: 9, 38; unit 2: 9; unit 3: 30, 35; unit 4: 30] has a worst case end-to-end delay of $7 * [1\, unit\ time]$, and a transmission cycle of [unit 1: 9, 38; unit 2: 30, 35; unit 3: 30; unit 4: 9] has delay of $10 * [1\, unit\ time]$.

To compute the optimal transmission order of a given transmission cycle, one solution is to consider all permutations of the order of the transmission, and record the best permutation. However it is not realistic since the time complexity is $O(n^n)$. A heuristic algorithm is proposed in this work to compute a suboptimal but reasonably accurate solution. The procedure of the algorithm is illustrated in Algorithm 3.

With the computed transmission cycle, interference may still exist due to the switching delay between different channels. For example, in Fig. 3.6, supernode 30 finishes transmitting at time

Algorithm 3 Heuristic algorithm for finding suboptimal scheduling order

1: **Input:** Graph $H(V_H, E_H)$,$Xsolu$,$XsoluSche$
2: **Output:** $suboptXsoluSche$
3: compute the level (distance from source TDsupernode) of all supernodes in Multicast tree Xsolu.
4: sort supernodes scheduled for transmitting in the same time units of $XsoluSche$ in increasing order respect to level value.
5: sort the time unit ID of a transmission cycle in increasing order of the minimal level value of supernodes scheduled for transmitting in each time unit.
6: if two time units have the same minimal level value, then continue comparing the second lowest level value.... if two time units has exactly the same level value, then order these two time units randomly.
7: return $suboptXsoluSche$

unit 4 of cycle 1, and supernode 38 starts transmitting at time unit 1 of cycle 2. In this case, interference exists. Since SU 5 of supernode 30 receive a message at channel 1, and it transmits on channel 2 in supernode 38. It takes $k * |f_i - f_j| = 10ms/10MHZ * |2 - 1| * 6MHZ = 6$ milliseconds to switch from channel 1 to channel 2. There is no time interval between time unit 4 in cycle 1 and time unit 1 in cycle 2 for SU 5 to switch channels, which will cause collisions in successive transmissions.

In Phase 2, we formulate an LP to compute the time interval needed between each time unit of a cycle. Csupernodes of different time units may have different operation channels, and the time intervals between each time unit are used for SUs to switch between channels. In the LP, our objective is to minimize the sum of all time intervals of a cycle. We use the previous example to illustrate the process of formulating the LP. Given the previous computed transmission cycle, we record the SUs that are scheduled for transmission or reception during each time unit, as shown in Fig. 3.7:

We use $[inter\, i]$ to represent the value of time interval i in seconds between time units i and $i + 1$, use $[unit\, time]$ to represent the period of one time unit in seconds and $[1\, data\, segment]$ to represent size of the transmitted data segment in bits. Notice that all different time units in the transmission cycle have the same period of time. In the following LP, $[inter\, i]$ of each time interval i are the decision variables, and they do not necessarily have the same period of time.

$Minimize:\ \ [inter\,1] + [inter\,2] + [inter\,3] + [inter\,4]$

$Subject\ To:$

A: for unit 1

$$[inter\,4] \geq 0.006 \times |2-1| \tag{3.2}$$

B: for unit 3

$$[inter\,1] + [unit\,time] + [inter\,2] \geq 0.006 \times |1-2| \tag{3.3}$$

C: for each time interval

$$[inter\,1] \geq 0 \tag{3.4}$$

$$[inter\,2] \geq 0 \tag{3.5}$$

$$[inter\,3] \geq 0 \tag{3.6}$$

$$[inter\,4] \geq 0 \tag{3.7}$$

D: for each time unit

$$[unit\,time] = \frac{[1\,data\,Segment]}{R_{max}} \tag{3.8}$$

The constraints above show that there should be sufficient time gap for SUs to switch channels between two successive time units where they are scheduled for operating. In the constraints above, $0.006 = k * 6MHz$ and k is $10ms/10MHz$.

This is a polynomial time solvable LP, and we use CPLEX [75] to solve it. Notice that a sufficiently large size of $[1\,data\,segment]$ can result in a large $[unit\,time]$, such that the sum of time intervals in the transmission cycle can be minimized. However, the size of 1 data segment can not be set too large, otherwise the retransmission cost will be huge when the network is unstable. A tradeoff needs to be considered here to set the size of 1 data segment, and is not studied in this work.

3.3.1.3 Computing the objective value

With the generation of the multicast routing tree, the computed transmission cycle and the computed time intervals between each time unit, we can compute the three objective values of the multicast tree.

delay is the time interval from the time unit where source SU sends the data segment and the time unit where all destination SUs finish receiving the data segment.

rate equals the size of 1 data segment divided by the duration of a transmission cycle, which includes the time intervals between each time unit in the cycle.

$numOfLinks$ equals the number of links of the generate multicast tree.

3.3.2 MOACS

In this subsection, we describe the metaheuristic algorithms for computing the Pareto Front of multicast routing trees in a CRN.

Ant colony optimization (ACO) algorithm is a metaheuristic inspired from the foraging behavior of some ant species. The ants leave pheromone at a path they select. The amount of pheromone on a path will cause the ants to eventually choose the shortest path to reach the destination.

Reference [72] proposed the ACO based MOACS to solve multi-objective multicast problem in wired networks. We adopt this algorithm with modifications to solve our problem. Algorithm 4 shows the general procedure of the modified MOACS.

Given the HyperGraph $H(V_H, E_H)$, we first initialize pheromone matrix τ_{ij} with τ_0 on all edges. In each iteration, with each X_{solu} found by an ant, a known Pareto Front Y_{known} is updated including the best non-dominated solutions that have been found so far. If Y_{known} changes, τ_{ij} will be re-initialized to improve the exploration in the decision space. Otherwise, τ_{ij} is updated using the solutions in Y_{known} to better exploit the knowledge of the best known solutions. In BuildTree() of Algorithm 4, we calculate the probabilities of selecting the next un-visited neighbor node using Equation 3.9. In this equation, τ_{ij} is the pheromone of link (i, j), $C_j = 1/(t_{ij} + t_j)$, t_{ij} is the cost of link (i, j), and t_j is the cost of supernode j. By setting α

Algorithm 4 General procedure of modified MOACS

Input: HyperGraph $H(V_H, E_H)$, Multicast Session(s, Des)
Output: Optimal multicast tree set Y_{known}
Initialize τ_{ij} with τ_0 for all edges, numOfIterations = 0, $Y_{known} = \phi$
while numOfIterations < Iteration Size **do**
    for i from 0 to popSize **do**
        BuildTree($Xsolu$)
        construct transmission schedule of $Xsolu$
        compute delay, rate, numOfLinks of $Xsolu$
        if $Xsolu$ is not dominated by any $Xsolu_k \in Y_{known}$) **then**
            $Y_{known} = Y_{known} \cup Xsolu - \{Xsolu_k | Xsolu \succ Xsolu_k\}, \forall Xsolu_k \in Y_{known}$
        end if
        if Y_{known} was modified **then**
            $\tau_{ij} = \tau_0 \ \forall (i,j) \in E_H$
        else
            for $\forall Xsolu_k \in Y_{known}$ **do**
                $\triangle\tau_{best} = w_1 * delay_k + w_2 * rate_k + w_3 * numOfLinks_k$
                $\tau_{ij} = (1-p)\tau_{ij} + p\triangle\tau_{best} \ \forall (i,j) \in Xsolu_k$
            end for
        end if
    end for
    numOfIterations++
end while

and β, we can adjust the weight of τ_{ij} and C_j in computing the probabilities. A pseudo-random procedure [72] is also used in selecting the next non-visited supernode.

$$p_{ij} = \begin{cases} \frac{\tau_{ij}^{\alpha} C_j^{\beta}}{\sum_{\forall g \in N_i} \tau_{ig}^{\alpha} C_g^{\beta}} & \text{if } j \in N_i \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

3.3.3 AMOSA

AMOSA is an algorithm proposed in reference [73] to use simulated annealing to solve multi-objective optimization problems. We adopt the code provided by reference [73] and modify the process of constructing a random multicast tree solution as introduced in previous subsections. AMOSA will return an archive which records all optimal multicast routing trees that can be found within a given number of iterations.

Notice that AMOSA starts with a set of initial solutions. An archiving process is applied to the initial solution set to select all Pareto optimal solutions within the solution set and put them into an archive. One solution (current solution) is randomly selected from the archive. Then the iteration process starts by generating a new solution from the current solution through mutation. After comparing the new solution with current solution and optimal solutions in the archive found so far, it will decide whether to put the new solution into the archive, drop the new solution or set the new solution as the current solution. The process repeats for a given number of iterations.

3.4 Simulation Result

In this section, we study the performance of MOACS and AMOSA in finding optimal multicast tree solutions in a hypergraph. Due to the dynamic nature of CRNs, we study the performance of MOACS and AMOSA after running them for a few seconds. We study the following three metrics of MOACS and AMOSA: 1) the percentage of solutions in true Pareto Front found by MOACS and AMOSA; 2) the distances of solutions found by MOACS and AMOSA from solutions in true Pareto Front; 3) running time comparison between MOACS and AMOSA.

Table 3.3: Percentages of solutions in APF that MOACS and AMOSA found through 200 and 1000 iterations and average relative distances from APF

Iterations	1,000,000	200	200	1000	1000
Solution Set	Y_{utrue}	Y_{MOACS}	Y_{AMOSA}	Y_{MOACS}	Y_{AMOSA}
Sum of Solutions in APF	62	26	28	39	28
Percentage of Y_{utrue}	—	41.94%	45.16%	62.9%	45.16%
Relative distance from Y_{utrue}	—	6.7%	8.81%	5.14%	9.1%

To approximate the true Pareto Front, we run a Uniformly Random Search Algorithm (URSA) for millions of iterations. The procedure of URSA is very similar to MOACS, except URSA does not consider the influence of pheromone. On a computer with 31 GB memory and 8 Core Intel Xeon CPU E5440@2.83GHz, we run a simulation of 20 randomly generate CRNs. Each network has 20 SUs (with ID from 0 - 19), 3 different channels, and 3 primary users, one for each channel. There are 1 source TDsupernode, 5 destination RDsupernodes for one multicast session in each CRN. For all SUs, we use three different transmission ranges corresponding to three different transmission rates. We set the size of 1 data segment as 1 Mbits. We use Y_{utrue} to denote the approximated true Pareto Front (APF) found using URSA, and Y_{MOACS} and Y_{AMOSA} to denote the solution set found by MOACS and AMOSA, respectively. For these 20 randomly generate CRNs, we run URSA for 1,000,000 iterations to approximate Y_{utrue}, and run MOACS and AMOSA both for 200 iterations and 1000 iterations to compute Y_{MOACS} and Y_{AMOSA}. In Table 3.3, we record sum of solutions in Y_{utrue} of 20 CRNs, sum of solutions in Y_{MOACS} and Y_{AMOSA} that exist in the Y_{utrue} of 20 CRNs, and compute the average percentages of solutions out of Y_{utrue} that Y_{MOACS} and Y_{AMOSA} have and the average relative distances of solutions in Y_{MOACS} and Y_{AMOSA} that are not in Y_{utrue} from solutions in Y_{utrue}.

Of all 20 simulation CRNs, the optimal solutions in APF that AMOSA finds within 200 iterations are the same as in 1000 iterations, and are an average of 45.16% of the APF. Moreover, the optimal solutions found by AMOSA are mostly from the archiving process of the initial solution set, and the iteration process can not improve its solutions significantly. However, MOACS finds more solutions of the APF by performing more iterations. Within 200 iterations,

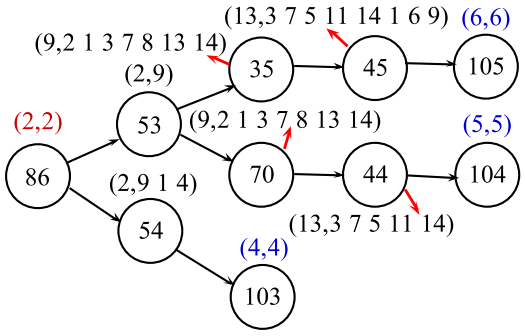
MOACS finds an average of 41.94% of the APF, and it finds an average of 62.9% of the APF within 1000 iterations. To study the closeness of solutions in APF and solutions found by MOACS and AMOSA, we use $(|delay_i - delay_j|/delay_i + |rate_i - rate_j|/rate_i + |numOfLinks_i - numOfLinks_j|/numOfLinks_i)/3$ to compute the relative distance between two solutions i and j, where solution i is in APF. The numbers in the bottom row of Table 3.3 show that the average relative distances of solutions find by AMOSA and MOACS that are not in APF to solutions in APF is small.

Fig. 3.8 shows the running time of AMOSA and MOACS in 20 simulated CRNs within 200 iterations and 1000 iterations, respectively. For all 20 CRNs, the running time of AMOSA and MOACS for 1000 iterations is within 5 seconds, and even less when running them for 200 iterations. Moreover, for most cases, the running time of AMOSA is 1 to 2 seconds longer than the running time of MOACS.

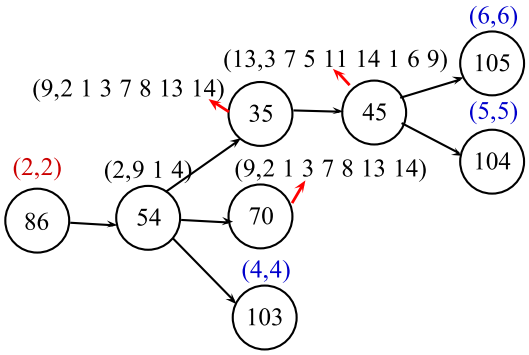
From the results above we can see that AMOSA and MOACS can find most solutions that are within or close to the APF in only a few seconds. However, one advantage of MOACS is that, unlike AMOSA, it is not limited by the annealing temperature, and it can be run continuously, hence adapting to the dynamic nature of CRNs.

3.5 Conclusions

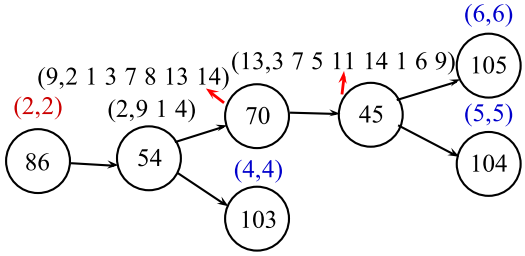
In this work, we proposed a new modeling method to model a CRN into a MMHG. Using this MMHG, we proposed an approach for building multicast trees, which attempting to optimize multiple objectives. We also studied scheduling algorithms to compute the *delay* and *rate* of a multicast tree. We applied two metaheuristic algorithms to find the Pareto Front of multicast routing trees with multiple objectives. Our simulation results show that in small CRNs, MOACS can find over 60% of the APF within only a few seconds, and AMOSA can find approximately 45%. The solutions found that are not in the APF are within 10% relative distances from the solutions in APF. Moreover, MOACS performs better than AMOSA as the speed of AMOSA in finding optimal solutions is limited by the annealing temperature.



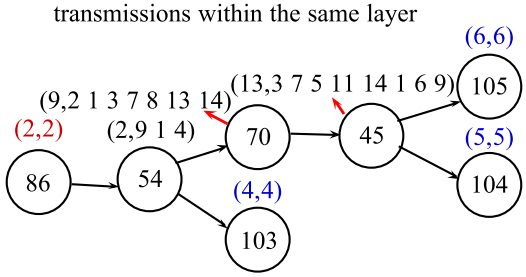
(a) Multicast tree build before ModifyTree()



(b) Multicast tree build after combining multiple transmissions within the same layer



(c) Multicast tree build after combining multiple transmissions within different layers



(d) Multicast tree build after checking all destination reachability

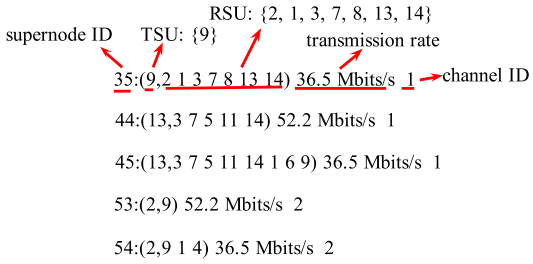


Figure 3.4: A detailed example of ModifyTree()

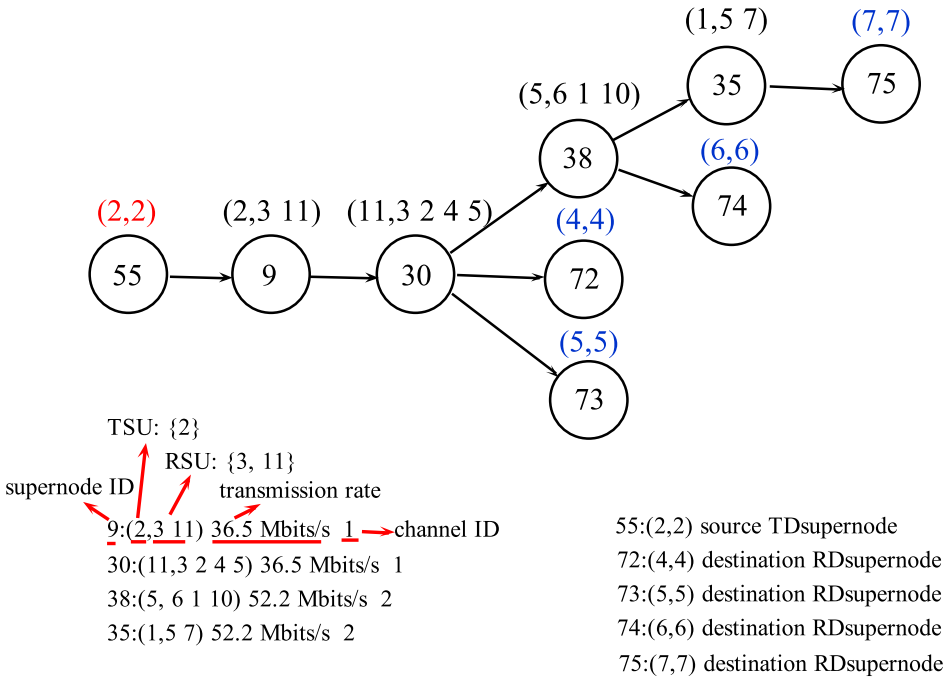


Figure 3.5: A random multicast tree

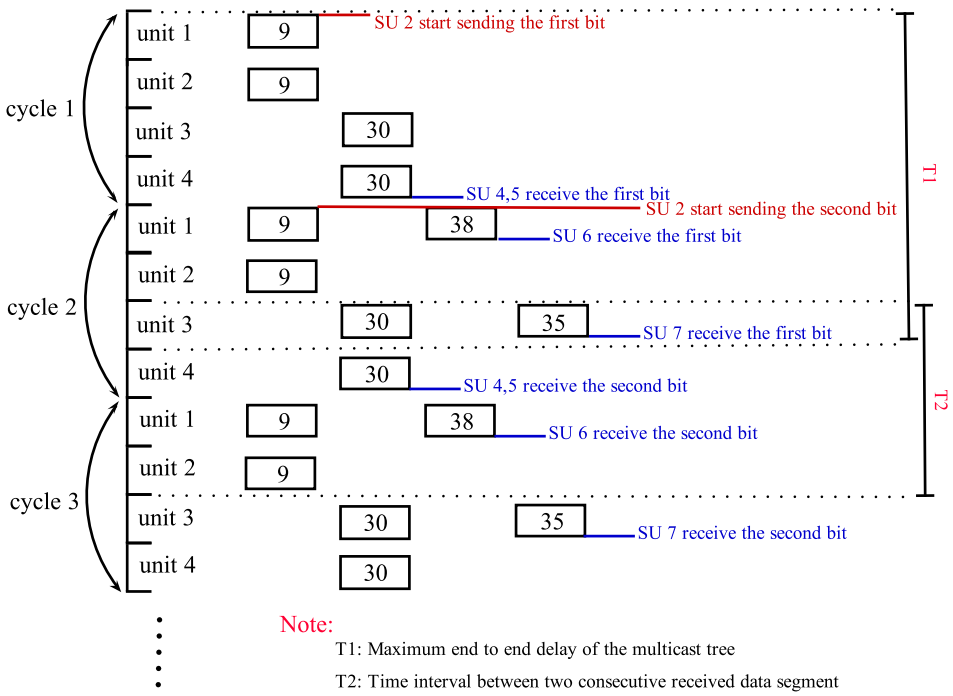


Figure 3.6: Example of transmission procedure of XsoluSche

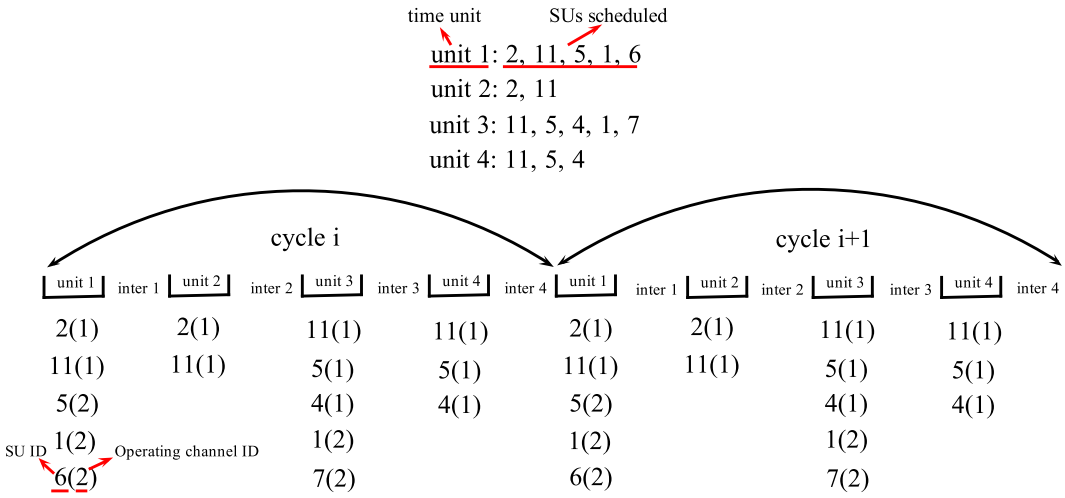


Figure 3.7: Active SUs in the transmission cycle

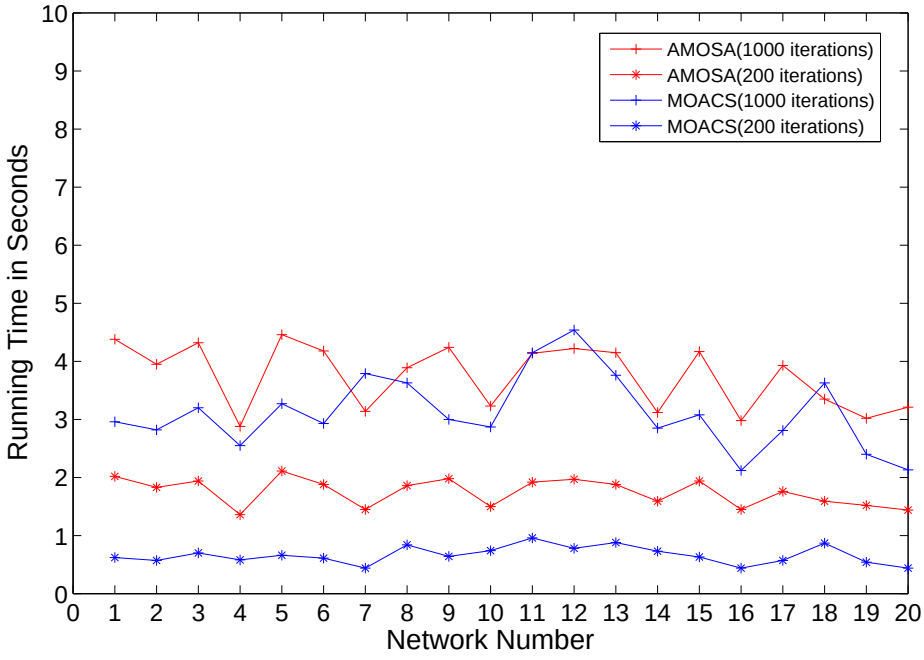


Figure 3.8: Running time comparison between AMOSA and MOACS

CHAPTER 4. A COMP-BASED SELF-HEALING SOLUTION FOR HETEROGENEOUS FEMTOCELL NETWORKS

4.1 System Model

We consider a HetNet where Macro BS (MBS), Pico BSs (PBSs) and FBSs overlay with each other. PBSs are located independently within the coverage of the MBS. FBSs are deployed within a dual strip apartment block [76], which is located within the coverage of MBS. The dual strip apartment (apt) block model is shown in Fig. 4.1. It is assumed that there are 40 apartment units in this apartment block, each apartment unit has an area of 10 by 10 meters, and a given probability of deploying one FBS.

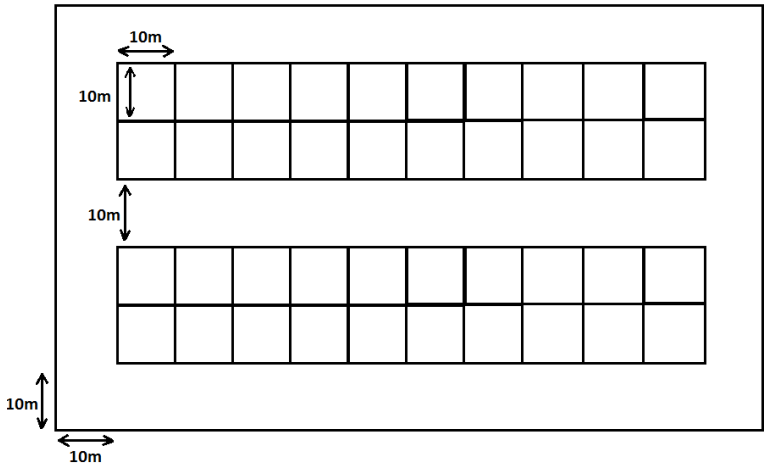


Figure 4.1: Dual stripe apartment block. Each stripe has 2 rows of apartment building. Each row has 10 apartment units. There is a 10 meters wide street separates the two strip apartments.

Fig. 4.2 plots an example of the system model. There are one MBS and one PBS. Users Equipment (UEs) (represented in blue +) are uniformly scattered over the field and some of them are uniformly dropped within the coverage of PBS to simulate a hotspot. There is one

dual strip apt block located within the field, and FBSs are randomly located within the apt units. There are also some UEs randomly deployed within each apartment unit.

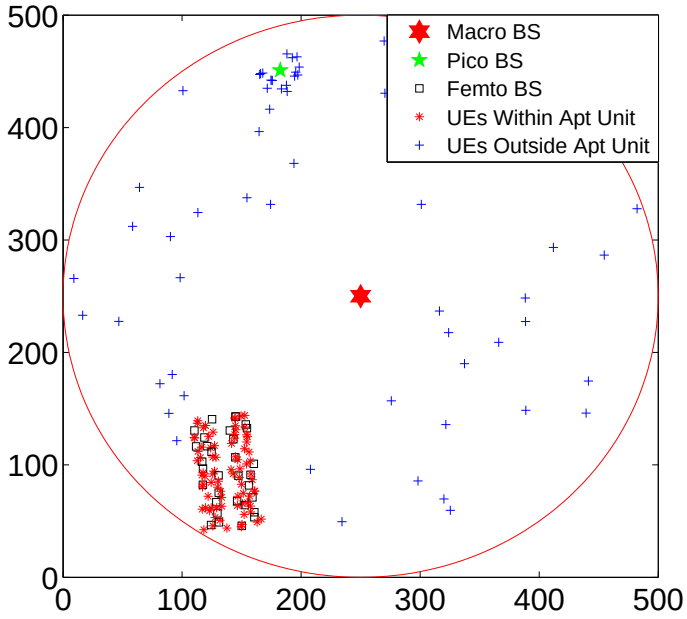


Figure 4.2: A sample plot of system model.

We define UEs located within apt units where a FBS is installed as FUEs. Other UEs are called non-Femto UEs (non-FUEs). FUEs are UEs that subscribed to femto cellular service, and they can be provided with higher data rate and better quality of service through the connections to FBSs. A FUE subscribes to FBS that located in the same unit as the FUE. In practice, FBSs operate in Closed Subscriber Group (CSG) where they can only provide service to UEs subscribed to it. In this work, we assume that FBS operate in hybrid mode where they can also be associated by FUEs that subscribe to other different FBSs, but not by UEs that do not subscribe to femto cellular service at all. A perfect knowledge of the channel gains between the BSs and their served UEs are known to the BSs. Also, we assume that each PBS, FBS, and UE has only one single radio antenna.

4.2 Resource Allocation Problem Formulation

In this section, we formulate the CoMP-JP operation of all FBSs in the system into a resource allocation problem. The received signal at FUE i from FBS j is given by

$$y_{ij} = h_{ij}x_{ij} \tag{4.1}$$

where h_{ij} and x_{ij} are the channel gain between FBS j and FUE i and the signal transmitted from FBS j towards FUE i, respectively. x_{ij} is the transmitted signals after applying a linear precoding scheme on s_i, which is a linear transformation of the information symbols:

$$x_{ij} = t_{ij}s_i \tag{4.2}$$

where s_i and t_{ij} are the information symbol for FUE i and the precoding coefficient that FBS applied to information symbol s_i, respectively. It is assumed that $E\{s_i s_i^H\} = 1$, where $(\cdot)^H$ is the conjugate transpose. We also assume that the information symbols of different UEs are orthogonal with each other.

In order to reduce intercell interference, improve system throughput and provide cell outage compensation capability, we allow FBSs in the apartment to form non-overlapping clusters, and coordinate in providing service to groups of FUEs using CoMP-JP.

A simple CoMP-JP transmission example is shown in Fig. 4.3. The total cellular channel spectrum is denoted by B. FBSs $\{f_1, f_2, f_3, f_4\}$ coorperate to provide service to FUEs $\{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$ using CoMP-JP. Since each FBS and FUE have one radio antenna, a FBS cluster $\mathcal{F} = \{f_1, f_2, f_3, f_4\}$ can only provide service to a set of FUEs g, such that $|\mathcal{F}| \geq |g|$, where $|\mathcal{F}|$ and $|g|$ are the cardinalities of sets $\mathcal{F}$ and g, respectively. In order to serve the 8 FUEs, we need to group the 8 FUEs into at least two sets and allocate to each set different spectrum. Therefore, we assume the FUEs are divided into two sets: $g_1 = \{u_1, u_3, u_6, u_7\}$ and $g_2 = \{u_2, u_4, u_5, u_8\}$ where B_1 and B_2 are allocated to g_1 and g_2, respectively. After the scheduling, FBS cluster $\mathcal{F}$ can coordinate transmitting signals to g_1 on spectrum B_1, and g_2 on spectrum B_2 using CoMP-JP. $\{\mathcal{F}, g_1\}$, $\{\mathcal{F}, g_2\}$ are two CoMP operation units.

Let us assume the FBSs and FUEs in the system have been clustered and grouped into many CoMP operation units. The group of FUEs in one CoMP operation unit that contains FUE i is denoted by $g(i)$. $\mathcal{F}_i$ is the FBS cluster in the CoMP operation unit that provide service to $g(i)$.

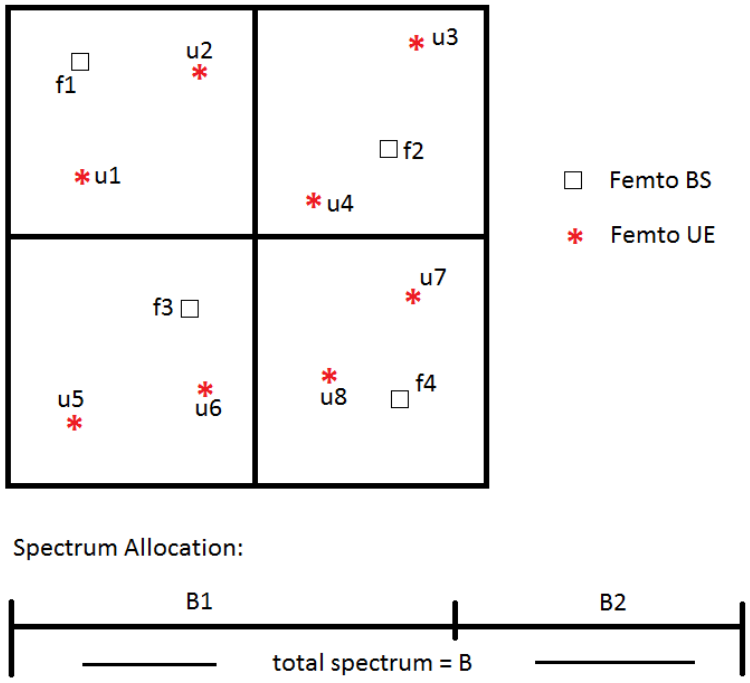


Figure 4.3: A simple FBS CoMP-JP operation example.

Using CoMP-JP, all FBSs within $\mathcal{F}_i$ transmit a linear combination of information symbols for all FUEs in $g(i)$. The transmitted signal of each FBS j within $\mathcal{F}_i$ is given by

$$x_j = \sum_{k \in g(i)} t_{kj} s_k \tag{4.3}$$

the received signal at FUE i is therefore given by

$$y_i = \sum_{j \in \mathcal{F}_i} h_{ij} x_j + Z_i \tag{4.4}$$

$$= \sum_{j \in \mathcal{F}_i} \sum_{k \in g(i)} h_{ij} t_{kj} s_k + Z_i \tag{4.5}$$

where Z_i combines the intercluster interference and noise for FUE i.

By applying zero forcing linear precoding technique on the transmitted signals, the interference to FUE i caused by information symbols of other UEs within $g(i)$ can be cancelled [77]. Therefore, the following should be satisfied when designing the precoding coefficients on transmitted information symbols:

$$\sum_{j\in\mathcal{F}_i}\sum_{k\in g(i),k\neq i} h_{ij}t_{kj}s_k = 0 \tag{4.6}$$

since s_k is orthogonal with s_l, $\forall k,l \in g(i), k \neq l$, we have

$$\sum_{j\in\mathcal{F}_i} h_{ij}t_{kj} = 0 \quad \forall k \in g(i), k \neq i \tag{4.7}$$

by combining (4.5) and (4.7), we have

$$y_i = \sum_{j\in\mathcal{F}_i} h_{ij}t_{ij}s_i + Z_i \tag{4.8}$$

the precoding coefficients are designed such that (4.7) and (4.8) can be satisfied for all FUEs within $g(i)$.

LTE and LTE-A spectrum are assigned in unit of resource blocks. Assume B_0 is the bandwidth of one resource block. $RB_{g(i)}$ denotes the number of resource blocks assigned to FUE set $g(i)$ by FBS cluster $\mathcal{F}_i$. According to Shannon-Hartley theorem, the data rate of FUE i is given by

$$Rate_i = RB_{g(i)}\, B_0 \log_2\left(1 + \frac{\left|\sum_{j\in\mathcal{F}_i} h_{ij}t_{ij}\right|^2}{I_i RB_{g(i)} B_0}\right) \tag{4.9}$$

where I_i is the interference plus noise spectrum density for FUE i.

The operation of each FBS j in the system is subject to power and spectrum constraints. Let $\mathcal{F}(j)$ be a FBS cluster that contains FBS j, and the set of FUEs served by $\mathcal{F}(j)$ are scheduled into different FUE groups $\mathcal{G}_j = \{g_1, g_2, \cdots, g_{|\mathcal{G}_j|}\}$ where $|\mathcal{G}_j|$ is the cardinality of set $\mathcal{G}_j$. Each FUE group $g_k \in \mathcal{G}_j$ together with $\mathcal{F}(j)$ forms one CoMP operation unit. Therefore, $\mathcal{U}_j = g_1 \cup g_2 \cup \cdots \cup g_{|\mathcal{G}_j|}$ is the set of total FUEs that are served by $\mathcal{F}(j)$. In the example of Fig. 4.3, for FBS f_1, $\mathcal{F}(1) = \{f_1, f_2, f_3, f_4\}$, and $\mathcal{G}_1 = \{g_1, g_2\}$, and its total served FUEs set is $\mathcal{U}_1 = g_1 \cup g_2 = \{u1, u2, u3, u4, u5, u6, u7, u8\}$.

Let P_j be the maximum transmit power of FBS j, then

$$\sum_{k\in\mathcal{U}_j} |t_{kj}|^2 \leq P_j \tag{4.10}$$

let RB_{g_k} be the number of resource blocks assigned to the FUE group $g_k \in \mathcal{G}_j$ by the FBS cluster $\mathcal{F}(j)$, then we have:

$$\sum_{g_k \in \mathcal{G}_j} RB_{g_k} \leq RB_m^{max} \quad \forall m \in \mathcal{F}(j) \tag{4.11}$$

where RB_m^{max} is the maximum available number of resource blocks to FBS m .

Finally, our resource allocation problem that maximize FUEs throughput can be formulated as follows:

$$\underset{\substack{t_{ij}, RB_{g_k}, \mathcal{G}_j, g(i), \mathcal{F}(j), \\ \mathcal{F}_i, \forall i \in \mathcal{U}, \forall j \in \mathcal{M}}}{\text{maximize}} \quad \sum_{i \in \mathcal{U}} Rate_i \tag{4.12}$$

$$\text{subject to:} \tag{4.13}$$

$$\sum_{i \in \mathcal{U}_j} |t_{ij}|^2 \leq P_j, \quad \forall j \in \mathcal{M} \tag{4.14}$$

$$\sum_{g_k \in \mathcal{G}_j} RB_{g_k} \leq \min_{m \in \mathcal{F}(j)} RB_m^{max}, \quad \forall j \in \mathcal{M} \tag{4.15}$$

$$\sum_{m \in \mathcal{F}_i} h_{im} t_{km} = 0, \quad \forall k \in g(i), k \neq i, \forall i \in \mathcal{U} \tag{4.16}$$

where $\mathcal{M}$ is the set of all FBSs in the system, $\mathcal{U}$ is the set of all FUEs in the system. Constraints (4.14) and (4.15) are the power constraint and spectrum constraint for each FBS, respectively. While, Constraint (4.16) is the zero-forcing precoding requirement for each FUE. $(\mathcal{F}(j), \mathcal{G}_j)$ and $(\mathcal{F}_i, g(i))$ are (FBS cluster, FUE group) defined from FBS and FUE point of view, respectively. This problem is considered NP-hard mixed integer mathematical programing problem, and it is difficult to solve, especially in practical situations as the size of FBSs and FUEs is large. Hence, in the sequel, we propose an efficient heuristic scheme that decompose the problem into sub-problems and solve each sub-problem separately.

4.3 Heuristic Divide and Conquer Solution

In this section, we propose a heuristic way (divide and conquer) to solve the formulated problem given in (4.12)-(4.16), where we need to find FBSs and FUEs CoMP operation units, precoding coefficients for FBSs and resource blocks allocation to optimize the objective function.

The FBSs and FUEs CoMP operation units formation involves FBS clustering and FUE scheduling. Firstly, we perform FBS clustering and FUE scheduling given the FBSs and FUEs topology and transmission channel conditions. After that, the CoMP operation units can be formed. We then design the precoding coefficients for all FBSs within each CoMP operation unit such that the throughput of the CoMP operation unit is maximized. Finally, spectrum allocation is performed based on the performance of each CoMP operation unit.

4.3.1 FBS Clustering

In this work, we study three different clustering approaches, 1) traffic density based clustering; 2) k-means clustering [78]; 3) spectral clustering [79].

4.3.1.1 Traffic density based FBS clustering

FBS with low transmission power has a transmission range of 20-30 meters. Based on the FBSs-FUEs deployment topology and FBSs' low transmission power feature, we can perform FBS clustering as shown in Fig. 4.4.

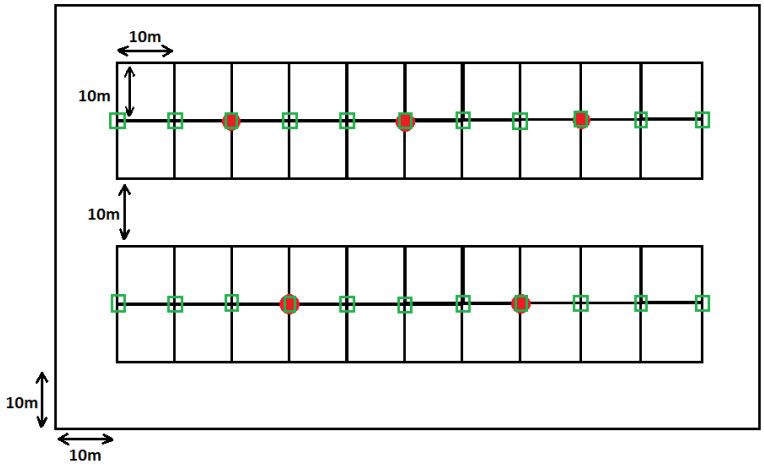


Figure 4.4: A simple FBS clustering example.

The red spots in the apartment model locate in the center of each dense femto service traffic area and their locations will accommodate the traffic dynamics. In order to reduce the search space for the red spots, in this work, we select the locations that are specified by the squares in Fig. 4.4 as the candidates locations for the red spots. The final red spots are selected from

squared points that have the strongest signals from FBSs, what's more, the locations of the red spots should be properly spaced and be at least three to four apartment units away from each other. Fig. 4.4 gives an example of red spots' locations.

After identifying the red spots, the FBS clustering can be implemented in the following two main steps:

Step (1): FBS clustering. Find the c UEs $\{e_1, e_2, \cdots, e_c\}$ that are closest to the c red spots locations, respectively. For each UE $e_i, \forall i \in \{1, 2, \cdots, c\}$, form a FBS cluster $\mathcal{C}_i$ of size S that have the strongest signal power toward UE e_i. S is an integer number given by users. Let $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_c\}$ be the set of clusters in the system.

Step (2): FBS clustering refinement. Step 1 may generate clusters that have the same FBS existing in two or more clusters, or have some FBSs not assigned to any cluster. If there exists FBS j that satisfies either of these two conditions, assign it to cluster $\mathcal{C}_i$, where it has the strongest average per-UE channel gain towards UEs located in the same apartment units of FBSs in $\mathcal{C}_i$.

$$\mathcal{C}_i = \arg\max_{\mathcal{C}_i \in \mathcal{C}} \frac{\sum_{k \in \mathcal{U}(\mathcal{C}_i)} |h_{kj}|^2}{|\mathcal{U}(\mathcal{C}_i)|} \tag{4.17}$$

where $\mathcal{U}(\mathcal{C}_i)$ is the set of UEs located within apartment units that deploy FBSs of $\mathcal{C}_i$. $|\mathcal{U}(\mathcal{C}_i)|$ is the cardinality of set $\mathcal{U}(\mathcal{C}_i)$. h_{kj} is the channel gain of FBS j towards UE k.

4.3.1.2 k-Means FBS clustering

k-Means clustering is to partition n points into k clusters, where each point belongs to a cluster with the minimal mean objective value. In our case, the points to be clustered are the femto base stations, and we use euclidean distance between base stations to be the clustering objective value. So after apply k-Means FBS clustering algorithm, each FBS belongs to a cluster with the minimal mean Euclidean distance. The algorithm is shown in Algorithm 5, where input *Loc* is n FBS's locations represented in (x axis, y axis) coordinate, output *OptCresult* is the optimal clustering result, c is the number of clusters and kmeans(Loc, c) is a function in MATLAB which implemented the k-means clustering algorithm. In the algorithm, we loop through different value of c, and select the one that gives the maximum FUE throughput.

Algorithm 5 K-Means Clustering Algorithm

Input: n by 2 matrix *Loc*
Output: n by 1 vector *OptCresult*
$OptR_F = 0$
for c = 2 : n **do**
    Cresult = kmeans(Loc, c);
    Perform FUE scheduling (4.3.2), and compute the Femto UEs throughput R_F (introduced in section 4.3.3 and 4.3.4)
    if $R_F > OptR_F$ **then**
        $OptR_F = R_F$
        OptCresult = Cresult
    end if
end for
return *OptCresult*

4.3.1.3 Spectral clustering

Even though Euclidean distance is a very important factor in deciding similarities between points, however, there are other factors that are important as well, for example, channel gain between two points. Similarity graph defines similarity matric between points, and the similarity matric is a function of the factors we want to study. Reference paper [79] introduced three different similarity graph representations: the ε-neighborhood graph, k-nearest neighbor graphs and the fully connected graph. Spectral clustering makes use of the eigenvalues of the similarity matrix of the data to perform clustering. Spectral clustering is considered as one of the most popular modern clustering algorithms and outperforms traditional clustering algorithms such as the k-means algorithm very often.

In this work, we transforms the FBS network to a graph $G(\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the set of FBSs and $\mathcal{E}$ is the set of links between FBSs. We use the fully connected graph, where we connect all FBSs with positive similarities, and weight all edges by the Gaussian similarity function of signal strength between two nodes:

$$s_{ij} = \begin{cases} \exp(-\frac{\|SS(i,j)\|^2}{2\delta_h^2}) & i \neq j \\ 0 & i == j \end{cases}$$

where $SS(i, j)$ is the signal strength (dB) from node i to node j when i is transmitting with the maximum power over maximum spectrum bandwidth. Signal strength represents not only the

physical distance between two nodes, but also the channel condition between two nodes. The higher the signal strength, the more similar the two nodes are. Note that the signal strength is measured in dB. In our simulation setup, the signal strength is always a negative value in dB.

The spectrum clustering algorithm is shown in Algorithm 6, where similarity matrix S is n by n matrix with each element s_{ij} defined above, and degree matrix D is defined as the diagonal matrix with degrees $d_1, ..., d_n$ on the diagonal, and d_i is defined as:

$$d_i = \sum_{j=1}^{n} s_{ij} \tag{4.18}$$

Algorithm 6 Spectrum Clustering Algorithm

Input: similarity matrix S and degree matrix D
Output: n by 1 vector $OptCresult$
Compute the unnormalized Laplacian $L = D - S$.
$OptR_F = 0$
for c = 2 : n **do**
    Compute the first c eigenvectors $v_1, ..., v_c$ of L
    Let $V \in \mathbb{R}^{n \times c}$ be the matrix containing the vectors $v_1, ..., v_c$ as columns.
    Cresult = kmeans(V, c);
    Perform FUE scheduling (4.3.2), and compute the Femto UEs throughput R_F (introduced in section 4.3.3 and 4.3.4)
    if $R_F > OptR_F$ **then**
        $OptR_F = R_F$
        OptCresult = Cresult
    end if
end for
return $OptCresult$

4.3.2 Femto UE Scheduling

After FBS clustering, $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_c\}$ is the set of FBS clusters in the system. Define $\mathcal{U}_i$ as the set of FUEs assigned to be served by FBS cluster $\mathcal{C}_i$, and an anchor FBS [80] of a FUE k as the FBS that has the strongest signal power towards the UE. Then $\mathcal{U}_i$ is the set of FUEs that their anchor FBSs are within the FBS cluster $\mathcal{C}_i$. For each $\mathcal{C}_i \in \mathcal{C}$, divide its UE set $\mathcal{U}_i$ into subgroups $\mathcal{G}_i = \{g_1, g_2, \cdots, g_{|\mathcal{G}_i|}\}$, with each subgroup having a UE size which is the same as the FBS size of the cluster $\mathcal{C}_i$:

$$|\mathcal{G}_i| = \left\lceil \frac{|\mathcal{U}_i|}{|\mathcal{C}_i|} \right\rceil \tag{4.19}$$

The dividing process of $\mathcal{U}_i$ is random, and we just need to make sure the maximum subgroup size is $|\mathcal{C}_i|$. The size of one subgroup can be less than $|\mathcal{C}_i|$. Now FUE set $\mathcal{U}_i$ is scheduled into $|\mathcal{G}_i|$ subgroups, and each subgroup $g_k \in \mathcal{G}_i$ together with the FBS cluster $\mathcal{C}_i$ form one CoMP operation unit. There is a total number of $\sum_{i\in\{1,2,\cdots,c\}} |\mathcal{G}_i|$ CoMP operation units in the system.

4.3.3 Power Allocation and Precoding Coefficients Calculation

After FBS clustering and FUE scheduling, we reformulate problem (4.12)-(4.16) as follows

$$\underset{\substack{t_{km},\, RB_{g_j},\, \forall k \in g_j,\\ \forall g_j \in \mathcal{G}_i, \forall m \in \mathcal{C}_i,\\ \forall \mathcal{C}_i \in \mathcal{C}}}{\text{maximize}} \quad \sum_{\mathcal{C}_i \in \mathcal{C}} \sum_{g_j \in \mathcal{G}_i} RB_{g_j} B_0 \sum_{k \in g_j} \log_2 \left(1 + \frac{\left|\sum_{m\in\mathcal{C}_i} h_{km} t_{km}\right|^2}{I_k RB_{g_j} B_0}\right) \tag{4.20}$$

subject to:

$$\sum_{g_j \in \mathcal{G}_i} \sum_{k \in g_j} |t_{km}|^2 \leq P_m, \quad \forall m \in \mathcal{C}_i, \forall \mathcal{C}_i \in \mathcal{C} \tag{4.21}$$

$$\sum_{g_j \in \mathcal{G}_i} RB_{g_j} \leq \min_{m \in \mathcal{C}_i} RB_m^{max}, \quad \forall \mathcal{C}_i \in \mathcal{C} \tag{4.22}$$

$$\sum_{m \in \mathcal{C}_i} h_{km} t_{lm} = 0, \forall l \in g_j, l \neq k, \forall k \in g_j, \forall g_j \in \mathcal{G}_i, \forall \mathcal{C}_i \in \mathcal{C} \tag{4.23}$$

The problem is still nonconvex, mixed integer mathematical programming problem. We can further decompose the problem. We first consider solving the above problem by setting $RB_{g_j} = 1, \forall g_j \in \mathcal{G}_i, \forall i \in \{1, 2, \cdots, c\}$, where problem (4.20)-(4.23) becomes to design precoding coefficients of all FBSs to maximize the sum rate of all CoMP operation units in the system on one resource block. We then do resource allocation on each CoMP operation unit later on.

After setting the $RB_{g_j} = 1, \forall g_j \in \mathcal{G}_i, \forall i \in \{1, 2, \cdots, c\}$, the resulting problem can be divided into $\sum_{i\in\{1,2,\cdots,c\}} |\mathcal{G}_i|$ subproblems, and each subproblem is a precoding coefficient design problem on FBSs in one CoMP operation unit. The problem of maximizing the throughput of one CoMP operation unit (FBS cluster $\mathcal{C}_i$, FUE group $g_j \in \mathcal{G}_i$) on one resource block is as follows:

$$\underset{\substack{t_{km}, \forall k \in g_j,\\ \forall m \in \mathcal{C}_i}}{\text{maximize}} \quad \sum_{k \in g_j} B_0 \log_2 \left(1 + \frac{\left|\sum_{m\in\mathcal{C}_i} h_{km} t_{km}\right|^2}{I_k B_0}\right) \tag{4.24}$$

subject to:

$$\sum_{k \in g_j} |t_{km}|^2 \leq \frac{P_m}{RB_m^{max}}, \quad \forall m \in \mathcal{C}_i \tag{4.25}$$

$$\sum_{m \in \mathcal{C}_i} h_{km} t_{lm} = 0, \quad \forall l \in g_j, l \neq k, \forall k \in g_j \tag{4.26}$$

In the above maximization problem, we assume that the FBS spreads its power on its available resource blocks evenly. Therefore, equation (4.25) has a power constraint of $\frac{P_m}{RB_m^{max}}$. In our model, we assume $RB_m^{max}, \forall m \in \mathcal{C}_i, \forall \mathcal{C}_i \in \mathcal{C}$ are the same, and are defined as RB^{max}. I_k is assumed as the thermal noise plus interferences spectrum density. The interference cames from BSs outside $\mathcal{C}_i$ transmitting at their maximum power. Problem (4.24)-(4.26) can be converted to a convex optimization problem according to reference [77]. Given perfect knowledge of the channel conditions between FBSs and FUEs, we can solve problem (4.24)-(4.26) using CVX [81] or other convex problem solvers.

4.3.4 Resource Allocation

We consider proportional fairness on the resource allocation problem. For a FBS cluster $\mathcal{C}_i$, the resource blocks scheduled to each UE group $g_j \in \mathcal{G}_i$ are proportional to spectrum efficiency sum of all UEs in each group g_j. Let SE_{g_j} be spectrum efficiency sum of all UEs within group g_j. Then

$$SE_{g_j} = \sum_{k \in g_j} \log\left(1 + SINR_k\right) \qquad \forall g_j \in \mathcal{G}_i \tag{4.27}$$

where $SINR_k$ is the signal to interference plus noise ratio of FUE k. The number of resource blocks assigned to each group of FUEs is

$$RB_{g_j} = \left\lfloor RB^{max} \frac{SE_{g_j}}{\sum_{g_k \in \mathcal{G}_i} SE_{g_k}} \right\rfloor \qquad \forall g_j \in \mathcal{G}_i \tag{4.28}$$

The remaining resource blocks $RB^{max} - \sum_{g_j \in \mathcal{G}_i} RB_{g_j}$ will be added to each subgroup in a round-robin way.

4.4 Cell Outage Compensation Solution

In our CoMP operation unit formation, different clusters of FBSs are formed independently from each other, and the FUEs scheduled within one CoMP operation unit are also nonoverlapping from other FUEs scheduled in another CoMP operation unit. Therefore, if any failure happens to a FBS, it will affect the operation of the CoMP operation units it works in. What's more, it will reduce the inter-cluster interference suffered by FUEs within other CoMP operation units.

To perform quick failure compensation, the affected CoMP operation units will maintain their FBS clustering groups unchanged, but re-perform FUEs scheduling, precoding coefficients calculation as well as resource allocation as introduced in previous sections. For example, in Fig. 4.3, we have two CoMP operation units: $\{\mathcal{C}_1, g_1\}$, $\{\mathcal{C}_1, g_2\}$, where $\mathcal{C}_1$ is the FBS cluster $\{f_1, f_2, f_3, f_4\}$, g_1 and g_2 are the FUEs groups $\{u_1, u_3, u_6, u_7\}$ and $\{u_2, u_4, u_5, u_8\}$, respectively. When failure happens to FBS f_1, to perform failure compensation, we maintain the FBS cluster, which is $\mathcal{C}_1 = \{f_2, f_3, f_4\}$, and we re-group the FUEs $\{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$ using the proposed scheduling scheme in Section IV. In this example, we re-group UEs to three groups: $g_1 = \{u_1, u_3, u_7\}, g_2 = \{u_2, u_6, u_8\}, g_3 = \{u_4, u_5\}$. Therefore, three different CoMP operation units are formed: $\{\mathcal{C}_1, g_1\}$, $\{\mathcal{C}_1, g_2\}$,$\{\mathcal{C}_1, g_3\}$. Each of these CoMP operation units also has to perform precoding coefficients calculation as well as resource allocation before they can operate successfully.

There are two different cases to consider in real implementation of failure compensation:

Case 1: in the FBS cluster that suffers from failure, at least one FBS is still functional. The compensation operation is as the example shown above: *keep the FBS cluster unchanged, and re-perform UE scheduling, precoding coefficients calculation and resource allocation.*

Case 2: in the FBS cluster that suffered from failure, all FBSs fail. Then *the FUEs served by these FBSs will be handed over to the MBS or PBS with the strongest signal strength.*

4.5 Numerical Results

In this section, we study the performance of our proposed CoMP-JP operation scheme. In the MATLAB simulation platform, we set up one MBS, one PBS and one dual strip apartment block. FBSs are deployed in the apartment block, and each apt unit has 0.8 probability of installing one FBS. We deploy two UEs in each of the apartment units (80 UEs in total located within the apartment block). There are another 60 UEs deployed outside the apartment block. UEs deployed outside the apartment block or deployed in apartment units that have no FBS installed can only connect to MBS or PBS.

The carrier center frequency used in the simulation is 2GHZ, and the total spectrum bandwidth is 10MHZ. The transmission powers for FBS, PBS and MBS are 20 dBm, 37 dBm, and 46 dBm, respectively and the antenna gains for FBSs, PBS and MBS are 5 dBi, 5 dBi, and 15 dBi, respectively. The UE antenna gain is 0 dBi. We use a noise power spectrum density of -174 dBm/HZ. The propagation loss model, shadowing fading model as well as the penetration loss model between BSs and UEs are taken from reference [76]. Please refer to reference article for details.

We assume that there is one Center Management Unit (CMU) with a strong calculation capability located within the apartment block. It connects to each of the FBSs through wired fiber. The FBSs pass the channel gain information between themselves and FUEs to the CMU. The CMU will perform calculations needed to perform the CoMP-JP operation scheme proposed in Section IV. After the calculation, it will pass the results back to FBSs. When a failure happens, the CMU will be notified automatically, and it performs failure compensation calculation as needed, then pass the results back to FBSs. FBSs will operate based on the information sent back by the CMU.

For comparison purpose, we also apply two non-cooperation solutions with the same simulation environment:

NonCoMP-ALL: in normal operation mode, a FUE will associate with the FBS located within the same apt unit as the FUE. Other non-FUEs will associate with the MBS or PBS with the highest signal strength. When a FBS failure happens, the FUEs suffering from failure will

be handed over to the BS (FBSs that are still functional, PBS or MBS) that provide them with the strongest signal power strength.

NonCoMP-NF: in normal operation mode, the UEs behave the same as in NonCoMP-ALL. However, when a FBS failure happens, the FUEs suffering from failure will only be handed over to MBS or PBS with highest signal strength. This comparison solution is to mimic most cases in practice, where FBSs work in CSG, and they only allow access from FUEs that belonging to the same owner(s) as the FBSs.

We assume the system environment is quasi-static (the devices locations and the channel gains remain unchanged during the process of calculation). For a given simulated system environment (deployment of the BSs and UEs in the system and realization of the channel conditions between different devices), we apply our proposed CoMP-JP operation schemes with three different clustering techniques (CoMP-Density, CoMP-kMeans, CoMP-spectral) and two non-cooperation solutions NonCoMP-ALL and NonCoMP-NF under different FBS failure probabilities, and calculate the achievable rate of each UE within the system. We do not consider the cases that MBS and PBS suffer from failure. We simulate 100 different system environments and calculate the sum rate of different groups of UEs in the system averaged over the 100 different simulated environments.

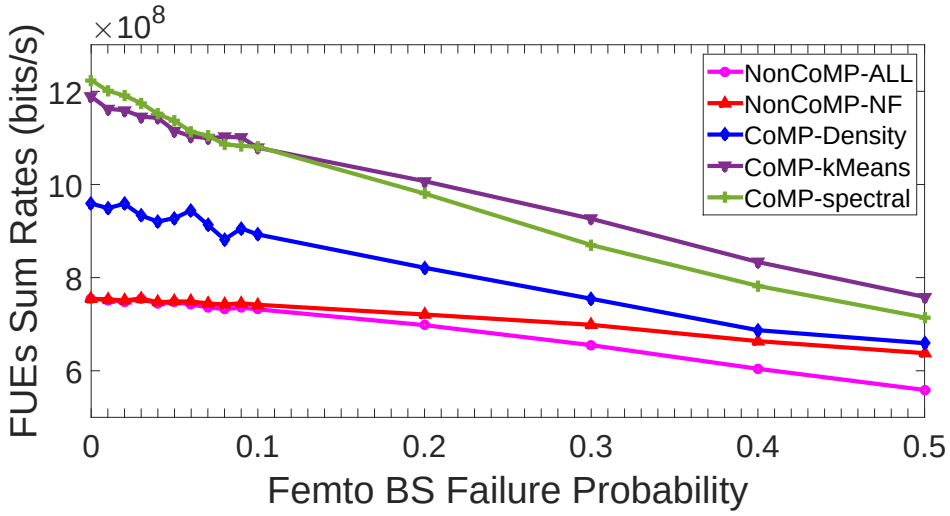


Figure 4.5: Sum rates of FUEs within the residential apartment units under different FBS failure probabilities.

Fig. 4.5 gives the results of sum rates of FUEs under different FBS failure probabilities. CoMP solutions show good advantage over Non-CoMP solutions in providing high sum rates, and they have at least 30% performance improvement compared to NonCoMP solutions. CoMP-kMeans and CoMP-spectral have around 50% rate improvement, and CoMP-spectral performs the best over all other solutions with FBS failure probability smaller than 8%. However, as the failure probability increases, the performance improvement decreases.

Table 4.1: FUE average rate loss with different FBS failure probability

Failure Probability	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07
Rate Loss (CoMP-Density)	0	0.010	0.000	0.027	0.040	0.033	0.015	0.048
Rate Loss (CoMP-kMeans)	0	0.022	0.025	0.036	0.038	0.062	0.072	0.075
Rate Loss (CoMP-spectral)	0	0.019	0.027	0.040	0.059	0.071	0.090	0.097
Failure Probability	0.08	0.09	0.1	0.2	0.3	0.4	0.5	-
Rate Loss (CoMP-Density)	0.081	0.056	0.069	0.144	0.213	0.284	0.313	-
Rate Loss (CoMP-kMeans)	0.073	0.073	0.092	0.153	0.221	0.299	0.362	-
Rate Loss (CoMP-spectral)	0.112	0.115	0.117	0.199	0.289	0.361	0.416	-

Table 4.1 lists the percentage of FUEs sum rate loss under different FBS failure probabilities when applying our proposed CoMP-JP operation schemes. The results show that when FBS failure probability is up to 50%, the FUEs sum rate has about 30% - 42% rate loss. Hence, our proposed schemes can prevent rate loss from having the same speed of radio resource loss, mostly by using available FBS resources within the system. What's more, NonCoMP-spectral has higher FUE rate loss speed. This is because that the resulted cluster size is smaller when using spectral clustering algorithm. With the increasing FBS failure probability, it has higher probability that all FBSs within the same cluster fail, and results FUEs being handed over to the MBS or PBS and served with lower rate.

On the other hand, using available FBS resources for compensation instead of using MBS or PBS radio resources can prevent other non-FUEs from rate loss, but will cause other FUEs not suffering from failure suffer from rate loss. NonCoMP-NF compensates the rate loss by handing all FUEs suffering from failure to MBS and PBS. Therefore, with NonCoMP-NF

scheme, non-FUEs associated with MBS or PBS suffer from great rate loss as shown in Fig. 4.6. NonCoMP-ALL allows FUEs suffering from failure to be handed over to other FBSs as well as MBS and PBS, hence non-FUEs does not suffer rate loss as bad as NonCoMP-NF scheme. However, NonCoMP-NF outperforms NonCoMP-ALL in terms of FUEs' throughput as shown in Fig. 4.5, which suggests that using available FBS resources to compensate rate loss without applying CoMP-JP, will cause more rate loss on FUEs sum rate than just using MBS/PBS resources.

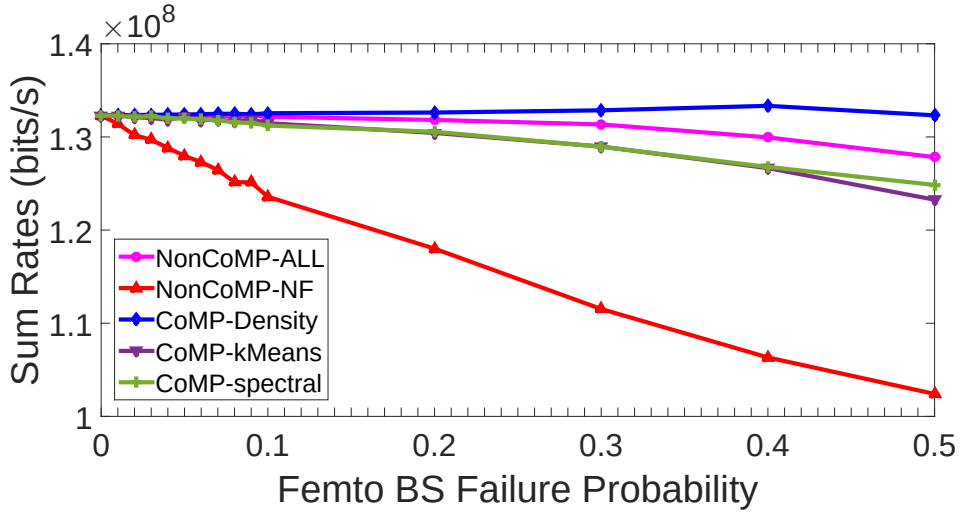


Figure 4.6: Sum rates of non-FUEs outside the residential apartment units under different FBS failure probabilities.

Fig. 4.6 shows the results of sum rate of non-FUEs outside residential area. These UEs are served by MBS and PBS, and do not suffer from any failure. The failure of FBSs will decrease the interference suffered by these non-FUEs. However, due to the failure of FBSs, more FUEs are handed over to MBS and PBS, resulting in decreased resource blocks allocated to each UEs served by MBS/PBS, hence resulting in rate loss. The decreasing trend of those rates with different failure probabilities suggests that the non-FUEs rate loss due to reduction of resource blocks per UE is more severe than the rate increment offered by decreased interference suffered from FBSs. Notice that non-FUEs have lower throughput when applying CoMP-kMeans and CoMP-spectral schemes compared to CoMP-Density and NonCoMP-ALL. This is because kMeans and spectral clustering algorithms result in FBS clusters with smaller size.

Hence with increasing FBS failure probability, it has higher chance that all FBSs within the same cluster fail, and results in more FUEs being handed over to the MBS or PBS, which results in higher rate loss for non-FUEs.

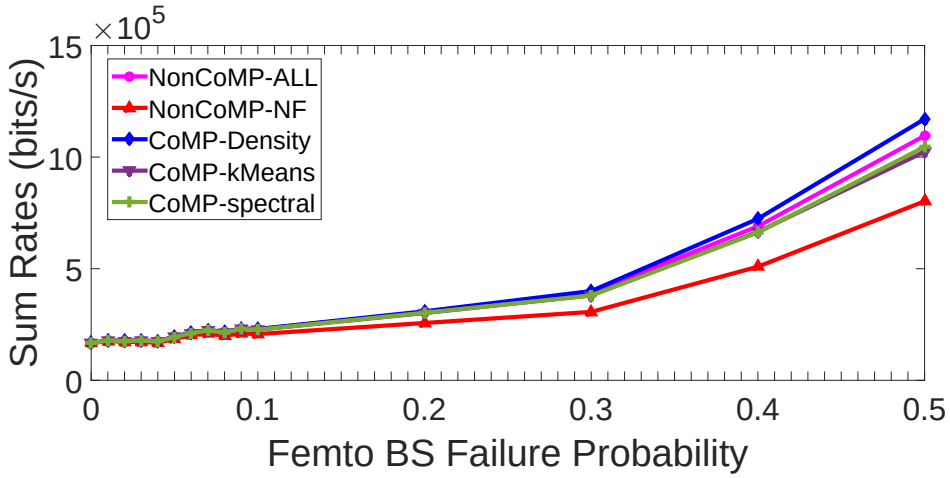


Figure 4.7: Sum rates of non-FUEs within the residential apartment units under different FBS failure probabilities.

Fig. 4.7 shows that the sum rate of non-FUE subscribers within residential area increases with the increment of FBS failure probability. The increasing trend suggests that due to FBS failure, the suffered interference by non-FBSs within the residential apartment units will decrease greatly. NonCoMP-NF gives lower sum rate than CoMP solutions and NonCoMP-ALL, because when a failure of FBS happens, NonCoMP-NF will hand over all FUEs suffering from failure to MBS and PBS, and results in more resource blocks reduction of non-FBSs within residential area, hence lowest sum rates.

Fig. 4.8 shows the total system throughput with different FBS failure probabilities. The sum rate of all UEs in the system have the similar trend as the FUEs sum rate shown in Fig. 4.5. Comparing the sum rate in Fig. 4.5, Fig. 4.6 and Fig. 4.7, we see that the sum rate of FUEs is much higher than the sum rate of non-FUEs outside residential apartment units, and even much higher than the sum rate of non-FUEs within the residential apartment units. This comparison shows that FBSs can provide higher rates than MBSs and PBSs, since FBS serves way less UEs than MBS and PBS. Therefore they support high rate consumption applications on mobile devices. This justifies the increasing popularity of deployment of small cells in reality.

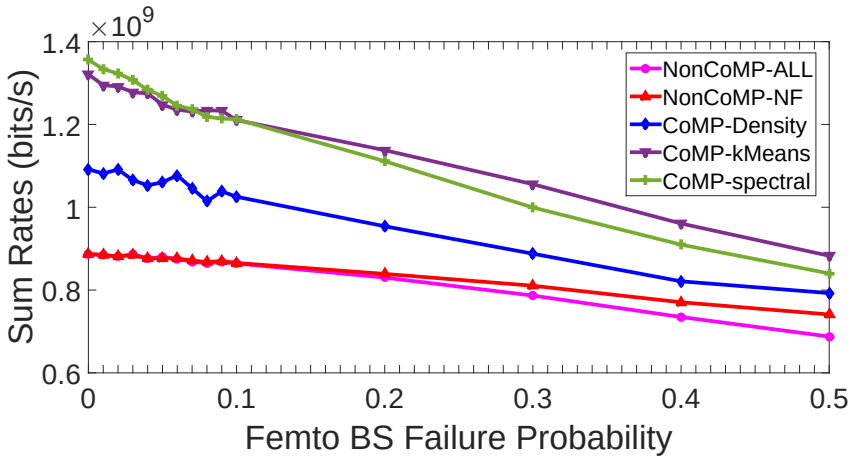


Figure 4.8: System throughputs of all UEs under different FBS failure probabilities.

4.6 Conclusions

In this work, we proposed a CoMP-JP operation scheme for FBSs to perform during both normal operation mode and cell outage compensation mode. We have also studied the performance of three different clustering algorithms applied during FBSs clustering process. The simulation results show that spectral clustering and k-Means clustering algorithms show more advantages in providing better clustering result than the traffic density based clustering algorithm, and spectral clustering algorithm has the best performance when FBS has low failure probability. By comparing our proposed CoMP-JP operation schemes to two other non-CoMP solutions, our proposed scheme shows great performance improvement. The schemes also prove its failure compensation capability by preventing the system sum rate loss from having the same speed of radio resource loss, and this is done without using additional radio resources and will not impair much of the performance of other UEs.

CHAPTER 5. COOPERATIVE MULTI-PATH ROUTING SOLUTION FOR STREAMING APPLICATION USING AUCTION THEORY

5.1 System Model

The system models a heterogeneous wireless network environment as shown in Fig. 5.1. A Macro Base Station (MBS) locates in the center of the cell, and a number of Femto Base Stations (FBS) overlaid within MBS's transmission range. We assume FBS operates in open group access mode, which means UEs can be served by any BSs in the system that provide it with the maximum received signal strength. What's more, The common cellular spectrum is evenly distributed to UEs that associated with BSs (whether it is associated with FBS or MBS). Therefore, there is no interference to UEs's cellular link. A number of User Equipments (UE) are deployed within the transmission range of the MBS. For each FBS, there are a number of UEs deployed within its transmission range as well. In this work, we assume UEs are enabled with device to device (D2D) communication technology and UEs' streaming application supports multi-path transport layer protocol. Therefore, when a UE i starts a streaming event, it first initiates an auction process, in which the idle neighbor UEs within i's transmission range participate to bid for relaying traffic for UE i. The winning neighbor UE helps relaying traffic for UE i, and UE i pays certain amount of credits to the wining neighbor UE for the service.

Fig. 5.1 illustrates four different multipath streaming scenarios using device to device relay transmission.

- **Scenario (a)**: UE e1 establishes two-path streaming traffic. One path is e1's cellular link connected to FBS f1, and the other path is its Wi-Fi link connected to UE e2. Both e1 and e2 are served by FBS f1.

- **Scenario (b)**: Similar to scenario (a), e2 relays traffic for e1 through Wi-Fi link, however,

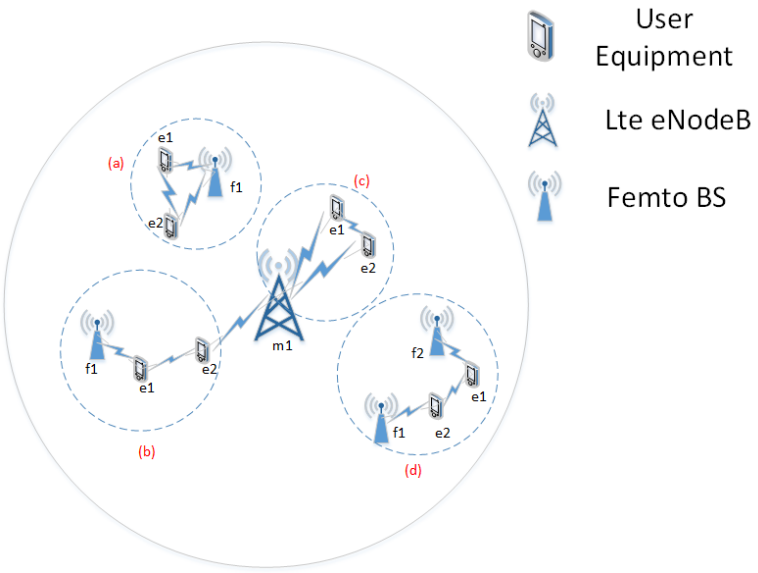


Figure 5.1: Network model

e1 connects to FBS f1, and e2 connects to MBS m1.

- **Scenario (c)**: UE e2 relays traffic for e1 through Wi-Fi link, however, both e1 and e2 connects to MBS m1.
- **Scenario (d)**: e1 and e2 connects to different FBSs, and e2 relays traffic for e1 through Wi-Fi link.

In this work, we assume UEs belong to different selfish users, therefore, UEs have no obligation in forwarding other UEs' traffic unless they enjoy positive utility with the payment made to them. We denote UEs with streaming event as auctioneers, and their neighbor idle UEs that within auctioneers' transmission range as bidders. The bidders report their transmission cost related information to the auctioneers, as if the bidders report their bids to auctioneers. We assume information is encoded and not known to other UEs (sealed bid). Auctioneers then run the auction following the rules proposed in Section 5.3. Our objectives of this work are as follows: 1) propose an operation scheme for auctioneers such that their service rate is maximized and the total operation power consumption can be minimized (discussed in Section 5.2 and 5.3); 2) propose real-time optimization and failure protection operation for auctioneers that further improve UEs' performance (discussed in Section 5.4).

To simplify this work, we assume the system operates in a timeline with a constant time step and all active UEs are connected to the BS whichever provides the strongest received signal power. At the beginning of each time slot, auctioneers broadcast message to neighbor UEs. The bidders that receive the request and are capable of relaying data report information back to the auctioneers. The auctioneers then decide the operation solutions based on the received information from bidders.

We assume each bidder i reports $Info_i$ to the auctioneer o, where $Info_i = \{H_{iLte}(d_i), H_{oiWifi}(d_{io}), P_{imaxLte}, k_i, \alpha_{i0}, \alpha_{i1}, E_{iFull}, \theta_i, E_i\}$. Each element in $Info_i$ is explained below:

- $H_{iLte}(d_i), H_{oiWifi}(d_{io}), P_{imaxLte}$: $H_{iLte}(d_i)$ is the channel gain in dB between bidder i and the BS it associated with. $H_{oiWifi}(d_{io})$ is the channel gain in dB between bidder i and auctioneer o. $P_{imaxLte}$ is the maximum transmission power of bidder i.

- $k_i, \alpha_{i0}, \alpha_{i1}, \theta_i, E_{iFull}$: in reality, UEs with lower energy balance are less likely to provide service. Therefore, we assume the energy price of every UE is a linear function of its current energy balance. Let $\mathcal{A}_i = \{e_k : e_k > \theta_i E_{iFull}\}$ be the set of energy balances larger than a threshold, then bidder i's energy unit price $\alpha_i = \mathbb{I}_{\mathcal{A}_i}(e_i)\alpha_{i0} + (1 - \mathbb{I}_{\mathcal{A}_i}(e_i))\alpha_{i1} + k_i e_i$, where $k_i < 0$, and

$$\mathbb{I}_{\mathcal{A}_i}(e_i) = \begin{cases} 1 & e_i \in \mathcal{A}_i \\ 0 & e_i \notin \mathcal{A}_i \end{cases}$$

  e_i is UE i's current energy balance, E_{iFull} is the energy balance when UE i is fully charged, θ_i is a threshold $\in (0, 1)$ and α_{i1} is a very large number close to infinity. k_i is the slope of the linear function. α_{i0} is a parameter controls the energy unit price when $e_i \in \mathcal{A}_i$. The cost function indicates that when UE i's current energy balance is below θE_{iFull}, it can not participate in the auction as bidders. This is to prevent UE from draining out its battery by acting as relay. The total energy cost when e_i of UE i drops from energy balance E_1 to energy balance E_2 is calculated as follows:

$$\int_{e_i=E_2}^{E_1} \Big(\mathbb{I}_{\mathcal{A}_i}(e_i)\alpha_{i0} + (1 - \mathbb{I}_{\mathcal{A}_i}(e_i))\alpha_{i1} + k_i e_i \Big) de_i$$

- E_i: bidder i's remaining energy at the beginning of the time slot.

With $Info_i$, auctioneer calculates the achievable rate through the cooperation communication with bidder i. When the achievable rate satisfies the lowest rate requirement, auctioneer optimizes the streaming traffic allocation between itself and bidder i to minimize total energy cost of the cooperation transmission. The formulation of this optimization problem is described in the next section.

5.2 Cooperative Multi-path Routing Design

In this section, we introduce auctioneers's operations after they received information reported from bidders. We assume the streaming application on the auctioneer supports three different constant bit rates $\{B_{low}, B_{med}, B_{high}\}$. If the application can not stream with a rate at least B_{low}, the streaming application can not stream data at all and suffers from failure. Based on the feedback from bidders, Auctioneers can encounter following two scenarios:

- Coop-selfServe: no bidder is able to provide relaying service that satisfies the lowest rate requirement.
- CooP: some bidders are able to provide relaying service that satisfies at least the lowest rate requirement.

In scenario CooP-selfServe, auctioneers suffer from failure if they can not stream with the lowest required rate by themselves, otherwise, they initiate a single path streaming traffic towards the BSs they associated with, and transmit at the highest streaming rate they can support. In scenario CooP, auctioneers compute the bid provided by each bidder by optimizing streaming traffic allocation between itself and the bidder. With each bidder, the streaming rate adopted by the streaming application on auctioneer is the highest streaming rate supported by the cooperative transmission of the auctioneer and the bidder. With the known streaming rate, auctioneer formulates an optimization problem to minimize total energy cost of cooperate communication with the bidder. The optimization variable is the portion $\beta \in [0, 1]$ of the total streaming traffic to be relayed by the bidder. The bid is calculated based on the optimized total energy cost, and is explained in more details in Section 5.3.

In the rest of this section, we formulate the optimization problem to be solved by auctioneer in scenario CooP. First, we decide the streaming rate, then we formulate the optimization problem.

5.2.1 Supported Streaming Rate

Let $r_{imax} = \min\{R_{oimaxWifi}, R_{imaxLte}\}$ be the maximum transmission rate supported by bidder i, where $R_{oimaxWifi}$ is the maximum transmission rate through the Wi-Fi connection between bidder i and auctioneer o and $R_{imaxLte}$ is the maximum transmission rate through i's cellular connection to BS. Given the maximum transmission rate $R_{omaxLte}$ supported by auctioneer o's cellular connection, the maximum transmission rate supported by the cooperative transmission of auctioneer o and bidder i is $TP_i = r_{imax} + R_{omaxLte}$. Applying Shannon-Hartley theorem, $R_{imaxLte}$, $R_{oimaxWifi}$ and $R_{omaxLte}$ is calculated as follows:

$$R_{imaxLte} = W_{iLte} \log_2 \left(1 + \frac{Pt_{imaxLte} \cdot H_{iLte}(d_i)}{N_0 W_{iLte}}\right) \tag{5.1}$$

$$R_{oimaxWifi} = W_{wifi} \log_2 \left(1 + \frac{Pt_{maxWifi} \cdot H_{oiWifi}(d_{io})}{N_0 W_{wifi} + I_{iWifi}}\right) \tag{5.2}$$

$$R_{omaxLte} = W_{oLte} \log_2 \left(1 + \frac{Pt_{omaxLte} \cdot H_{oLte}(d_o)}{N_0 W_{oLte}}\right) \tag{5.3}$$

Here W_{iLte}, W_{wifi} and W_{oLte} are the spectrum bandwidth allocated to bidder i's cellular connection, Wi-Fi connection between bidder i and auctioneer o and auctioneer o's cellular connection, respectively. Since we assume that cellular spectrum is evenly distributed among UEs no matter which BS they are associated with, therefore, $W_{iLte} = W_{oLte} = \frac{W_{total}}{N_{ue}}$, where W_{total} is the total bandwidth of the cellular spectrum and N_{ue} is the total number of UEs that are associated with BSs in the system at the current time slot. H_{iLte} and H_{oLte} are the channel gain of bidder i's cellular connection and auctioneer o's cellular connection, respectively. H_{oiWifi} is the channel gain of the Wi-Fi connection between bidder i and auctioneer o. $Pt_{imaxLte}$, $Pt_{maxWifi}$ and $Pt_{omaxLte}$ are the maximum transmission power of cellular interface of bidder i, Wi-Fi interface of auctioneer o and cellular interface of auctioneer o, respectively. N_0 is noise power spectrum density. d_o is the distance in meters between auctioneer o and its associated BS. I_{iWifi} is the interference towards the Wi-Fi receiver in bidder i.

With the computed maximum transmission rate TP_i, auctioneer o always selects the highest streaming rate B_o restricted by TP_i:

- $TP_i < B_{low} \rightarrow$ the streaming event is not supported.
- $B_{low} \leq TP_i < B_{med} \rightarrow B_o = B_{low}$.
- $B_{med} \leq TP_i < B_{high} \rightarrow B_o = B_{med}$.
- $B_{high} \leq TP_i \rightarrow B_o = B_{high}$.

5.2.2 Optimization Problem Formulation

The optimization problem to minimize the total energy cost of the cooperative transmission among bidder i and auctioneer o is formulated as follows:

$$\underset{\beta}{\text{Minimize}} \quad Cost_o + Cost_i \tag{5.4}$$

$$\text{s.t.} \quad \beta B_o <= \min\{R_{oimaxWifi}, R_{imaxLte}\} \tag{5.5}$$

$$(1-\beta)B_o <= R_{omaxLte} \tag{5.6}$$

$$0 \leq \beta \leq 1 \tag{5.7}$$

$Cost_o$ is the energy cost on auctioneer o, and $Cost_i$ is the energy cost on bidder i. Constraint (5.5) and (5.6) state that the streaming rate relayed by bidder i and auctioneer o should not exceed their corresponding transmission capacity.

In this work, we study two different energy cost functions. The first one is the Linear Cost Function (LCF) introduced in Section 5.1, and the second one uses consumed Energy As the energy Cost (EAC). The closed form expressions for $Cost_o$ and $Cost_i$ are as follows when using LCF approach:

$$\begin{aligned} Cost_i &= \int_{e_i=E_i-Pt_{iLte}T_o}^{E_i} \Big(\mathbb{I}_{\mathcal{A}_i}(e_i)\alpha_{i0} + (1-\mathbb{I}_{\mathcal{A}_i}(e_i))\alpha_{i1} + k_i e_i\Big) de_i \\ &= cnst_{i1} Pt_{iLte} + cnst_{i2} Pt_{iLte}^2 \end{aligned} \tag{5.8}$$

$$Cost_o = \int_{e_o=E_o-(Pt_{oWifi}+Pt_{oLte})T_o}^{E_o} \Big(\mathbb{I}_{\mathcal{A}_o}(e_o)\alpha_{o0} + (1-\mathbb{I}_{\mathcal{A}_o}(e_o))\alpha_{o1} + k_o e_o\Big) de_o$$
$$= cnst_{o1}(Pt_{oWifi} + Pt_{oLte}) + cnst_{o2}(Pt_{oWifi} + Pt_{oLte})^2$$

where $cnst_{i1}$, $cnst_{i2}$, $cnst_{o1}$, $cnst_{o2}$ are:

$$cnst_{i1} = (\mathbb{I}_{\mathcal{A}_i}(e_i)\alpha_{i0} + (1-\mathbb{I}_{\mathcal{A}_i}(e_i))\alpha_{i1})T_o + k_i E_i T_o \tag{5.9}$$

$$cnst_{i2} = -\frac{k_i T_o^2}{2} \tag{5.10}$$

$$cnst_{o1} = (\mathbb{I}_{\mathcal{A}_o}(e_o)\alpha_{o0} + (1-\mathbb{I}_{\mathcal{A}_o}(e_o))\alpha_{o1})T_o + k_o E_o T_o \tag{5.11}$$

$$cnst_{o2} = -\frac{k_o T_o^2}{2} \tag{5.12}$$

T_o is the duration of one time slot. $Pt_{iLte}T_o$ is the energy consumption of bidder i during the time slot and $(Pt_{oWifi} + Pt_{oLte})T_o$ is the energy consumption of auctioneer o in the same time slot. In this work, we only considered the transmission energy of the wireless interface.

When using EAC approach, The expressions for $Cost_o$ and $Cost_i$ are as follows:

$$Cost_i = Pt_{iLte}T_o \tag{5.13}$$

$$Cost_o = (Pt_{oWifi} + Pt_{oLte})T_o \tag{5.14}$$

5.2.3 Problem Analysis and Discussion

The optimization variable in problem (5.4)-(5.7) is β. The constraints of the problem is convex. Therefore, the convexity of the optimization problem depends on the convexity of the objective function. To analysis the convexity of the objective function, we first compute the transmission power Pt_{iLte}, Pt_{oWifi} and Pt_{oLte}. Consider following five pivot values of β, $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$:

$$(1-\beta_0)B_o = R_{omaxLte} \rightarrow \beta_0 = 1 - \frac{R_{omaxLte}}{B_o} \tag{5.15}$$

$$\beta_1 B_o = W_{wifi}\log_2(1 + \frac{Pr_{minWifi}}{N_0 W_{wifi} + I_{iWifi}}) \rightarrow \beta_1 = \frac{W_{wifi}}{B_o}\log_2(1 + \frac{Pr_{minWifi}}{N_0 W_{wifi} + I_{iWifi}}) \tag{5.16}$$

$$\beta_2 B_o = W_{iLte}\log_2(1+\frac{Pr_{minLte}}{N_0 W_{iLte}}) \rightarrow \beta_2 = \frac{W_{iLte}}{B_o}\log_2(1+\frac{Pr_{minLte}}{N_0 W_{iLte}}) \tag{5.17}$$

$$(1-\beta_3)B_o = W_{oLte}\log_2(1+\frac{Pr_{minLte}}{N_0 W_{oLte}}) \rightarrow \beta_3 = 1-\frac{W_{oLte}}{B_o}\log_2(1+\frac{Pr_{minLte}}{N_0 W_{oLte}}) \tag{5.18}$$

$$\beta_4 B_o = \min\{R_{oimaxWifi}, R_{imaxLte}\} \rightarrow \beta_4 = \frac{\min\{R_{oimaxWifi}, R_{imaxLte}\}}{B_o} \tag{5.19}$$

Where $Pr_{minWifi}$, Pr_{minLte} are the sensitivity power levels of the Wi-Fi receiver and cellular receiver, respectively. $\beta_1 B_o$ is the channel capacity of the Wi-Fi link between auctioneer o and bidder i with lowest transmission power of auctioneer o's Wi-Fi interface. Similarly, $\beta_2 B_o$ is bidder i's cellular link channel capacity with the lowest transmission power. $(1-\beta_3)B_o$ is auctioneer o's cellular link channel capacity with the lowest transmission power.

β is constrained between β_0 and β_4. If $\beta_0 > \beta_4$, the problem has no feasible solution. The computation of Pt_{iLte}, Pt_{oWifi} and Pt_{oLte} depend on the β's relation with β_1, β_2 and β_3:

$$\text{If } \beta \le \beta_1 : Pt_{oWifi} = \frac{Pr_{minWifi}}{H_{oiWifi}(d_{io})} = C_{oWifi} \tag{5.20}$$

$$\text{If } \beta \le \beta_2 : Pt_{iLte} = \frac{Pr_{minLte}}{H_{iLte}(d_i)} = C_{iLte} \tag{5.21}$$

$$\text{If } \beta \ge \beta_3 : Pt_{oLte} = \frac{Pr_{minLte}}{H_{oLte}(d_o)} = C_{oLte} \tag{5.22}$$

$$\text{If } \beta \ge \beta_1 : Pt_{oWifi} = \frac{(2^{\frac{\beta B_o}{W_{wifi}}}-1)(N_0 W_{wifi}+I_{iWifi})}{H_{oiWifi}(d_{io})} = 2^{\frac{B_o\beta}{W_{wifi}}} D_{oWifi} - D_{oWifi} \tag{5.23}$$

$$\text{If } \beta \ge \beta_2 : Pt_{iLte} = \frac{(2^{\frac{\beta B_o}{W_{iLte}}}-1)(N_0 W_{iLte})}{H_{iLte}(d_i)} = 2^{\frac{B_o\beta}{W_{iLte}}} D_{iLte} - D_{iLte} \tag{5.24}$$

$$\text{If } \beta \le \beta_3 : Pt_{oLte} = \frac{(2^{\frac{(1-\beta)B_o}{W_{oLte}}}-1)(N_0 W_{oLte})}{H_{oLte}(d_o)} = 2^{\frac{-B_o\beta}{W_{oLte}}} F_{oLte} - D_{oLte} \tag{5.25}$$

The only unknown variable in the above equations is β, hence we use C_{oWifi}, C_{iLte}, C_{oLte}, D_{oWifi}, D_{iLte}, F_{oLte} and D_{oLte} to denote the constant parts of the equations above. Equations (5.20) to (5.25) indicate that when the streaming rate is smaller than the channel capacity with the minimal transmission power, the transmission power is constant and is the minimal transmission power that guarantees the signal being received at the receiver side is at its sensitivity level. Otherwise, the transmission power is a function of the streaming rate.

Since the transmission power Pt_{iLte}, Pt_{oWifi} and Pt_{oLte} depend on β's relation with β_1, β_2 and β_3, the optimization problem is different with different feasible ranges of β. From permutation theory we know that there are 6 different relations of β_1, β_2 and β_3. For each relation, the optimization problem can be divided into four problems, each of which optimizes on a different feasible range of β.

As an example, in the case of $\beta_1 \leq \beta_2 \leq \beta_3$, the four feasible ranges of β are $[\max\{\beta_0, 0\}, \min\{\beta_1, \beta_4, 1\}]$, $[\max\{\beta_1, \beta_0, 0\}, \min\{\beta_2, \beta_4, 1\}]$, $[\max\{\beta_2, \beta_0, 0\}, \min\{\beta_3, \beta_4, 1\}]$, $[\max\{\beta_3, \beta_0, 0\}, \min\{\beta_4, 1\}]$. If we apply LCF energy cost function, the objective function when $\max\{\beta_1, \beta_0, 0\} \leq \beta \leq \min\{\beta_2, \beta_4, 1\}$ becomes:

$$
\begin{aligned}
&Cost_i + Cost_o \qquad (5.26)\\
&= cnst_{i1} Pt_{iLte} + cnst_{i2} Pt_{iLte}^2 + cnst_{o1} Pt_{oWifi} + cnst_{o1} Pt_{oLte} + cnst_{o2}(Pt_{oWifi} + Pt_{oLte})^2 \\
&= cnst_{i1} C_{iLte} + cnst_{i2} C_{iLte}^2 + cnst_{o1} 2^{\frac{B_o \beta}{W_{wifi}}} D_{oWifi} + cnst_{o1} 2^{\frac{-B_o \beta}{W_{oLte}}} F_{oLte} \\
&- cnst_{o1}(D_{oWifi} + D_{oLte}) + cnst_{o2} \cdot (2^{\frac{B_o \beta}{W_{wifi}}} D_{oWifi} - D_{oWifi} + 2^{\frac{-B_o \beta}{W_{oLte}}} F_{oLte} - D_{oLte})^2
\end{aligned}
$$

the optimization problem becomes:

$$\underset{\beta}{\text{Minimize}} \quad Cost_i + Cost_o \qquad (5.27)$$

$$\text{s.t.} \quad \max\{\beta_1, \beta_0, 0\} \leq \beta \leq \min\{\beta_2, \beta_4, 1\} \qquad (5.28)$$

the objective function is not convex. Let $X = 2^{\frac{B_o \beta}{W_{wifi}}}$, $Y = 2^{\frac{-B_o \beta}{W_{oLte}}}$, (5.27)-(5.28) become:

$$
\begin{aligned}
\underset{X,Y}{\text{Minimize}} \quad & cnst_{o1}(D_{oWifi} X + F_{oLte} Y) + cnst_{o2} \cdot (D_{oWifi} X - D_{oWifi} \\
& + F_{oLte} Y - D_{oLte})^2 + cnst_{i1} C_{iLte} + cnst_{i2} C_{iLte}^2 - cnst_{o1}(D_{oWifi} + D_{oLte})
\end{aligned} \qquad (5.29)
$$

$$\text{s.t.} \quad Y X^{\frac{W_{wifi}}{W_{oLte}}} = 1 \qquad (5.30)$$

$$2^{\max\{\beta_1, \beta_0, 0\} \cdot B_o / W_{wifi}} \leq X \leq 2^{\min\{\beta_2, \beta_4, 1\} \cdot B_o / W_{wifi}} \qquad (5.31)$$

now the objective function is convex, but constraint (5.30) is non-convex, we relax (5.30) to the two constraints below:

$$Y \geq X^{\frac{-W_{wifi}}{W_{oLte}}} \qquad (5.32)$$

$$Y \leq (2^{\max\{\beta_1,\beta_0,0\}\cdot B_o/W_{wifi}})^{\frac{-W_{wifi}}{W_{oLte}}} \tag{5.33}$$

This relaxed optimization problem (5.29),(5.31),(5.32)-(5.33) gives a lower bound for the original problem (5.29)-(5.31). Use the optimal β computed from this relax optimization problem, we calculate the actual energy cost $Cost_o + Cost_i$. In the simulation, we calculate the ratio of actual energy cost computed over the optimal energy cost computed from the relaxed optimization problem, and the average ratio is 1.0211 with standard deviation of 0.1371. Moreover, of all the relaxed optimization problems solved in the simulation, around 51% of them give the actual optimal result. This number indicates that our relaxation problem gives a solution that is very close to the optimal solution's lower bound, hence is close to the optimal solution.

When we apply LCF as the energy cost function, there are 6 different relations among β_1, β_2 and β_3, and four different settings for Pt_{iLte}, Pt_{oWifi} and Pt_{oLte} in each relation, there are total of 24 possible representations of the optimization problems (only 8 distinct representations). Some of the optimization problems are convex, some are not. We use similar relaxation approach used above to solve the non-convex optimization problems. If we apply EAC as the energy cost function, the objective function is convex in β's entire feasible range. So the optimization problem using EAC approach is solvable using convex solver without any relaxation. The convex solver we used in this work is CVX [81].

5.3 Auction Mechanism Design

With multiple bidders participating in the auction, the auctioneer needs to select the right bidder and makes the appropriate payment to the selected bidder. In this section, we introduce the auction mechanism designed for our user cooperative communication auction. The following selection rule and payment rule are modified version of the selection rule and payment rule from the classic Vickrey auction mechanism [82].

Selection Rule: Auctioneer selects the bidder i_o that provides the minimal bid, where b_i is the bid offered by bidder i, which is the $Cost_i + Cost_o$ optimized in Section 5.2.3:

$$i_o = \arg\min_i b_i \tag{5.34}$$

If the energy cost of the cooperative transmission with the selected bidder is higher than the energy cost when auctioneer self serves itself (transmits streaming traffic without cooperation) with the same streaming rate or the optimized β value is 0, the auctioneer will self serve itself.

Payment Rule: the auctioneer pays the winning bidder i_o the lowest bid that is higher than b_{i_o} minus the auctioneer's energy cost $Cost_o$ when cooperates with bidder i_o. It pays 0 to bidders that lose in the auction:

$$P_i = \begin{cases} min\{b_j : b_j > b_i\} - Cost_o & \text{if i} = i_o \\ 0 & \text{if i} \neq i_o \end{cases}$$

In the case where there is only one bidder i or there is no bid higher than b_i, and auctioneer chooses to cooperate with the bidder, auctioneer then pays bidder $P_i = b_{max} - Cost_o$, where $b_{max} = Cost_o^{max} + Cost_i^{max}$, $Cost_o^{max}$ and $Cost_i^{max}$ are calculated with maximum transmission power on auctioneer o's cellular interface, Wi-Fi interface and bidder i's cellular interface.

The payment is paid in credit. The auctioneer reduces the same amount of the payment from its credit balance, and the winning bidder adds the same amount of payment to its credit balance. The credit is used by bidders in this work to decide which auctioneer to serve in order to enjoy a higher utility and compensate mobile users that provide relay service to other mobile users. There are other usages of the credit system, such as 1) preventing mobile users without any credits from requesting other mobile users for relay; 2) preventing UEs from falsely reporting information to gain extra credits, etc. In this work, we do not study enforcement strategies that can use the credit to prevent mobile users from cheating or taking advantage of the credit system, instead, we focus on studying honest users' behaviors when participating in auctions in the user cooperative system and how different energy cost functions affect users' energy usage differently. Therefore, the initial credit level for all UEs in the system is set as zero, and UEs are not forbidden from participating into the auction even when they have negative credits levels.

The utility of bidder i is

$$U_i = P_i - Cost_i \tag{5.35}$$

if bidder does not cheat by providing false information $Info_i$, its utility will always be non-negative. In the case where bidder i wins, $P_i = (second\ lowest\ bid - Cost_o)$, then $U_i =$

$(second\ lowest\ bid - Cost_o - Cost_i) = (second\ lowest\ bid - lowest\ bid) \geq 0$. In the case where bidder i loses, $P_i = 0$, $U_i = (0 - 0) = 0$.

In this work, we assume UEs are honest and report information truthfully. What's more, more than one UE is allowed to start the auctions simultaneously. So it is possible that a bidder participates into multiple auctions, and is selected as the winner in multiple auctions. To resolve this confliction, we set up the following iterative matching process:

- Auctioneer:
  - Sort bidders in an increasing order of bids provided.
  - Iteratively confirms with bidders in the increasing order. If the bidder accepts the offer, the auctioneer stops and selects the bidder for relay transmission; If the bidder temporarily rejects the offer, the auctioneer continues confirming with the next lowest bidder on the ordered list.
  - When auctioneer finishes confirming with all bidders in the list without any offer, it refreshes its' bidder list (remove bidders that reject its offer permanently from the list), and continue confirming with bidders in an increasing order of the refreshed list.
  - It ends the auction when it reaches an agreement with one bidder or all of its bidders reject its offer permanently. When it ends the auction, it broadcasts 'end auction' signal to other bidders.
- Bidder:
  - Sorts auctioneers (in the auctions bidder participates in) in a decreasing order of utilities bidder can enjoy.
  - When being contacted by one auctioneer, it checks if the auctioneer is the one that can provide it with the highest utility. If it is, bidder accepts the offer, and broadcasts 'reject permanently' signal to other auctioneers. If it is not, bidder temporarily rejects the auctioneer.

– Bidder removes the auctioneer from its auctioneers list whenever it receives 'end auction' signals from the auctioneer.

5.4 Real-time Optimization and Failure Protection

In reality, UEs' locations as well as the channel conditions are changing over time, which can result in UEs' streaming services suffering from rate loss or total service failure. Therefore, the failure protection functionality is needed for better quality of service. What's more, the changing may improve the channel condition as well, in which case the streaming application should be able to adjust the streaming rate to improve quality of service accordingly.

In this work, UEs are assumed to operate with constant time step, and UEs' new movement and channel condition are updated at the beginning of each time slot. During each time slot, UEs' movement and channel condition remain static. UEs' streaming service lasts for more than one time slot. When UEs stream traffic, they perform real-time optimization and failure protection in slots other than the first slot.

5.4.1 Operations When UEs Self Serve in Streaming

In this case, UEs are self serving themselves with rate B^k when stream traffic at time slot k. At the beginning of time slot $k+1$, UEs first update the channel capacity TP^{k+1}. If $TP^{k+1} >= B^{low}$, UEs continue streaming with the highest streaming rate that TP^{k+1} supports. If $TP^{k+1} < B^{low}$, UEs start an auction to retry the streaming service. When no bidder in the auction can help with the streaming, UEs end the streaming service and announce streaming service disconnection failure.

5.4.2 Operations When UEs Cooperate Transmit with Other UEs in Streaming

In this case, UEs cooperate transmit with other UEs at time slot k. For example, auctioneer o establishes two-path streaming connections. One path goes through auctioneer's cellular connection with base station BS_o, and another path goes through auctioneer's Wi-Fi connection with bidder i. Below are the possible failures during streaming.

- **AucLinkFailure**: Cellular link between auctioneer o and base station BS_o suffers from quality drop or disconnection .
- **BidLinkFailure**: Cellular link between bidder i and base station BS_i suffers from quality drop or disconnection.
- **WiFiLinkFailure**: Wi-Fi link between auctioneer o and bidder i suffers from quality drop or disconnection.

At time slot k, auctioneer o streams with rate B^k, and β^k portion of the rate is relayed by bidder i. At the beginning of time slot $k+1$, the updated channel capacity is TP^{k+1}. Different values of TP^{k+1} result in different operations for auctioneer o:

- $TP^{k+1} < B_{low}$: auctioneer o ends the cooperation with bidder i, and retries the streaming service by starting another auction process. If there is no bidder available during the new auction, auctioneer o ends the streaming service and announces streaming service disconnection failure.
- $TP^{k+1} \geq B_{low}$: auctioneer o updates B^{k+1} to be the highest streaming rate that TP^{k+1} supports, meanwhile re-optimizes β^{k+1} to minimize the total energy cost.
  - $B^{k+1} \leq R^{k+1}_{omaxLte}$: auctioneer o self serves itself if self-serving consumes lower energy cost. otherwise, auctioneer cooperates with bidder i with the optimized β^{k+1}. At the end of time slot $k+1$, the credits paid to bidder i are updated accordingly.
  - $B^{k+1} > R^{k+1}_{omaxLte}$: auctioneer cooperates with bidder i, where $\beta^{k+1}B^{k+1}$ is relayed by bidder i.

As a summary, when UEs perform real-time optimization and failure protection, they will try to continue the cooperation with the neighbor UEs the same as the ones relay the traffic in previous time slot until the streaming service fails and needs retry. This is to make sure that UEs' streaming service goes smoothly throughout the entire streaming event, since starting a new auction is time costly.

5.5 Simulation Results

In this section, we study the performance of our proposed user cooperative multi-path routing scheme. We first introduce the wireless network environment set up and UEs' battery model and streaming model in Section 5.5.1 and Section 5.5.2, respectively. The detailed simulation results are shown in Section 5.5.3.

5.5.1 Wireless Network Environment

In the MATLAB simulation platform, we set up one MBS, 5 FBSs and 100 UEs. The FBSs are uniformly distributed within 200 meters from the MBS, and for each FBS, there are 10 UEs uniformly deployed within 50 meters from the FBS. What's more, there are another 50 UEs uniformly deployed within 250 meters from the MBS. This simulation setup simulates a two-layer heterogeneous wireless network environment.

The cellular carrier center frequency used is 2 GHz, and the total cellular spectrum bandwidth is 20 MHz. UE associates with the BS that provides it with the strongest signal strength. The cellular channel spectrum is evenly distributed among UEs that are associated with BSs. Wi-Fi channel has 40 MHz bandwidth [83]. UE's maximum transmission power of cellular interface and Wi-Fi interface are 23 dBm and 10 dBm, respectively. The sensitivity power level for Wi-Fi interface and cellular interface are -40 dBm and -101.5 dBm, respectively. The noise power spectrum density is -174 dBm/Hz. The path loss model for cellular channel is $PL(dB) = 15.3 + 37.6 \log_{10} R + X_c$, where R is in meters [76], and X_c is Gaussian random variable with zero mean and standard deviation $\delta_c = 15$. X_c reflects the attenuation caused by flat fading. The Wi-Fi signal path loss model is $PL(dB) = 32.2 \log_{10}(d) + X_w$, where d is in meters [84], and X_w is Gaussian random variable with zero mean and standard deviation $\delta_w = 8$. Because each UE uses different portion of the cellular spectrum, there is no interference to BS's cellular interface. We assume each UE's Wi-Fi interface suffers interference from half of signal power received from all the other UEs transmitting at maximum transmission power. It is a reasonable assumption since in the simulation, the number of UEs that are transmitting using Wi-Fi interface is less than half of the total UEs in the system.

To simulate the dynamic of channel conditions, at the beginning of each time slot, we re-randomize the flat fading variable X_c and X_w, and each UE walks randomly within 1 meter from its position in previous time slot. At the beginning of each time slot, FBS suffers from cell outage with probability Pr_f. We study network performance with Pr_f selected from a number of different values [0, 0.015, 0.03, 0.045, 0.06, 0.075, 0.09, 0.105, 1].

5.5.2 UE's Battery Model and Streaming Model

In practise, only a portion of UE's battery is used for wireless transmission. Therefore, we set the initial energy balance for all UEs as 2150 J (10% of Iphone 5s battery capacity [1570 mAh, 3.8 V].). α_{i0} and k_i used in the LCF energy cost approach are 2150 and -1, respectively.

The three constant bit rates that streaming application supports are $B_{low} = 1Mbps$, $B_{med} = 2Mbps$ and $B_{high} = 3Mbps$.

The simulation lasts for 90 time slots, and each time slot lasts for 30 seconds. In each time slot, each UE starts streaming event with certain probability. Each streaming event lasts for 3 time slots. For each UE i, the number of streaming events x_i generates during the entire simulation (90 time slots) is poisson distributed with mean $\lambda_i = 10$. At the beginning of time slot k, each UE streams as in following different cases:

- UE i is idle, it starts a streaming event with probability $\frac{x_i}{S/3}$, where $S = 90 - k + 1$ is the number of remaining time slots. When S is smaller than 3, the probability of starting streaming event is 0.
  - If UE i starts the streaming event in time slot k, then it continues the streaming event in the following 2 time slots, and at the end of time slot $k + 2$, the streaming event is finished, UE i updates $x_i = x_i - 1, S = S - 3$, and becomes idle in the beginning of time slot $k + 2$;
  - If UE i does not generate streaming event in time slot k, it updates $x_i = x_i, S = S-1$ at the end of time slot k, and remains idle.
- UE i continues streaming the event generated in previous time slot.

- UE i announces the streaming event disconnection failure. It updates $x_i = x_i - 1, S = S - 1$ at the end of time slot k.

5.5.3 Performance Analysis

In the simulation, we assume all bidders are honest and report information truthfully to auctioneers. In the rest of the section, we compare the performance of three operation schemes: self serving scheme where UE transmits its streaming traffic without cooperation during the entire simulation (non-cooperation scheme); user cooperative multi-path routing scheme with LCF energy cost approach (LCF scheme), and user cooperative multi-path routing scheme with EAC energy cost approach (EAC scheme).

Fig. 5.3 and Fig. 5.8 give the energy balance and credit balance of 10 selected UEs throughout a selected simulation round using EAC scheme. Fig. 5.5 and Fig. 5.10 give the energy balance and credit balance of the same 10 selected UEs throughout the same selected simulation round using LCF scheme. The failure probability of FBS in the selected simulation round is 1, which means all FBSs are failed from the beginning of the simulation and UEs are served by MBS throughout the simulation. Fig. 5.3 and Fig. 5.5 show that the energy balance for all UEs are decreasing as simulation goes under both LCF and EAC schemes. What's more, some of the UEs have higher remaining energy balance than average remaining energy balance of all UEs in the system, and other UEs have lower remaining energy balance. The UEs further away from the MBS consume more energy than UEs closer to MBS. The energy level is also affected by UEs' total number of generated streaming events throughout the simulation, and how the generated streaming traffic is transmitted (cooperative or self-served). UE 2 in Fig. 5.3 and UE 2, UE 6 in Fig. 5.5 are such examples.

In both credit balance figures of LCF scheme and EAC scheme, the credit balance is fluctuated as time goes. This is due to the fact that throughout the simulation, UE sometimes acts as the auctioneer and sometimes acts as the bidder. UEs further away from MBS have more dramatic fluctuation in credit balance. This is because their neighbor UEs are also very far way from the MBS, and the energy costs are much higher as well. Fig. 5.6 and Fig. 5.11 do not show a direct relation of higher energy cost results in higher credit balance. In fact, the

credit balance fluctuation is jointly determined by UE's streaming frequency, UE's locations from BSs and neighbor UEs' streaming frequency. In Fig. 5.10, the LCF scheme results in very large credit value due to its used cost function, and we can control the credit value by tuning the parameters in the cost function.

Fig. 5.16 shows the accumulated streaming event count at each time slot of the same 10 selected UEs as in Fig. 5.11. In the figure, UEs that streaming event successfully increase the accumulated event count in a continuous 3 time slots, otherwise, UEs fail the streaming event. Comparing Fig. 5.6, Fig. 5.11 and Fig. 5.16, UEs with lower streaming event are more likely to have higher energy balance. UEs with higher streaming events, higher energy balance are more likely to have lower credit balance. UEs with lower streaming events, lower energy balance are more likely to have higher credit balance. As a summary, both UEs' energy level fluctuation and credit balance fluctuation are affected jointly by UE's streaming frequency, UE's locations from BSs, neighbor UEs' streaming frequency and UEs' approach for transmission.

Fig. 5.21 shows the performance comparison between cooperation schemes and non-cooperation scheme (Self Serving) with two different FBS failure probabilities. The green circled line shows a rate ratio of cooperation scheme over non-cooperation scheme over different simulation runs. The rate is a summation of all UEs' streaming rate over 90 simulation slots with FBS failure probability being 0. It shows that cooperation schemes and non-cooperation scheme provide the same throughput when FBS failure probability is 0. In fact, given the three different streaming rate levels (B_{low}, B_{med}, B_{high}), cooperation schemes and non-cooperation scheme can support most of UEs with B_{high} when there is no FBS failure. When FBS failure probability becomes 1, cooperation schemes can provide UEs with higher rate compared to non-cooperation scheme, and this is shown by the light green circled dash line.

Fig. 5.21 also shows UEs' total energy consumption ratio of both cooperation schemes over non-cooperation scheme. UEs with cooperation schemes consume less energy compared to UEs with non-cooperation scheme. When FBS failure probability is 0, UEs with cooperation schemes and non-cooperation scheme enjoy the same sum rate, however, in cooperation schemes, UEs consume about 35% less energy compared to that in non-cooperation scheme. With FBS failure probability being 1, the energy efficiency improvement of cooperation schemes over non-

cooperation scheme is around 30%, and cooperation schemes provide UEs with higher rate compared to non-cooperation scheme.

Fig. 5.26 shows UEs' streaming event success rate of cooperation schemes and non-cooperation scheme. When FBS failure probability is 0, UEs' streaming event success rate is 1 for both cooperation scheme and non-cooperation scheme. With FBS failure probability being 1, the success rate for both cooperation scheme and non-cooperation scheme is lower than 1, and cooperation schemes provide higher success rate.

Fig. 5.27 shows the sum rate of all UEs over 90 simulation slots averaged over 30 simulation runs under different FBS failure probabilities. The averaged sum rate decreases as FBS failure probability increases. UEs with cooperation schemes have better throughput compared to non-cooperation scheme when FBSs are suffering from failure, and the improvement increases as FBS failure probability increases.

Fig. 5.27 also shows the total energy consumption of all UEs over 90 simulation slots averaged over 30 simulation runs. The energy consumption increases as FBS failure probability increases. FBSs are located closer to UEs, therefore, UEs consume lower energy when transmitting to FBSs. When FBSs fail, UEs need to transmit to MBS, and consume more energy. The figure also shows that UEs with cooperation schemes consume less energy compared to UEs with non-cooperation scheme, and the energy consumed is around 30% less compared to non-cooperation scheme under different FBS failure probabilities.

To compare the performance between EAC energy cost function and LCF energy cost function in user cooperation scheme, we run simulation with a different simulation set up [85], where there is no FBS in the system, and all UEs are served by MBS. What's more, UEs' streaming event lasts for one time slot. The network environment are static throughout the simulation (UEs's location and channel condition remain static). This new simulation set up makes sure the simulations run with EAC and LCF scheme are only different in the energy cost function. Comparing the performance between EAC scheme and LCF scheme under this new simulation set up is much fair when we compare EAC scheme and LCF scheme under previous simulation set up. Because in previous simulation setup, streaming event lasts for 3 time slots. Longer streaming event results in different streaming behaviour throughout the simulation for

the same UE under different cooperation schemes. This difference can be seen in Fig. 5.13 and Fig. 5.15.

Fig. 5.28 shows the performance comparison between two user cooperation schemes LCF and EAC under the new simulation environment. The upper circled line shows the ratio of total UEs' energy consumption applying LCF scheme over total UEs' energy consumption applying EAC scheme. The ratio is close to 1, and LCF scheme consumes more energy than EAC scheme by about 2-3%. The lower line is the ratio of standard deviation of UEs' remaining energy with LCF scheme over the standard deviation of UEs' ramaining energy with EAC scheme. The standard deviation of UEs' remaining energy indicates how balanced are the energy consumption among all UEs in the system. The ratio shows that LCF scheme outperforms EAC scheme in balancing UE's energy consumption. In LCF scheme, UEs with lower energy will have higher energy cost as indicated by the cost function introduced in Section 5.1. As a result, the auctioneer is more likely to select the bidder with higher remaining energy. However, in EAC scheme, auctioneer selects bidder which can provide it with minimal energy consumption.

5.6 Conclusions

In this work, we proposed user cooperative multi-path routing solution for streaming application. We designed the auction mechanism which incentivize UEs to provide relay service for other UEs. We also designed UEs' real-time optimization and failure protection operations to provide better quality of service. With the assumption that UEs are all honest players, our proposed solution showed great advantage in improving service rate, improving streaming event success rate and reducing energy consumption compared to non-cooperation solution. The proposed auction mechanism motivates UEs into participating in the user cooperation auction game with non-negative utilities. The comparison study also shows that LCF energy cost function has more potential in balancing the UE's remaining energy across all UEs in the system.

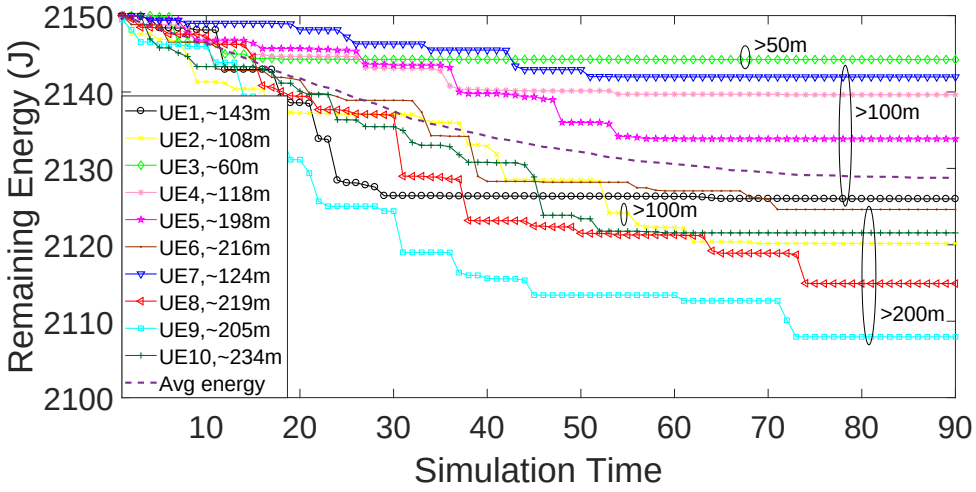


Figure 5.3 With EAC energy cost function.

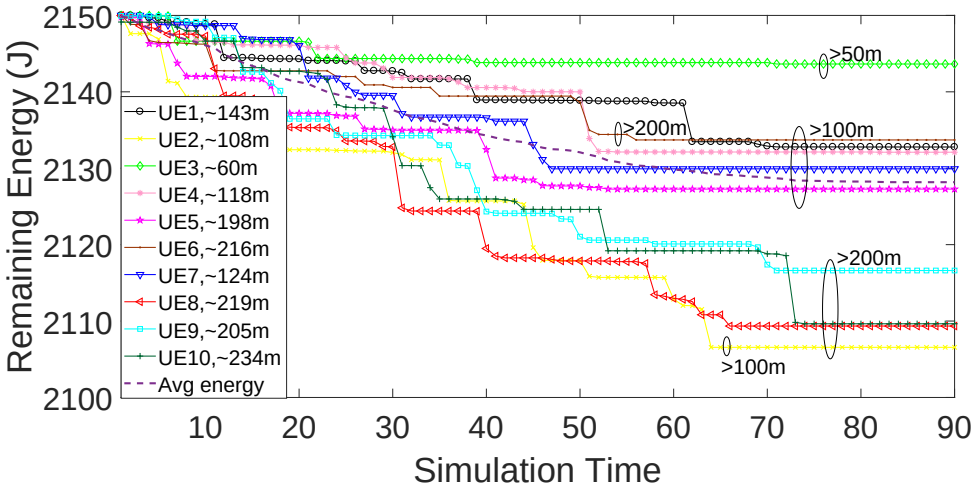


Figure 5.5 With LCF energy cost function.

Figure 5.6: Selected UEs' available energy vs simulation time with cooperation scheme.

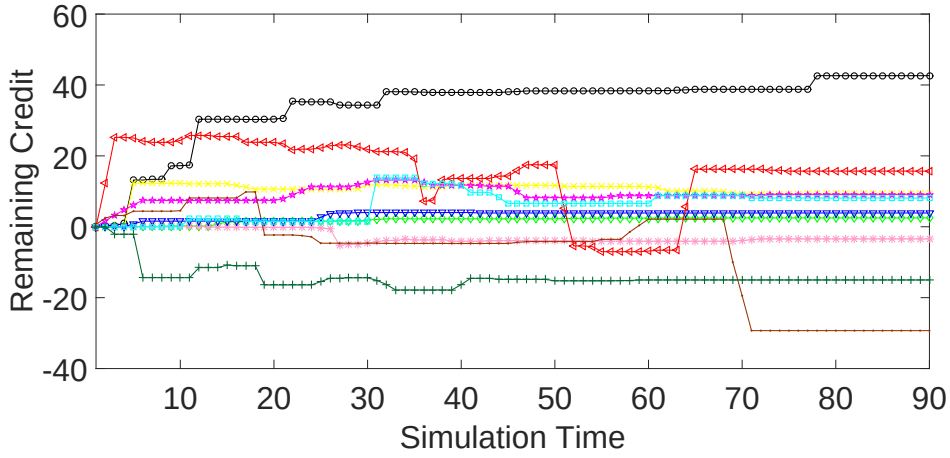


Figure 5.8 With EAC energy cost function.

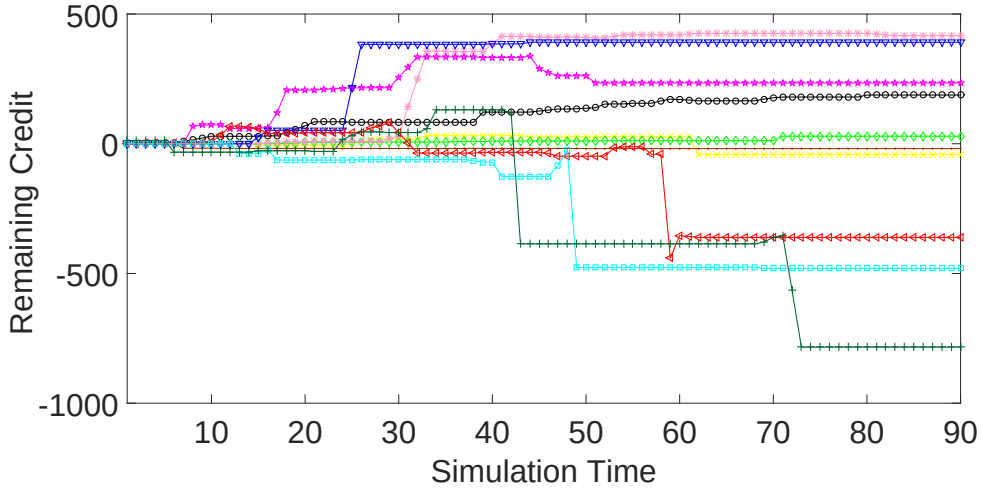


Figure 5.10 With LCF energy cost function.

Figure 5.11: Selected UEs' available credit vs simulation time with cooperation scheme.

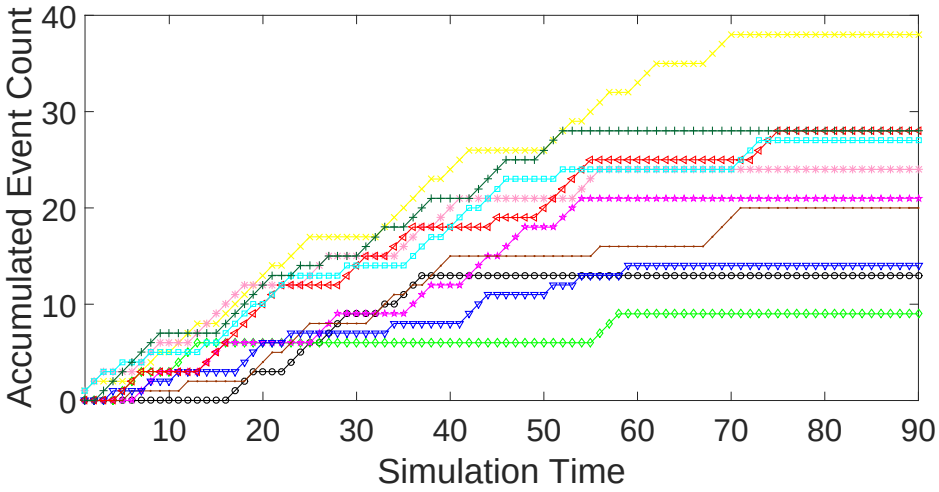


Figure 5.13 With EAC energy cost function.

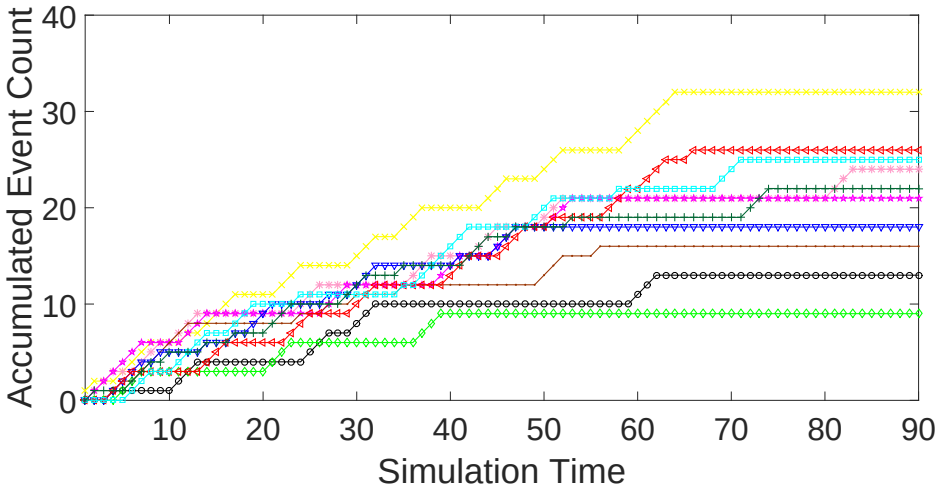


Figure 5.15 With LCF energy cost function.

Figure 5.16: Selected UEs' accumulated event count vs simulation time with cooperation scheme.

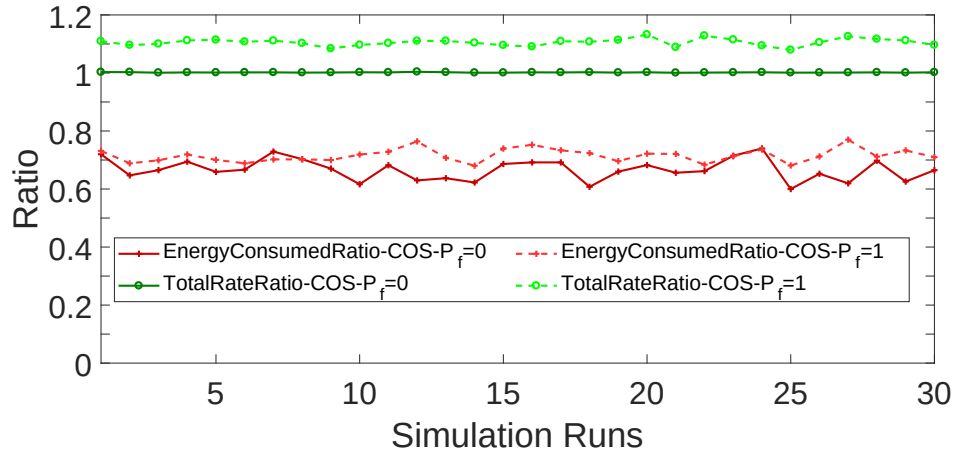


Figure 5.18 With EAC energy cost function.

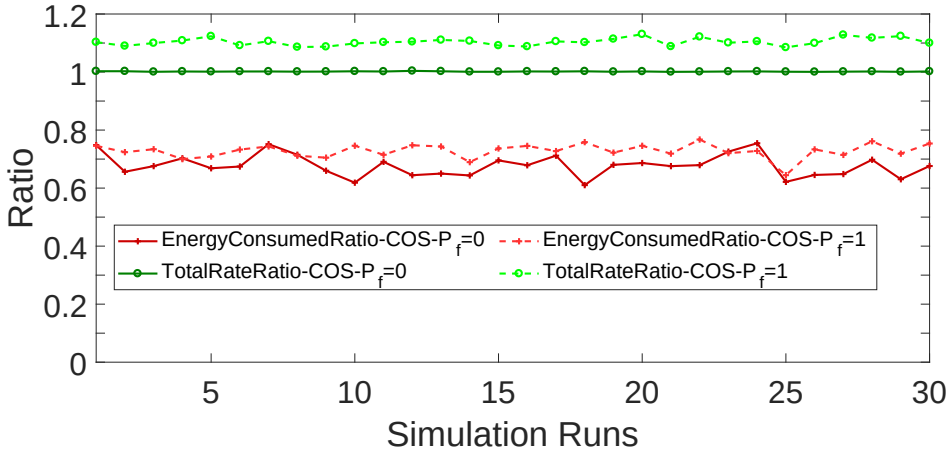


Figure 5.20 With LCF energy cost function.

Figure 5.21: Ratio of UEs' rate consumption with cooperation schemes over rate consumption of non-cooperation scheme and energy consumption with cooperation schemes over energy consumption with non-cooperation scheme.

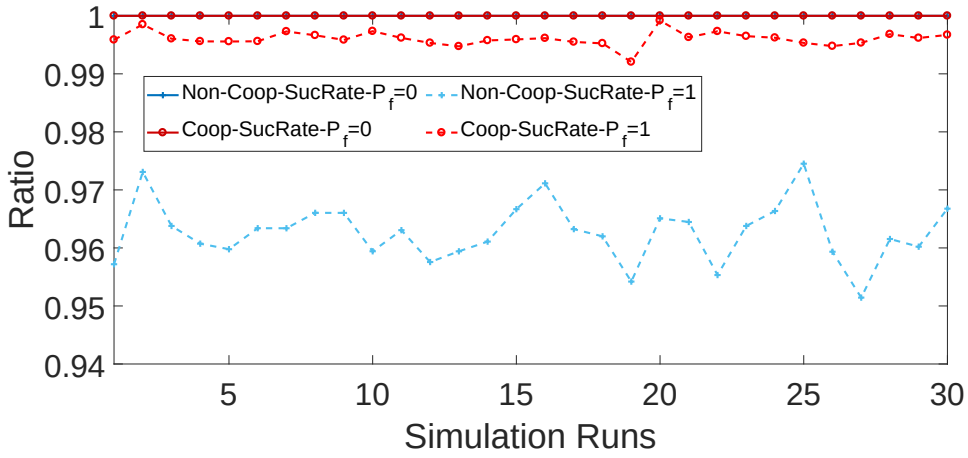


Figure 5.23 With EAC energy cost function.

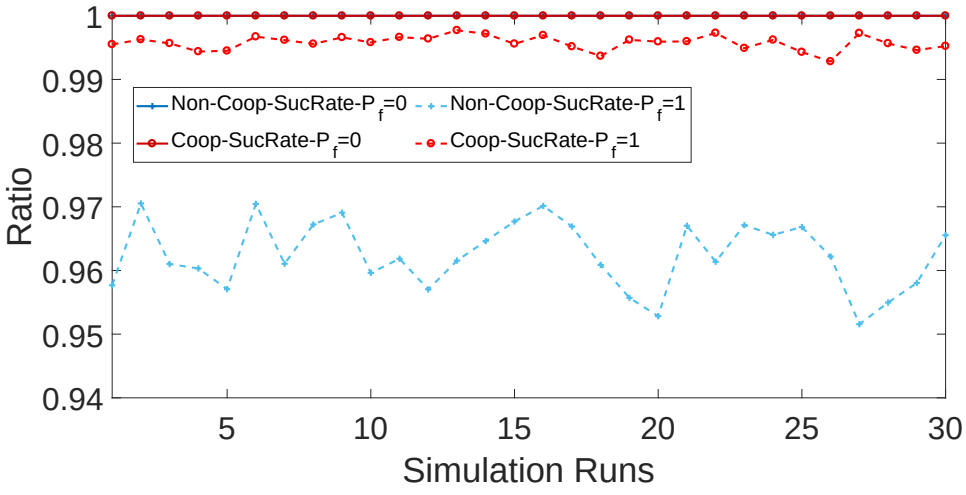


Figure 5.25 With LCF energy cost function.

Figure 5.26: UEs' streaming event success rate with cooperation scheme over success rate with non-cooperation scheme.

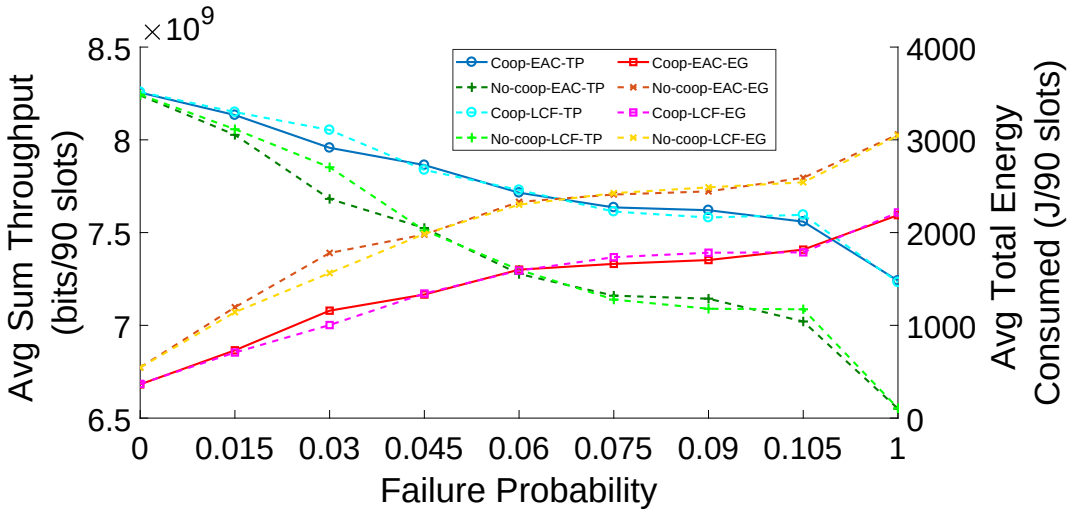


Figure 5.27: UE's sum throughput (TP) and total energy consumed (EG) averaged over 30 time slots vs different FBS failure probability

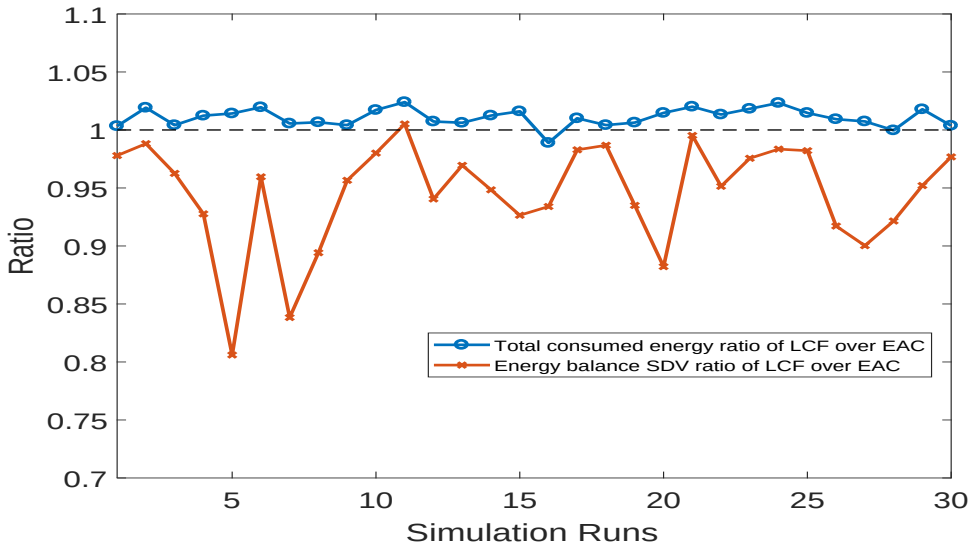


Figure 5.28: Total UEs' energy consumption and UEs' remaining energy standard deviation comparison between LCF scheme and EAC scheme

CHAPTER 6. CONCLUSION AND FUTURE RESEARCH

In this thesis, we studied optimal operation of wireless networks including CRNs and heterogeneous cellular wireless networks. The objective was to reuse existing radio spectrum in wireless networks to improve network performance. We first studied multicast routing in CRNs. We proposed a new modeling approach for CRNs, which simplifies the multicast routing problem to find multicast routing trees in the network graph. We also proposed multicast routing optimization algorithms to find the Pareto Front of the multicast routing tree. Simulation results show that the proposed algorithms can find most of the Pareto Front in a short amount of time, and the proposed multicast routing schemes are efficient to operate in small CRNs. In the second and third work, we studied cooperation solutions among network devices in a heterogeneous network to improve network performance and provide network failure protection capability. In our study, our proposed network cooperation and user cooperation approaches are shown to improve UEs' throughput compared to non-cooperation schemes. The network cooperation approach using CoMP provides failure compensation capability by preventing the system sum rate loss from having the same speed of radio resource loss, and this is done without using additional radio resources and will not impair much of the performance of other UEs. The user cooperation approach shows great advantage in improving service rate, improving streaming event success rate and reducing energy consumption compared to non-cooperation solutions. The proposed auction mechanism in the user cooperation solution motivates UEs into participating in the user cooperation auction game with non-negative utilities.

We studied cooperation solutions among base stations and users equipment in this thesis. There are some other interesting ideas can be explored for further network performance improvement. The network cooperation approach using CoMP is computationally intensive. An online version of the approach can be studied to reduce algorithm response time. Moreover,

solutions are needed to cancel interference between CoMP operation units to improve network performance further. In our proposed multi-path user cooperation approach, we studied one hop relay and at most two streaming paths. A multi-hop relay scheme with multi-path streaming service has greater potential and is more promising to improve wireless network performance. In addition, enforcement schemes need to be proposed together with the credit system to prevent mobile users from taking advantage of the credit system.

BIBLIOGRAPHY

[1] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, *et al.*, "Scenarios for 5g mobile and wireless communications: the vision of the metis project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, 2014.

[2] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.

[3] B. R, "How to build a national cellular wireless network for 50m," 2011.

[4] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for lte-advanced [coordinated and distributed MIMO]," *IEEE Wireless Communications*, vol. 17, no. 3, pp. 26–34, Jun. 2010.

[5] T. R. 3GPP, "36.819: Coordinated multi-point operation for LTE physical layer aspects (release 11)." www.3gpp.org, 2011.

[6] U. Jang, H. Son, J. Park, and S. Lee, "Comp-csb for ici nulling with user selection," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 2982–2993, 2011.

[7] D. Jaramillo-Ramirez, M. Kountouris, and E. Hardouin, "Coordinated multi-point transmission with imperfect csi and other-cell interference," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 1882–1896, 2015.

[8] C.-I. Kuo, S.-H. Wu, and C.-K. Tseng, "Robust linear beamformer designs for coordinated multi-point af relaying in downlink multi-cell networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3272–3283, 2012.

[9] J.-B. Seo, S.-Y. Kim, and V. Leung, "Outage probability characterization of comp-joint transmission with path-loss and rayleigh fading," *Communications Letters*, vol. 19, no. 1, pp. 78–81, 2015.

[10] Z. Xu, C. Yang, G. Y. Li, Y. Liu, and S. Xu, "Energy-efficient comp precoding in heterogeneous networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 1005–1017, 2014.

[11] S. Fu, B. Wu, H. Wen, P.-H. Ho, and G. Feng, "Transmission scheduling and game theoretical power allocation for interference coordination in comp," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 112–123, 2014.

[12] H. Pilaram, M. Kiamari, and B. H. Khalaj, "Distributed synchronization and beamforming in uplink relay asynchronous ofdma comp networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3471–3480, 2015.

[13] Y. Cheng, M. Pesavento, and A. Philipp, "Joint network optimization and downlink beamforming for comp transmissions using mixed integer conic programming," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 3972–3987, 2013.

[14] S. Shalmashi and S. B. Slimane, "Cooperative device-to-device communications in the downlink of cellular networks," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pp. 2265–2270, IEEE, 2014.

[15] R. A. Loodaricheh, S. Mallick, and V. K. Bhargava, "Joint resource optimization for ofdma cellular networks with user cooperation and qos provisioning," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, pp. 1264–1269, IEEE, 2014.

[16] H. Ju and R. Zhang, "User cooperation in wireless powered communication networks," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, pp. 1430–1435, IEEE, 2014.

[17] A. Ferragut and F. Paganini, "Network resource allocation for users with multiple connections: fairness and stability," *IEEE/ACM Transactions on Networking*, vol. 22, no. 2, pp. 349–362, 2014.

[18] D. Zhou and W. Song, *Multipath TCP for User Cooperation in Wireless Networks*. Springer, 2014.

[19] S. E. Deering, "Host extensions for ip multicasting." https://tools.ietf.org/html/rfc1054, 1988.

[20] D. Waitzman, S. Deering, and C. Partridge, "Distance vector multicast routing protocol." https://tools.ietf.org/html/rfc1075, 1988.

[21] J. Moy, "Multicast extensions to ospf." https://tools.ietf.org/html/rfc1584, 1994.

[22] "Multicast routing." Cisco Technology White Paper, 1995.

[23] T. Ballardie, P. Francis, and J. Crowcroft, "Core based trees (cbt)," *ACM SIGCOMM Computer Communication Review*, vol. 23, no. 4, pp. 85–95, 1993.

[24] N. Meghanathan, "Survey of topology-based multicast routing protocols for mobile ad hoc networks," *International Journal of Communication Networks and Information Security*, vol. 3, no. 2, p. 124, 2011.

[25] V. T. M. Do, L. Landmark, and Ø. Kure, "A survey of qos multicast in ad hoc networks," *Future Internet*, vol. 2, no. 3, pp. 388–416, 2010.

[26] E. M. Royer and C. E. Perkins, "Multicast operation of the ad-hoc on-demand distance vector routing protocol," in *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pp. 207–218, ACM, 1999.

[27] T. Ozaki, J. B. Kim, and T. Suda, "Bandwidth-efficient multicast routing protocol for ad-hoc networks," in *Computer Communications and Networks, 1999. Proceedings. Eight International Conference on*, pp. 10–17, IEEE, 1999.

[28] C.-K. Toh, G. Guichal, and S. Bunchua, "Abam: On-demand associativity-based multicast routing for ad hoc mobile networks," in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, vol. 3, pp. 987–993, IEEE, 2000.

[29] X. Zhang and L. Jacob, "Mzrp: an extension of the zone routing protocol for multicasting in manets," *J. Inf. Sci. Eng.*, vol. 20, no. 3, pp. 535–551, 2004.

[30] C.-C. Chiang, M. Gerla, and L. Zhang, "Shared tree wireless network multicast," in *Computer Communications and Networks, 1997. Proceedings., Sixth International Conference on*, pp. 28–33, IEEE, 1997.

[31] C.-W. Wu and Y. Tay, "Amris: A multicast protocol for ad hoc wireless networks," in *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, vol. 1, pp. 25–29, IEEE, 1999.

[32] J. Xie, R. R. Talpade, A. McAuley, and M. Liu, "Amroute: ad hoc multicast routing protocol," *Mobile networks and Applications*, vol. 7, no. 6, pp. 429–439, 2002.

[33] S.-J. Lee, M. Gerla, and C.-C. Chiang, "On-demand multicast routing protocol," in *Wireless Communications and Networking Conference, 1999. WCNC. 1999 IEEE*, vol. 3, pp. 1298–1302, IEEE, 1999.

[34] E. L. Madruga and J. Garcia-Luna-Aceves, "Scalable multicasting: the core-assisted mesh protocol," *Mobile Networks and Applications*, vol. 6, no. 2, pp. 151–165, 2001.

[35] J. Jin, H. Xu, and B. Li, "Multicast scheduling with cooperation and network coding in cognitive radio networks," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–9, IEEE, 2010.

[36] D. T. Ngo, C. Tellambura, and H. H. Nguyen, "Resource allocation for ofdm-based cognitive radio multicast networks," in *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*, pp. 1–6, IEEE, 2009.

[37] W. Kim, S. Y. Oh, M. Gerla, and J.-S. Park, "Cocast: multicast mobile ad hoc networks using cognitive radio," in *Military Communications Conference, 2009. MILCOM 2009. IEEE*, pp. 1–7, IEEE, 2009.

[38] C. Gao, Y. Shi, Y. T. Hou, H. D. Sherali, and H. Zhou, "Multicast communications in multi-hop cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 784–793, 2011.

[39] H. M. Almasaeid and A. E. Kamal, "Exploiting multichannel diversity for cooperative multicast in cognitive radio mesh networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 770–783, 2014.

[40] 3GPP, "TS 32.500: Telecommunication management; self-organizing networks (SON); concepts and requirements (release 12)." www.3gpp.org, Oct. 2014.

[41] S. Hämäläinen, H. Sanneck, and C. Sartori, *LTE self-organising networks (SON): network management automation for operational efficiency.* John Wiley & Sons, 2012.

[42] T. R. 3GPP, "32.541: Self-organizing networks (SON); Self-healing concepts and requirements (release 11)." www.3gpp.org, Sep. 2012.

[43] P. Koopman, "Elements of the self-healing system problem space." http://repository.cmu.edu/isr/675/, 2003.

[44] M. Amirijoo, L. Jorguseski, R. Litjens, and L.-C. Schmelz, "Cell outage compensation in lte networks: algorithms and performance assessment," in *Proc. of the 73rd IEEE Vehicular Technology Conference (VTC Spring 2011), Yokohama, Japan*, May 2011.

[45] F.-q. Li, X.-s. Qiu, L.-m. Meng, H. Zhang, and W. Gu, "Achieving cell outage compensation in radio access network with automatic network management," in *Proc. of the IEEE Global Communications Conference Workshops, Huston, TX, USA*, Dec. 2011.

[46] M. Putzke and C. Wietfeld, "Self-organizing ad hoc femtocells for cell outage compensation using random frequency hopping," in *Proc. of the 23rd IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2012), Sydney, Australia*, Sep. 2012.

[47] A. Apostolidis, D. Siouras, and N. Alonistioti, "An outage compensation algorithm for wlans," in *Proc. of the International Conference on Smart Communications in Network Technologies, Paris, France*, Jun. 2013.

[48] L. Kayili and E. Sousa, "Cell outage compensation for irregular cellular networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC 2014), Istanbul, Turkey*, Apr. 2014.

[49] W. Sang, D. Shen, W. Ren, and X. Shuai, "A survey of capacity in cooperative relay networks," in *Mobile Congress (GMC), 2011 Global*, pp. 1–8, IEEE, 2011.

[50] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in lte-advanced: deployment scenarios and operational challenges," *Communications Magazine*, vol. 50, no. 2, pp. 148–155, 2012.

[51] "4g lte comp, coordinated multipoint tutorial." http://www.radio-electronics.com/info/cellulartelecomms/lte-long-term-evolution/4g-lte-advanced-comp-coordinated-multipoint.php, [last online access: April. 2016].

[52] W.-F. Alliance, "Wi-Fi Direct: Portable Wi-Fi that goes with you anywhere." http://www.wi-fi.org/discover-wi-fi/wi-fi-direct.

[53] Qualcomm, "LTE Direct: Operator enabled proximity services." https://www.qualcomm.com/invention/technologies/lte/direct.

[54] B. Hamdaoui, T. Alshammari, and M. Guizani, "Exploiting 4g mobile user cooperation for energy conservation: challenges and opportunities," *IEEE Wireless Communications*, vol. 20, no. 5, pp. 62–67, 2013.

[55] H. Niu, L. Sun, M. Ito, and K. Sezaki, "User cooperation analysis under eavesdropping attack: A game theory perspective," in *Personal, Indoor, and Mobile Radio Communication (PIMRC), 2014 IEEE 25th Annual International Symposium on*, pp. 139–144, IEEE, 2014.

[56] F. Jiang, H. Tian, and P. Zhang, "A user cooperation stimulating strategy based on cooperative game theory in cooperative relay networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, p. 294942, 2009.

[57] S. Zhong, J. Chen, and Y. R. Yang, "Sprite: A simple, cheat-proof, credit-based system for mobile ad-hoc networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, pp. 1987–1997, IEEE, 2003.

[58] L. Anderegg and S. Eidenbenz, "Ad hoc-vcg: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents," in *Proceedings of the 9th annual international conference on Mobile computing and networking*, pp. 245–259, ACM, 2003.

[59] S. Zhong, L. E. Li, Y. G. Liu, and Y. R. Yang, "On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks: an integrated approach using game theoretic and cryptographic techniques," *Wireless networks*, vol. 13, no. 6, pp. 799–816, 2007.

[60] T. LeAnh, S. Ullah, N. H. Tran, S. S. Kim, S. I. Moon, and C. S. Hong, "Coalitional game theoretic approach for cooperation in heterogeneous cognitive wireless networks," in *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*, p. 86, ACM, 2016.

[61] O. A. Adeleke and M. F. Salleh, "Relay selection mechanism for user cooperation networks using a game-theoretic approach," *Journal of Electrical and Electronics Engineering Research*, vol. 5, no. 3, pp. 50–56, 2013.

[62] A. Ford, C. Raiciu, M. Handley, O. Bonaventure, *et al.*, "Tcp extensions for multipath operation with multiple addresses," *Internet Engineering Task Force, RFC*, vol. 6824, 2013.

[63] "Use Multipath TCP to create backup connections for iOS." https://support.apple.com/en-us/HT201373.

[64] "MultiPath TCP - Linux Kernel implementation." http://multipath-tcp.org/.

[65] V. Singh, S. Ahsan, and J. Ott, "Mprtp: multipath considerations for real-time media," in *Proceedings of the 4th ACM Multimedia Systems Conference*, pp. 190–201, ACM, 2013.

[66] "Multipath rtp (mprtp)." https://tools.ietf.org/html/draft-ietf-avtcore-mprtp-03, 2016.

[67] "White spaces (radio)." http://en.wikipedia.org/wiki/White_spaces_(radio).

[68] S. H. Alnabelsi and A. E. Kamal, "Resilient multicast routing in crns using a multilayer hyper-graph approach," pp. 2910–2915, 2013.

[69] S. Krishnamurthy, M. R. Thoppian, S. Kuppa, S. Venkatesan, R. Chandrasekaran, N. Mittal, and R. Prakash, "Time-efficient layer-2 auto-configuration for cognitive radios.," in *IASTED PDCS*, pp. 459–464, Citeseer, 2005.

[70] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[71] E. Zitzler, M. Laumanns, L. Thiele, E. Zitzler, E. Zitzler, L. Thiele, and L. Thiele, "Spea2: Improving the strength pareto evolutionary algorithm." Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK) Zürich, Switzerland, 2001.

[72] D. Pinto and B. Barán, "Solving multiobjective multicast routing problem with a new ant colony optimization approach," in *Proceedings of the 3rd international IFIP/ACM Latin American conference on Networking*, pp. 11–19, 2005.

[73] S. Bandyopadhyay, S. Saha, U. Maulik, and K. Deb, "A simulated annealing-based multiobjective optimization algorithm: Amosa," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 269–283, 2008.

[74] W. Wang, Y. Wang, X.-Y. Li, W.-Z. Song, and O. Frieder, "Efficient interference-aware tdma link scheduling for static wireless networks," in *Proceedings of the 12th annual international conference on Mobile computing and networking*, pp. 262–273, 2006.

[75] "cplex optimizer." http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/.

[76] T. R. 3GPP, "TR 36.814: Evolved universal terrestrial radio access, further advancements for E-UTRA physical layer aspects (release 9)." www.3gpp.org, Mar. 2010.

[77] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4409–4418, Aug. 2008.

[78] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[79] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[80] P. Baracca, F. Boccardi, and N. Benvenuto, "A dynamic clustering algorithm for downlink comp systems with multiple antenna ues," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–14, Aug. 2014.

[81] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming." http://cvxr.com/cvx/download/, 2008.

[82] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," *The Journal of finance*, vol. 16, no. 1, pp. 8–37, 1961.

[83] IEEE-SA, "Ieee 802.11n-2009amendment 5: Enhancements for higher throughput." https://standards.ieee.org/findstds/standard/802.11n-2009.html, 2009.

[84] D. B. Faria *et al.*, "Modeling signal attenuation in ieee 802.11 wireless lans-vol. 1," *Computer Science Department, Stanford University*, vol. 1, 2005.

[85] Y. Jie, A. Kamal, and M. Alnuem, "User cooperation solution of multipath streaming application using auction theory," in *Global Communications Conference (GLOBECOM), 2016 IEEE*, pp. 1–6, IEEE, 2016.