Optimizing the Transaction Times for NOMA-Based Mobile Edge Computing

By

Anam Yasir Kiani

Fall 2017-MS(EE-9)-00000204070

Supervisor

Dr. Syed Ali Hassan

Department of Electrical Engineering

A thesis submitted in partial fulfillment of the requirements for the degree

of Masters of Science in Electrical Engineering (MS EE)

In

School of Electrical Engineering and Computer Science,

National University of Sciences and Technology (NUST),

Islamabad, Pakistan.

(October 2019)

Approval

It is certified that the contents and form of the thesis entitled "**Optimizing the Transaction Times for NOMA-Based Mobile Edge Computing**" submitted by **Anam Yasir Kiani** has been found satisfactory for the requirement of the degree.

Advisor: **Dr. Syed Ali Hassan**
Signature: ____________________
Date: ____________________

Committee Member 1: **Dr. Hassaan Khaliq Qureshi**
Signature: ____________________
Date: ____________________

Committee Member 2: **Dr. Fahd Ahmad Khan**
Signature: ____________________
Date: ____________________

Committee Member 3: **Dr. Sajjad Hussain**
Signature: ____________________
Date: ____________________

Abstract

Non orthogonal multiple access (NOMA) and mobile edge computing (MEC) are evolving as key enablers for future fifth generation (5G) networks as this combination can provide high spectral efficiency, massive connectivity, improved quality-of-service (QoS), and lower latency. In this research, we aim to minimize the transaction time difference of NOMA paired users, where the users have disparate amount of data to send to the MEC servers, which have variable computing capabilities. The transaction time includes both the transmission time of data and the computational time of the servers. The equalization of transaction time for paired users reduces the wastage of both frequency and computational resources. The transaction time difference is minimized by optimizing the transmission power of the users and the computational resources of severs using a successive convex approximation method. The results show that the proposed scheme reduces the transaction time difference of the two paired users from hundreds of seconds to just a few seconds; thereby conserving the resources. .The percentage improvement in effective throughput of the system is shown to be 19% on average.

Dedication

To Mamu and his whole family who have sincerely and selflessly paved ways to my destination.
To Papa, my best friend, who made me dream beyond limits and utilized all of his resources to curve them to reality.
To Mama, the lady who compromised everything to embellish her children with jewel of education.
To my siblings, Samman, Aiman, Izan and Qais who bring me joy and happiness.

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: **Anam Yasir Kiani**

Signature: ________________

Acknowledgment

All praises to the Lord of wisdom (Al - Hakeem) and knowledge (Al - Aleem), who showered countless blessings on me and brought such things to the canvas of reality which were beyond my imaginations. The heartiest reverence and respect to the Holy Prophet Hazrat Muhammad (Peace Be Upon Him), who took humanity from darkness to light and showed the righteous path.

My appreciation and gratitude to National University of Sciences and Technology for well preparing students for challenging practical life ahead.

I am honored to work as MS thesis student under the supervision of Dr. Syed Ali Hassan, who has been an inspiration for me. You made the hardest things, simple and easy with your expertise and knowledge. You have been source of constant motivation and support. Your valuable advises and sincere guidance is regarded at highest esteem. I am thankful to you for making the whole experience worthy and commendable.

I would seek this opportunity to gratify Rida Zia-ul-Mustafa for always being there for me and helping me at her best. Lastly, I am indebted to Warda Zahoor for affection and support through the stressful times.

Table of Contents

List of Figures

List of Tables

List of Acronyms

OFDM	Orthogonal frequency division multiplexing
OMA	Orthogonal Multiple Access
NOMA	Non-orthogonal Multiple Access
MEC	Mobile Edge Computing
5G	Fifth Generation
QoS	Quality-of-Service
D2D	Device-to-Device
EHGRA	Energy Harvesting and Gain-Base Resource Allocation
SWIPT	Simultaneous Wireless Information and Power Transfer
mmWave	Millimeter-Wave
IoT	Internet-of-things
DBs	Data Base Stations
EE	Energy Efficiency
UHF	Ultra High Frequency

MPA	Message Passing Algorithm
ML	Maximum Likelihood
SIC	Successive Interference Cancellation
BS	Base Station
AP	Access Point
AWGN	Additive White Gaussian Noise
CR-NOMA	Cognitive Radio Non-orthogonal Multiple Access
SCMA	Sparse Code Media Access
PDMA	Pattern Division Multiple Access
EECO	Energy Efficient Computation Offloading
FiWi	Fiber-Wireless
STBC	Space Time Block Coding
SBD	Successive Bandwidth Division
MIMO	Multiple-Input Multiple-Output

Chapter 1

Introduction

1.1 Motivation

Some of the important concerns in today's wireless networks are limited resources and media arbitration. The medium used for the transmissions is being shared by millions of devices and immense traffic is flowing over it, which is expected to increase by 1000 folds in the next decade. To better serve the access mechanism, a number of techniques such as time, frequency and code division multiplexing have been used until now. Orthogonal frequency division multiplexing (OFDM) is also being used in order to serve more number of users with reduced interference offering high spectral efficiency. In all these techniques, a single user is served in a single orthogonal resource block as in Fig. 1.1, however, these techniques are unable to reach the lower bound of capacity for broad-cast and multiple access channels [1]. The bottleneck lies in random nature of the channel, i.e., if some bandwidth resources are allocated to a user with bad channel conditions, any orthogonal

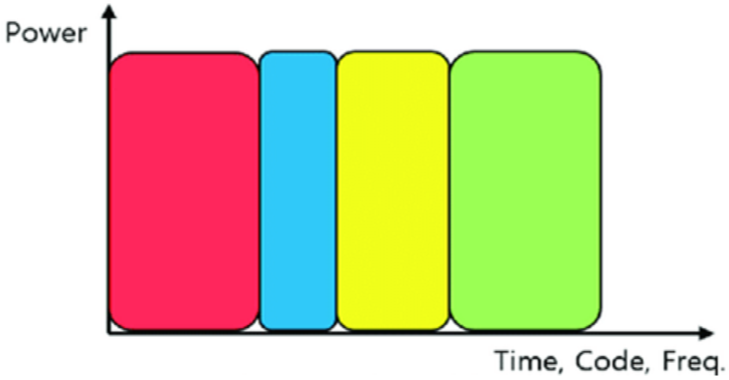

Figure 1.1: Orthogonal Multiple Access (OMA)

technique would result in reduced spectral efficiency. Also note that spectral efficiency is not the only issue; with the evolution of technologies, the demands for quality-of-service (QoS) are constantly varying. The need of the time encompasses high connectivity, reliability, ultra-low latency, improved fairness and high throughput. Non-orthogonal multiple access (NOMA) has been introduced to muddle through the demands of the epoch. One of main purposes of NOMA is to serve multiple users by utilizing the same resource block as shown in Fig. 1.2. It can also be integrated with the existing orthogonal media access technologies. In NOMA, multiple signals at the receiver

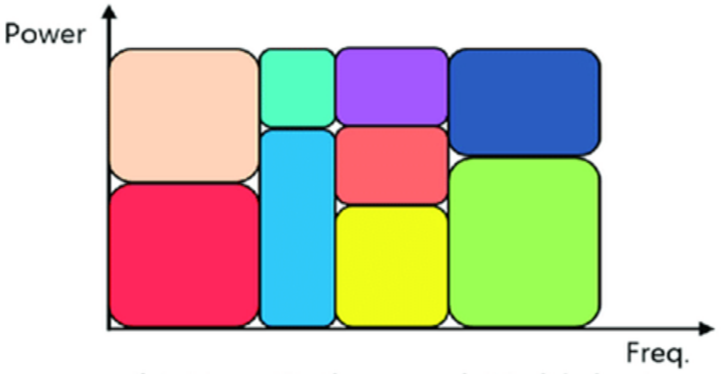

Figure 1.2: Non-orthogonal Multiple Access (NOMA)

are separated with message passing algorithm (MPA), maximum likelihood (ML) and successive interference cancellation (SIC) techniques [2]. NOMA provides a balanced trade-off between the system throughput and user fairness [3] and is being envisioned as a key technology in 5G networks [4].

5G enabled devices are expected to have latency constraint and have computationally complex applications running on them [5]. For such applications, limited power and computational capacity of mobile devices pose a problem, which can be solved by using mobile edge computing (MEC) [6]. MEC offloads computationally intensive data to base stations (BSs) and access points (APs) that are equipped with powerful servers. Servers being available at the edges result in reduction of delay and improvement of computational efficiency [7]. The advantages of both techniques (i.e., NOMA and

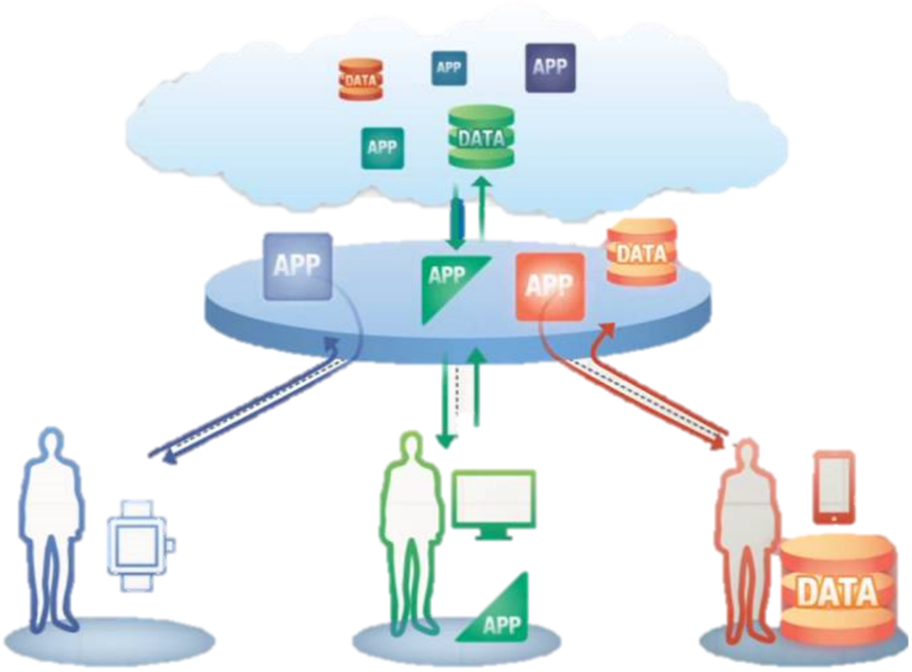

Figure 1.3: Mobile Edge Computing (MEC)

MEC) have drawn considerable attention of researchers recently. In NOMA-MEC, paired users offload their data to mobile edge computing servers by using non-orthogonal multiple access technique.

The time taken to process data (i.e., offloaded to MEC servers) is not equal for each NOMA paired user (i.e., dependent upon amount of offloaded data, channel conditions of paired users etc.). This inequality leads to under utilization of resources and reduced spectral efficiency. Hence a mechanism is needed to conserve scare resources and improve the system performance.

1.2 Contribution

A lot of work is being done to investigate different aspects of NOMA-MEC. In this work, we focus on the transaction time difference of NOMA paired users, i.e., offloading data to MEC servers. Transaction time is the sum of transmission time and computational time. We propose a scheme to reduce the transaction time difference between paired users as well as optimize the transaction time of individual users to conserve both frequency and computational resources. The prominent research contributions are as under:

1. We provide an overview of work have been done in context of NOMA-MEC (Chap. 2).

2. We study the transaction time in term of transmission and computation time, where we reduce the difference between the transmission and computation time separately, to minimize the difference between the transaction time of paired users.

3. We formulate the problem of minimizing the transaction time difference of NOMA paired users as two independent convex optimization problems. Firstly, we optimize the power allocation to reduce the difference between the transmission time of paired users, secondly we optimize the server's core allocation with the objective of computational time difference minimization.

4. We evaluate the performance of our proposed scheme (i.e., NOMA-MEC with optimized power and core allocation) by comparing with conventional NOMA power allocation and equal core allocation, optimized power and equal core allocation and lastly with optimized power and random core allocation schemes. The performance metrics used for evaluation are: transaction time difference and system effective throughput. The transaction time difference for the paired users is computed for aforementioned schemes, for a range of offloaded data amounts and complexities ratios. However system effective throughput is investigated for fixed offloaded data complexities and a range of offloaded data amounts. The proposed scheme outperforms the other schemes in both transaction time difference minimization and increase in system's effective throughput.

5. Lastly, we summarize the extracts of our research and identify the future directions for study in this area.

Chapter 2

Groundwork on NOMA-MEC

The biggest challenge to today's communication world is the scarcity of resources as number of users, devices and traffic is growing by leaps and bounds. Along with this the demands for QoS are continuously varying (i.e., improved user fairness, massive connectivity, higher reliability, minimal delays etc.) for future 5G networks. A lot of work in done in this regard. [8] evaluated the performance of three-tier heterogeneous 5G network by using model developed by using stochastic geometry. Device-to-device (D2D) communication is also studied for 5G networks along its potential and challenges [9]. Energy harvesting and gain-base resource allocation (EHGRA) algorithm is proposed in [10] to increase sum rates in D2D networks enabled with energy harvesting. Resources are shared dynamically in [11], to increase energy efficiency (EE) of hybrid D2D network in micro as well as millimeter Wave (mmWave). A simultaneous wireless information and power transfer (SWIPT) approach for energy harvesting is used for cooperative D2D [12]. A limit on number of active D2D connections in a circular region with a given outage probability is determined in [13], minimum transmission power is also computed to guaranty required QoS. [14] proposed energy efficient algorithm for fast neighbour discovery in mmWave based D2D network. By fulfilling QoS requirements of D2D pairs, users are either served by mmWave or micro wave with the objective of maximizing EE in [15]. In [16] coding is applied to D2D relay based network to improve quality. In cooperative D2D networks relay selection is very important task. In [17] energy efficient relay selection mechanism is devised. Game theory is also exploited to improve the performance of 5G networks. Energy efficiency for heterogeneous 5G networks (HetNet) is improved by using game theory for optimization of resources allocation [18]. In [19] hierarchy based game theory technique is used for optimal resource allocation in resilient 5G networks to optimize sum rates and coverage probability. NOMA based D2D is studied in [20], to make single transmitter communicate to multiple receives. MEC architecture having D2D abilities is proposed by [21] for 5G networks. Dual band allocation scheme is utilized at data base stations (DBs) for 5G internet-of-things (IoT) networks to improve system performance [22]. In this study it is shown that the dynamic management of spectrum (i.e., licensed and unlicensed mmWave) helps in efficient spectrum utilization. The performance of hybrid network in which mmWave and ultra high frequency (UHF) coexist is studied in [23] and it is established that mmWave achieves higher rates and coverage.

NOMA's potential for future 5G networks is investigated by [24]. In [24] in addition to research directions standardization is also discussed. NOMA is investigated as key technology for 5G networks by [25], as it improves spectral efficiency, EE and sum capacity [26]. In [27] NOMA is discussed as medium access technology for 5G radio access networks. [28] investigated the potential of MEC for 5G IoT networks and provided an overview of its architecture. In [29] technical perspective of MEC is studied. In short, extensive research is being carried out on NOMA and MEC with future 5G networks in view. The literature on NOMA-MEC is being categorized into three sections that are as under:

2.1 Non-orthogonal Multiple Access

The requirements for future networks have made NOMA a potential candidate for future 5G networks [30]. The main advantage of NOMA is to serve multiple users by utilizing the same resource block i.e. frequency, time. NOMA uses different domains for medium access like power, patterns, sparse code, lattice partition etc. The novelty in NOMA is removal of orthogonality. It can also be integrated with the existing orthogonal media access technologies. In NOMA multiple signals at the receiver are separated with message passing algorithm (MPA), maximum likelihood (ML) and successive interference cancellation (SIC) techniques. NOMA as number of variants with different characteristics and advantages.

NOMA is being used for single resource block as well as multi resource blocks. NOMA used for single resource block is referred as Power Domain NOMA. In this technique multiple users are served with in same time slot, Orthogonal Frequency Division Multiple Access (OFDMA) subcarrier frequency etc. by assigning different powers to different users [31]. Higher power is given to the user with the bad channel conditions and low power is allotted to the user with good channel conditions but still target QoS is not guaranteed. To ensure QoS a variant of power domain NOMA exist known as Cognitive Radio NOMA (CR-NOMA) [2]. In CR-NOMA the power allocation is done in such a manner to fulfill the predefined needs of the QoS. The user with the poor channel conditions is considered as the primary user of the cognitive radio. After allocating enough power to the primary user the left over power is assigned to the other user; one with better channel. At the receiver the primary user decodes the signal with the higher power intended for it and treat the residue signal as noise. The receiver with the good channel conditions first decodes the signal of the primary user; subtract it from the originally received signal in order to get the signal meant for it. In single carrier NOMA all the users have to be compensated within a single resource block. Accommodation of all the users that is immense in number within same block results in decoding delays and the system complexity. The receiver of the weak signal first has to decode all the powerful signals and subtract them from the mixture of the signals received in order to get its signal (i.e. Successive Interference Cancellation).

In Multi-carrier NOMA, also referred as hybrid NOMA, users are divided into number of groups and each group is allocated with an orthogonal resource block that is used to serve multiple users with in the group. Those users are placed in the same group that have most different channel conditions. Multi-carrier NOMA includes techniques like Sparse Code Media Access, Pattern Division Multi Access etc. [2]. In Sparse Code Media Access (SCMA) the bit stream is mapped to different sparse code words and different code words for all users are spread across multiple subcarriers. The number of sub carriers allocated to each user is less than the total number of subcarriers. In Pattern Division Multiple Access (PDMA) the performance is determined by the scheme of resource allocation matrix that is known as pattern matrix. Multiplexing is based on code, power, spatial domain or combination of these, which increases the diversity and reduces the overlap among different users. Successive Bandwidth Division (SBD) NOMA is another hybrid NOMA approach. It combines both conventional NOMA and OMA technique to model divide and allocate mechanism [3]. It reduces the complexity of receiver and increases signal-to-interference plus noise ratio [32]. SBD is also applied to multiple-input multiple-output (MIMO) uplink NOMA to improve the system performance [33]. NOMA is also studied in combination with other emerging technologies like mmWave [34] to improve performance of cellular systems, backscatter communication for green IoT networks [35] etc. NOMA is inherited with receiver side complexity, to reduce the number of SICs, outage probability and improve sum rates distributed space time block coding (STBC) is introduced in cooperative NOMA [36], [37]. NOMA performance in term of outage probability and ergodic sum rates is better than OMA if power coefficients are chosen appropriately [38].

In short NOMA is a potential technology for the next generation systems i.e. 5G systems.

2.2 Mobile Edge Computing

The evolution of technology has enabled the users to perform storage and computation extensive tasks on their mobile devices. The performance of mobile devices is limited by their processing power and storage capabilities. This deficiency is aided by the cloud services. Mobile devices offload their complex tasks to the cloud. The cloud servers are hosted at centralized locations, hence inducing delays and network congestion. This problem is resolved by MEC. In MEC computing and storage resources are provided at BSs and APs, reducing delay and network congestion [39]. MEC is declared as key technology for 5G networks in [40]. MEC is being investigated at number of frontiers like energy efficient computation offloading (EECO) [41], real time and context aware applications as use case [42], for fiber-wireless (FiWi) access networks [30] etc. MEC has found a number of interesting combinations with existing technologies like NOMA, wireless power transfer, energy harvesting etc. [43]. MEC is being foreseen as key technology in 5G networks.

2.3 NOMA-MEC

NOMA-MEC has recently drawn considerable attention of researchers. A lot of work is being done to fully utilize the potential of both the techniques (i.e., NOMA and MEC) to achieve massive connectivity, improved QoS, lower latency, user fairness, higher reliability and enhanced spectral efficiency. The aspects of NOMA-MEC investigated so far are mostly associated with energy conservation and delay minimization. For instance, [44] formulated delay minimization for NOMA-MEC data offloading as a form of fractional programming. In this scheme the pure NOMA is also compared with hybrid NOMA and OMA for data offloading purpose. In [45], energy consumption of MEC users utilizing uplink NOMA is reduced by optimizing user clustering, power, frequency and computational resource allocation. [46] proved that the total energy (i.e., sum of transmission and computation energy) minimization is a convex problem and the authors solved it by an iterative algorithm. [47] reduced the total system energy by optimizing allocated power, transmission time and offloaded task portions. Optimal time allocation and offloading task partitions are obtained by an iterative algorithm, where the algorithm is implemented by using successive convex optimization. In [48], energy consumption is reduced by jointly optimizing power and time allocation. This is achieved by formulating the problem to a form of geometric programming. [49] studied energy harvesting for full duplex NOMA-MEC, where total energy consumption is minimized by efficient power allocation, time scheduling and computing resources allocation. In [50] NOMA-MEC offloading scheme is devised that can work in three modes, successful computation probability is also increased by optimizing power allocation and offloading time consumption. NOMA transmission time and and MEC work offloading is optimized by layered structured algorithm to reduce the delay in [51]. Bisection search iterative algorithm is proposed by [52], to reduce latency in partial offloading scheme for NOMA-MEC.

Chapter 3

Minimizing the Transaction Time Difference for NOMA-MEC

3.1 System Model

The aim of devised scheme is to reduce the transaction time difference of NOMA paired users offloading data to MEC servers, in order to conserve the resources (i.e., frequency and computation) and to increase system throughput.

A single cell is considered with $2N$ number of users, which are served by a single BS. The BS is equipped with MEC server having C_T number of cores each with a computational capability of f cycles/sec and total system bandwidth is B_T. Hybrid NOMA technique is used to pair the users into N NOMA clusters, where each cluster has two users. A single cluster with users u_1 and u_2 is considered for study.

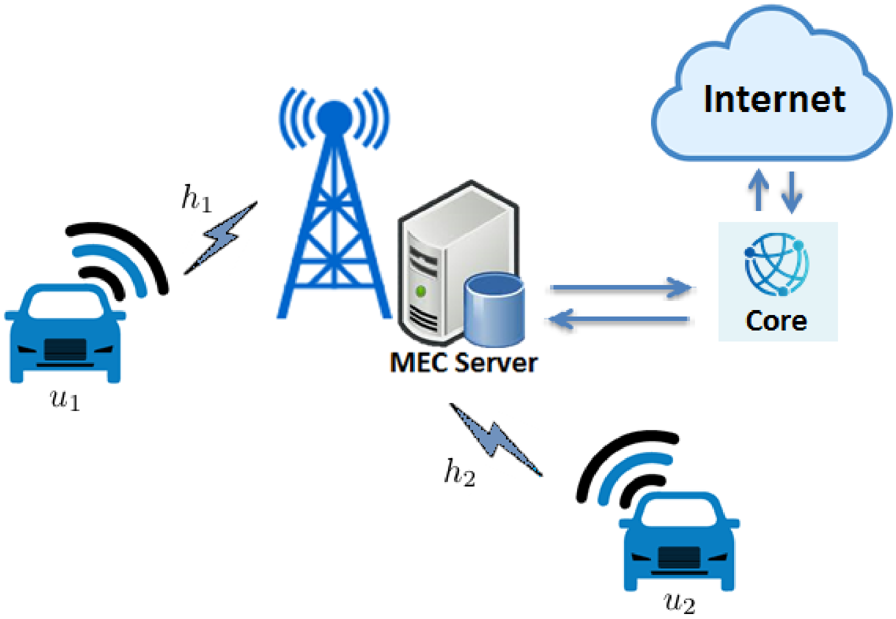

Figure 3.1: A snapshot of the network model

The user u_1 is located at a distance $dist_1$ from BS whereas u_2 is located at a distance $dist_2$ from BS, such that $dist_1 < dist_2$. Without the loss in generality, we assume that u_1 is a strong user with allocated power p_1 and u_2 is a weak user with allocated power p_2, such that $p_1 < p_2$. Let $p_{1,\max}$ and $p_{2,\max}$ are the maximum transmission powers that can be allocated to u_1 and u_2, respectively. We assume that d_1 bits are offloaded by u_1 and d_2 bits are offloaded by u_2 to the MEC server. Complete offloading scheme is considered, where no local computation is being performed. Each bit offloaded by u_1 requires c_1 cycles and that of u_2 requires c_2 cycles for computation at MEC server. The computational complexity of offloaded data is dependent upon offloaded data type (i.e, video data requires more CPU cycles as compared to text data). The total system bandwidth is divided into N number of fre-

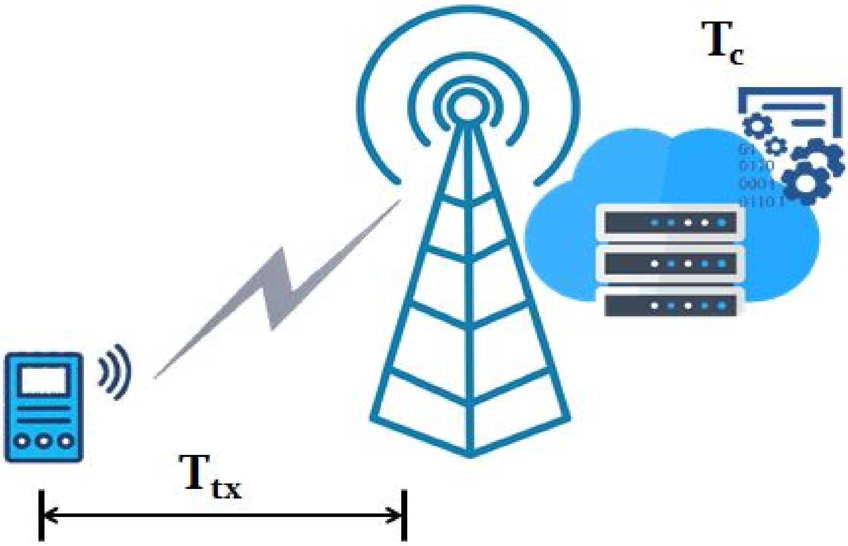

Figure 3.2: Transaction time

quency resource blocks. A single frequency resource block with bandwidth $B_w = B_T/N$ is allocated to a NOMA cluster and shared by paired users, similarly the cores at MEC servers are divided into N number of computational resource blocks and are allocated to NOMA clusters. The allocated computational resources of a cluster (i.e., $C_t = C_T/N$) are divided among the paired users, depending upon the complexity and amount of data being offloaded by them. Let u_1 and u_2 are allocated with n_1 and n_2 cores, respectively. The transaction time of the i^{th} user illustrated in Fig. 3.2 is given as

$$T_i = T_{tx_i} + T_{c_i}, \qquad i \in \{1, 2\}, \tag{3.1}$$

where T_{tx_i} is the transmission time and T_{c_i} is the computational time of the i^{th} user, respectively. The transmission time for the i^{th} user is

$$T_{tx_i} = \frac{d_i}{R_i}, \tag{3.2}$$

where R_i is the data rate of the i^{th} user. The data rates are dependent upon the effective channel gains (i.e., h_1, h_2) and the allocated powers (i.e., p_1, p_2), such that

$$R_1 = B_w \log_2 \left(1 + \frac{p_1 h_1}{p_2 h_2 + \sigma^2}\right), \tag{3.3}$$

$$R_2 = B_w \log_2 \left(1 + \frac{p_2 h_2}{\sigma^2}\right), \tag{3.4}$$

where σ^2 is the power spectral density of noise and the effective channel gain for i^{th} user is

$$h_i = \frac{\widetilde{h_i}}{dist_i^{\rho}}, \tag{3.5}$$

where $\widetilde{h_i}$ is the exponential channel gain (corresponding to Rayleigh fading) of i^{th} user and ρ is the path loss exponent. The amount of data offloaded and the effective channel gains are associated with the paired users, however, the powers are optimized to reduce their transaction time difference. Similarly, the computational time for the i^{th} user is given by

$$T_{c_i} = \frac{d_i c_i}{n_i f}, \tag{3.6}$$

where n_i is the number of cores allocated to the user i and f is the computational capacity of each MEC core. For a given paired users, d_i, c_i and f are fixed. The number of computational resources allocated to the i^{th} user is optimized to balance the load across the cores in order to reduce the difference between transaction time. By manipulating, it is inferred that T_1 is equal to T_2 if

$$\begin{aligned}
\frac{d_1}{R_1} + \frac{d_1 c_1}{n_1 f} &= \frac{d_2}{R_2} + \frac{d_2 c_2}{n_2 f} \\
\frac{n_1 f d_1 + d_1 c_1 R_1}{n_1 f R_1} &= \frac{d_2 n_2 f + d_2 c_2 R_2}{n_2 f R_2} \\
\frac{d_1 \left(n_1 f + c_1 R_1\right)}{n_1 R_1} &= \frac{d_2 \left(n_2 f + c_2 R_2\right)}{n_2 R_2} \\
\frac{d_1}{d_2} = \left(\frac{R_1}{R_2}\right) &\left(\frac{n_1 n_2 f + n_1 c_2 R_2}{n_1 n_2 f + n_2 c_1 R_1}\right) \qquad (3.7)
\end{aligned}$$

From (3.7), we can divide the original formulated problem into two independent sub-problems and reformulate it as T_1 is equal to T_2, if T_{tx_1} is equal to T_{tx_2} as well as T_{c_1} is equal to T_{c_2}.

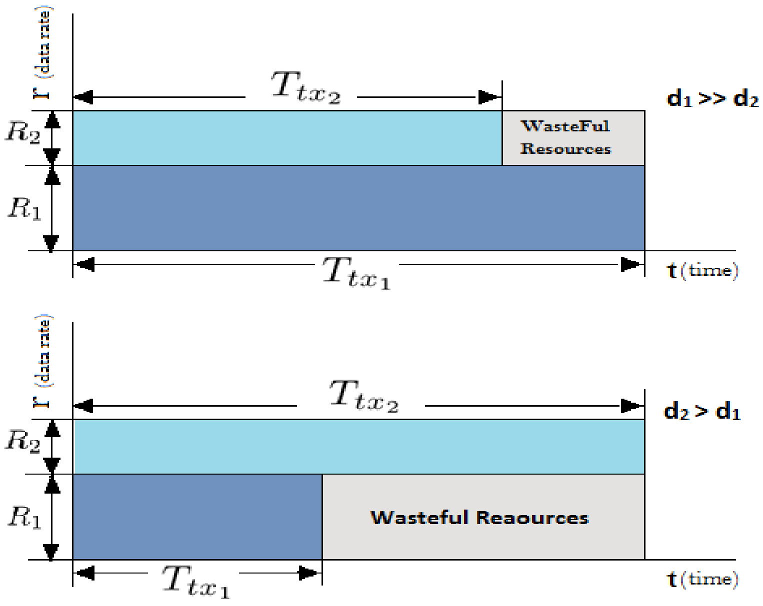

Figure 3.3: Unequal Transmission Time and Wasteful Resources

It is evident from Fig. 3.3 that unequal transmission time results in wastage of allocated frequency resources. It can be seen that for $\delta_{tx} = |T_{tx_1} -$

$T_{tx_2}|$ amount of time, the resources are under-utilized, i.e., a new NOMA signal cannot be initiated. As this difference increases, the spectral efficiency of the network decreases. The transmission time, T_{tx}, is equal for both the users, if

$$\frac{d_1}{d_2} = \frac{R_1}{R_2}, \tag{3.8}$$

where R_1 and R_2 are the data rates of both users u_1 and u_2, respectively. Similarly, the disparity in amount and computational complexity of data offloaded by paired users (i.e., allocated with equal number of cores) results in wastage of allocated computational resources as shown in Fig. 3.4.

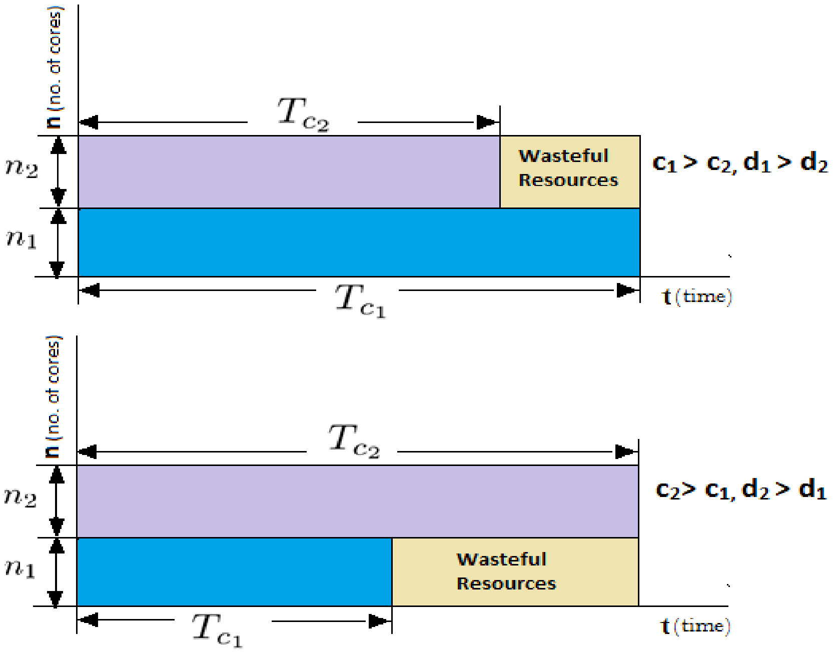

Figure 3.4: Unequal Computational Time and Wasteful Resources

The computational time difference $\delta_{tc} = |T_{tc1} - T_{tc2}|$ for both the users is zero, if

$$\frac{d_1 c_1}{n_1 f} = \frac{d_2 c_2}{n_2 f}, \tag{3.9}$$

which can be written in simplified form as

$$\frac{d_1}{d_2} = \frac{n_1 c_2}{n_2 c_1} \tag{3.10}$$

3.2 Optimization Problem Formulation

The original problem described in the previous section is given by

$$\begin{aligned}
(\mathbf{P}) \qquad & \min \quad \lambda, && (3.11a)\\
& s.t. \quad \left(\frac{d_1}{R_1} + \frac{d_1 c_1}{n_1}\right) - \left(\frac{d_2}{R_2} + \frac{d_2 c_2}{n_2}\right) \leq \lambda, && (3.11b)\\
& \left(\frac{d_2}{R_2} + \frac{d_2 c_2}{n_2}\right) - \left(\frac{d_1}{R_1} + \frac{d_1 c_1}{n_1}\right) \leq \lambda, && (3.11c)\\
& \lambda \geq 0 && (3.11d)
\end{aligned}$$

The problem is formulated as two independent optimization sub-problems.The results for original optimization problem and sub-problems are equivalent. The objective of the first optimization problem is to minimize the transmission time difference of given paired users with known d_i's, by optimizing the power allocation. From equations (3.3) and (3.4), we have

$$R_1 + R_2 = B_w \log_2\left(1 + \frac{p_1 h_1 + p_2 h_2}{\sigma^2}\right) \tag{3.12a}$$

$$\begin{aligned}
R_1 &\leq B_w \log_2\left(1 + \frac{p_1 h_1}{p_2 h_2 + \sigma^2}\right)\\
&= B_w \log_2\left(1 + \frac{p_1 h_1 + p_2 h_2}{\sigma^2}\right) - R_2,
\end{aligned} \tag{3.12b}$$

$$R_2 \leq B_w \log_2 \left(1 + \frac{p_2 h_2}{\sigma^2}\right), \tag{3.12c}$$

where $0 \leq p_i \leq p_{i,\max}$ $i \in \{1, 2\}$, the power allocated to both the users, p_i, is positive and less than the respective maximum, i.e., $p_{i,\max}$. For the objective with power allocation, we introduce a new variable μ and hence the problem of minimizing the transmission time difference of paired users can be reformulated as

$$\begin{aligned}
(\mathbf{P1}) \qquad & \min \quad \mu, && (3.13\text{a})\\
& s.t. \qquad \frac{d_1}{R_1} - \frac{d_2}{R_2} \leq \mu, && (3.13\text{b})\\
& \qquad\qquad \frac{d_2}{R_2} - \frac{d_1}{R_1} \leq \mu, && (3.13\text{c})\\
& \qquad\qquad \mu \geq 0 && (3.13\text{d})
\end{aligned}$$

where the objective function (3.13a) is subjected to data rate (3.12b, 3.12c) and power constraints. By manipulating (3.13b), we get

$$\mu R_1 R_2 \geq \mu \alpha_1 \geq \alpha_2^2 \geq d_1 R_2 - d_2 R_1, \tag{3.14}$$

where α_1 and α_2 are real valued variables, having values such that inequality holds. The equation (3.14) is equivalent to

$$R_1 R_2 \geq \alpha_1, \tag{3.15a}$$

$$\begin{bmatrix} \mu & \alpha_2 \\ \alpha_2 & \alpha_1 \end{bmatrix} \succeq 0, \tag{3.15b}$$

$$\alpha_2^2 \geq d_1 R_2 - d_2 R_1, \tag{3.15c}$$

where (3.15b) is positive semi-definite matrix, which is solved by applying successive convex approximation to (3.15c) as

$$2\alpha_2^{(j)}\alpha_2 - (\alpha_2^{(j)})^2 \geq d_1 R_2 - d_2 R_1, \tag{3.16}$$

where α_2 is updated in each iteration and j shows the number of iteration. The equation (3.15a) is rewritten as $R_1 R_2 \geq \beta^2$ and $\beta^2 \geq \alpha_1$. Similarly, the objective of the second optimization problem is to minimize the computational time difference of given paired users with known d_i's and c_i's by optimizing the core allocation. The computational resources allocation problem can be formulated by introducing new variable ζ, that represents the computational time difference such that

$$\begin{aligned}
(\mathbf{P2}) \qquad \min \quad & \zeta, && (3.17a)\\
s.t. \quad & \frac{d_1 c_1}{n_1} - \frac{d_2 c_2}{n_2} \le \zeta, && (3.17b)\\
& \frac{d_2 c_2}{n_2} - \frac{d_1 c_1}{n_1} \le \zeta, && (3.17c)\\
& \zeta \ge 0, && (3.17d)
\end{aligned}$$

where the objective function (3.17a) is subject to constraints $0 < n_i < C_t$, the number of cores allocated to individual user is greater than zero and less than total cores allocated to the cluster and $n_1 + n_2 \le C_t$, the sum of cores allocated to both the users is less than total cores allocated to the cluster. The number of cores allocated to individual user must be greater than zero to ensure the minimum requirement of the user. By manipulating (3.17b), we get

$$\zeta n_1 n_2 \ge \zeta \gamma_1 \ge \gamma_2^2 \ge (d_1 c_1) n_2 - (d_2 c_2) n_1, \tag{3.18}$$

where γ_1 and γ_2 are variables with real values. The equation (3.18) implies

$$n_1 n_2 \ge \gamma_1, \tag{3.19a}$$

$$\begin{bmatrix} \zeta & \gamma_2 \\ \gamma_2 & \gamma_1 \end{bmatrix} \succeq 0, \tag{3.19b}$$

$$\gamma_2^2 \ge (d_1 c_1) n_2 - (d_2 c_2) n_1, \tag{3.19c}$$

The positive semi-definite matrix (3.19b) is solved by applying successive convex approximation to (3.19c) as

$$2\gamma_2^{(j)}\gamma_2 - (\gamma_2^{(j)})^2 \geq (d_1c_1)n_2 - (d_2c_2)n_1, \tag{3.20}$$

where γ_i's are updated in each iteration and j shows the number of iteration. From (3.19a), we have $n_1n_2 \geq \eta^2$ and $\eta^2 \geq \gamma_1$. For a given pair of users, we obtain optimal values of power and number of cores once the optimization is performed. These parameters result in minimization of transaction time difference, which is illustrated in next chapter.

Chapter 4

Performance Evaluation

4.1 Simulation Playground

The efficiency of the proposed scheme is evaluated by comparative analysis with other schemes through numerical simulations. The software used for simulations is MATLAB (R2015b). The maximum power for u_1, $p_{1,\max}$, is 2W and of u_2, $p_{2,\max}$, is 4W. Initially p_1 is 1W and p_2 is 2W, as in NOMA more power is given to weak user as compared to the stronger one. The $dist_1$ and $dist_2$ are 200 m and 600 m, respectively. The cluster bandwidth is 200 kHz and the path loss exponent is 3.8. The transaction time difference is considered for three different approaches namely: Power Optimization with Equal Core Allocation (A), Power Optimization with Random Core Allocation (B) and proposed Power Optimization with Optimal Core Allocation (C). The power and core optimization is achieved by successive convex approximation as discussed in the previous chapter. These schemes are also investigated in term of impact on system effective throughput.

Parameter	**Value**	**Parameter**	**Value**
p_{1max}	2 W	p_{2max}	4 W
C_t	20	B_w	200 kHz
$dist_1$	200 m	$dist_2$	600 m

Table 4.1: Simulation Parameters

4.2 Transaction Time Difference

In equal core allocation, the cores are equally divided between the paired users, i.e., $n_1 = n_2$. In random core allocation, the cores are randomly divided between the paired users $n_1 = \kappa C_t$ and $n_2 = (1 - \kappa)C_t$, where, κ is from uniform random distribution varying from 0 to 1. While for proposed scheme both powers and cores are obtained by solving optimization problems. The ratio of offloaded data amount, i.e., d_2/d_1 and complexity, i.e., c_2/c_1 is varied to study their impact on the transaction time.

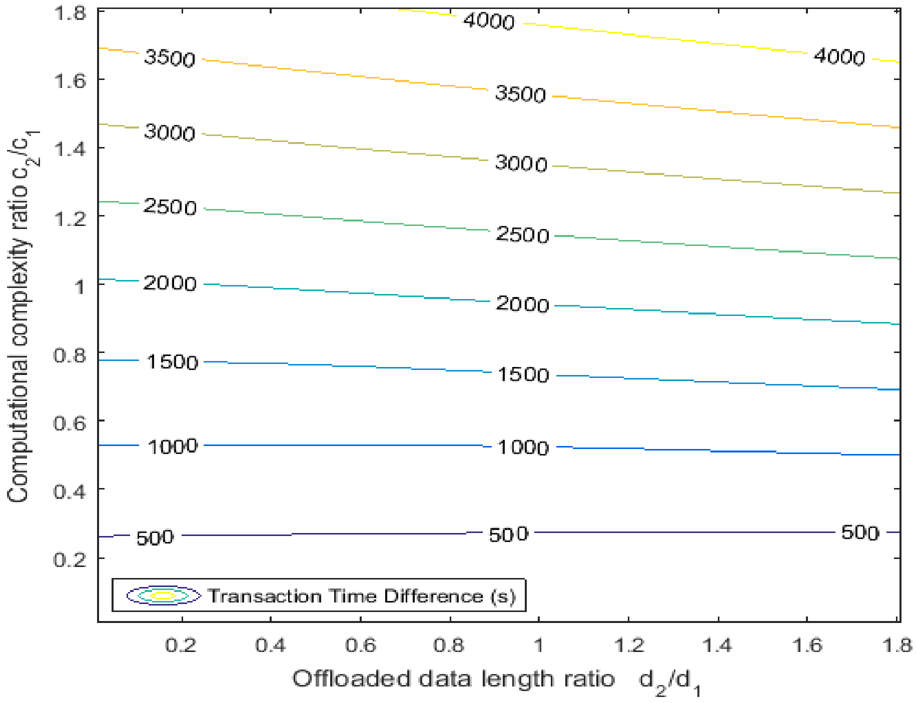

Figure 4.1: Impact of Computational Complexity ratio (c_2/c_1) and Offloaded Data Length ratio (d_2/d_1) on Transaction Time Difference without Power Optimization for Equal Core Allocation [*Benchmark (or Non-Optimization) Case*]

Fig. 4.1 depicts the transaction time difference without any power optimization and equal number of core allocation is shown. It can be observed that for a fixed value of d_2/d_1, different values of c_2/c_1 result in different transaction time differences. The larger the transaction time difference, the more the under utilized resources.

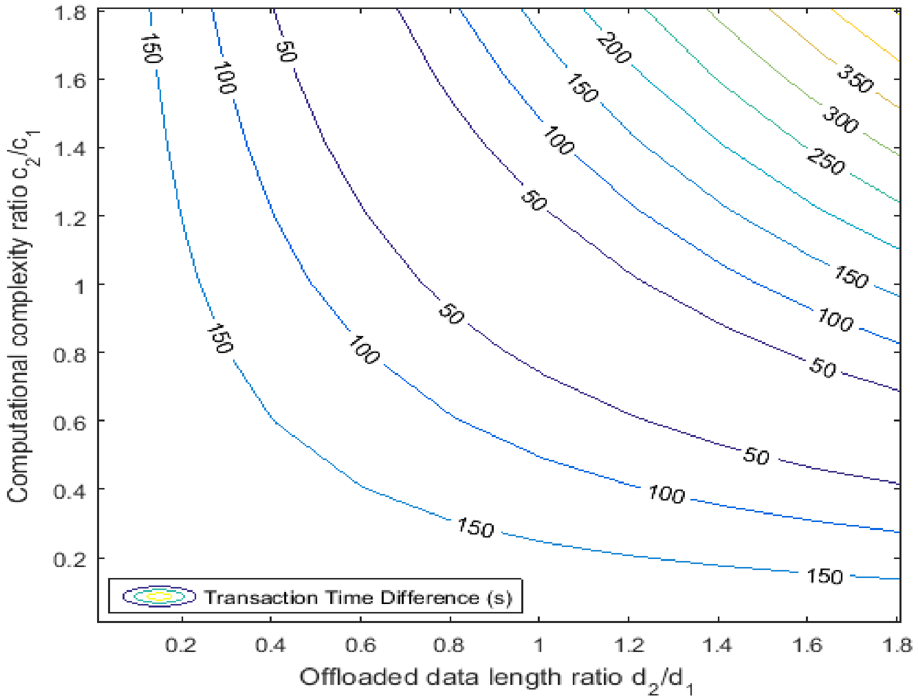

Figure 4.2: Impact of Computational Complexity ratio (c_2/c_1) and Offloaded Data Length ratio (d_2/d_1) on Transaction Time Difference with Power Optimization for Equal Core Allocation [*Approach A*]

Fig. 4.2 illustrates the transaction time difference for Approach A. It can be observed that the time difference is overall lesser than the previous case as shown in Fig 4.1. For the same ratios of d_2/d_1 and c_2/c_1, the transaction time difference is reduced by optimizing only the power allocations. In general, the transaction time difference has reduced from thousand of seconds to hundred of seconds.

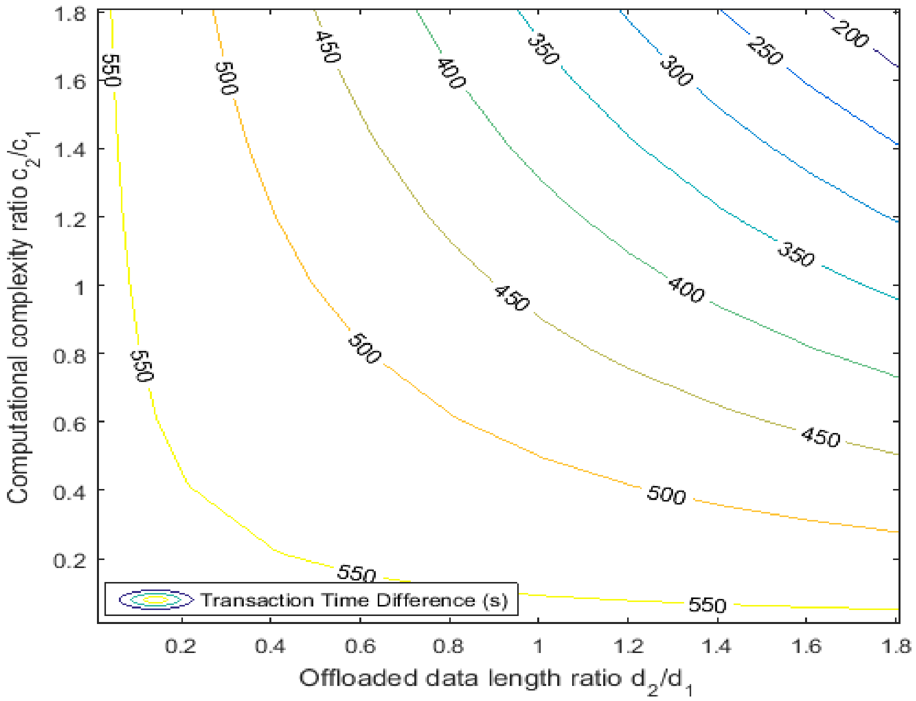

Figure 4.3: Impact of Computational Complexity ratio (c_2/c_1) and Offloaded Data Length ratio (d_2/d_1) on Transaction Time Difference with Power Optimization for Random Core Allocation [*Approach B*]

The transaction time difference for Approach B is shown in Fig. 4.3. This difference is lesser than the transaction time difference shown in Fig. 4.1, i.e., without any optimization. However, this difference is comparable with Approach A, as the only difference is in the core allocation. The transaction time difference for Approach C is shown in Fig. 4.4, where both the power and cores are optimized.

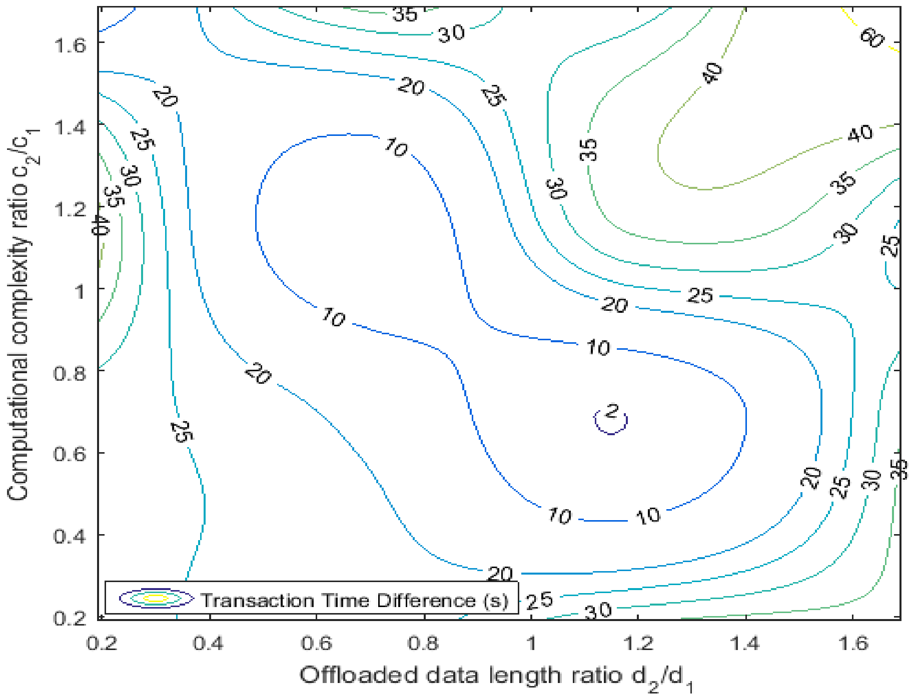

Figure 4.4: Impact of Computational Complexity ratio (c_2/c_1) and Offloaded Data Length ratio (d_2/d_1) on Transaction Time Difference for Power Optimization with Optimal Core Allocation [*Approach C*]

It is evident from the graphs that the transaction time difference is minimum for the proposed scheme, i.e., Approach C. It is also clear that the paired users have optimal values of d_2/d_1 and c_2/c_1 for which the transaction time difference is minimum. The optimal values are different for different user pairs, depending upon their relative channel conditions. For instance in Fig. 4.4, when $d_2/d_1 = 1.2$, the transaction time difference is 2 seconds for c_2/c_1 of 0.7. As d_2/d_1 is increased to 1.7, the transaction time gap jumps to 35 seconds for the same ratio of c_2/c_1. Similarly, when d_2/d_1 is decreased to 0.7, the transaction time difference increases to 20 seconds. Hence the proposed scheme provides least transaction time difference for a given ratio of d_2 and d_1 by optimizing the cores and well as the powers.

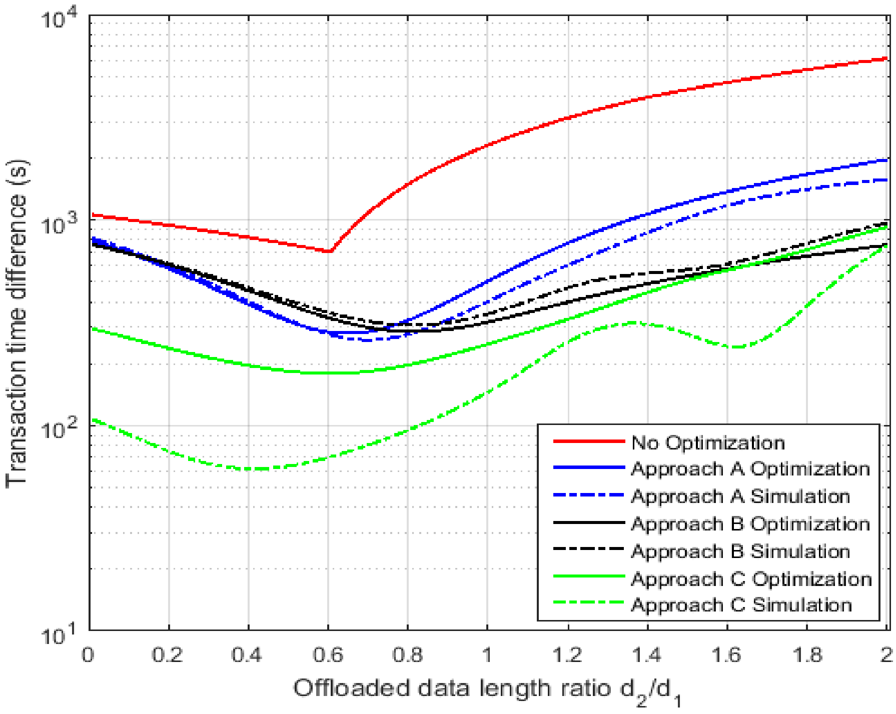

Figure 4.5: Comparison of Simulation and Optimization Results

To validate the proposed solution, the approaches A, B and C are also solved heuristically by searching over the whole solution space labelled as "Simulation" and compared with the results obtained for the approaches A, B and C using successive convex optimization method labelled as "Optimization" in Fig. 4.5. The maximum number of iterations for the Optimization results for the Approaches A, B and C is set to 100. In Fig. 4.5 when d_2/d_1 is 0.4, the transaction time difference without optimization is 822 seconds, approximately 402 seconds for both heuristic and optimized solutions of Approach A and 467 seconds for Approach B. The transaction time difference for Approach C goes to 196.1 and 61.58 seconds for optimized and heuristic solutions, respectively.

4.3 System Effective Throughput

We now illustrate the effect of reducing the transaction time on the *effective throughput* of the system. The effective throughput of the system is given by

$$\Phi_{eff} = \frac{\sum_{i=1}^{2} R_i}{\max\left(T_1, T_2\right)}, \tag{4.1}$$

where the numerator is the sum of achieved data rates by the paired users while the denominator is the maximum of transaction times of the paired users. As both the users are paired, therefore, the resources allocated to them are free only when both of them complete their transactions, hence the denominator is characterized by the max(.) operator. A decrease in effective transaction time increases the system's effective throughput.

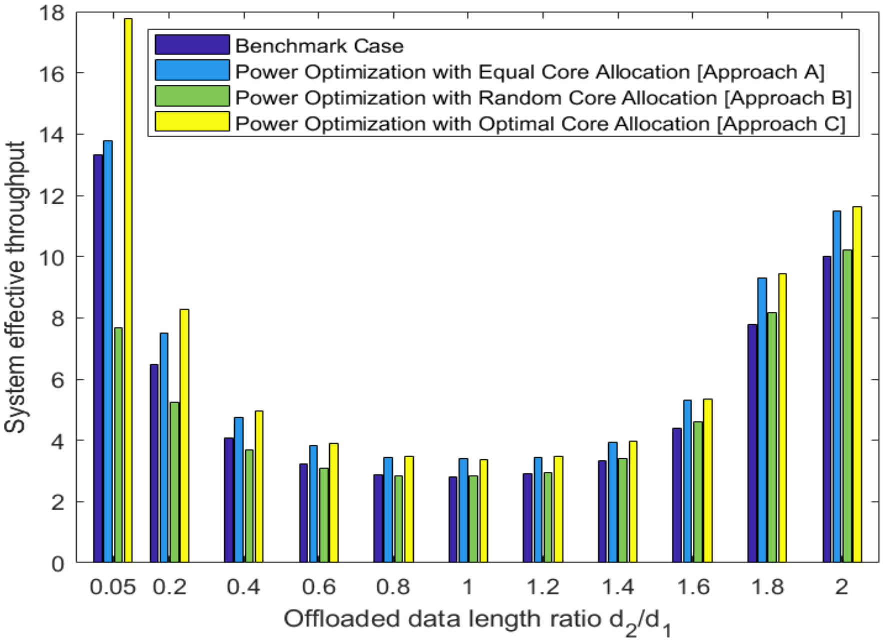

Figure 4.6: Effective System Throughput

It is clear from Fig. 4.6 that for a fixed value of c_2/c_1 and a range of d_2/d_1, the system effective throughput for proposed Approach C is greater than the other approaches. It is also evident that larger is the offloaded data disparity, the larger is the difference between the system's effective throughput for the compared schemes. The reason behind this trend is the optimal core allocation. When the offloaded data is same in characteristic (i.e., amount and complexity is same), the cores allocation for the schemes are same (i.e., equal number of cores for no optimization, Approach A and Approach C) and the difference in the throughput appears only because of the power allocation. However, as the offloaded data disparity increases, the proposed scheme outperforms others. The average increase in the system effective throughput is 19% for the case shown in Fig. 4.6.

Chapter 5

Conclusion and Future Work

In this dissertation, we made the reader to get insight of two evolving technologies, i.e., NOMA and MEC. We identified the potential problem of wasteful resources both in term of cores and frequency for NOMA-MEC. A novel approach is proposed to conserve the resources and increase the system effective throughput. The proposed scheme is evaluated through comparative analysis with other schemes via numerical simulations. The impact of offloaded data amounts and complexities ratios of paired users on transaction time difference is also discussed. The simulation results have shown that the proposed scheme contributes towards the improvement of system performance.

5.1 Future Work

The proposed approach can be extended to multiple users in a single cluster. As in this study we have considered a NOMA cluster with two users only. A data aware NOMA clustering scheme can be used where the users are paired considering both the power disparity as well as their data offloading requirements.The proposed scheme can also be investigated for energy efficiency.

Bibliography

[1] J. Zeng, B. Li, X. Su, L. Rong, and R. Xing, “Pattern division multiple access (PDMA) for cellular future radio access,” in *International Conference on Wireless Communications & Signal Processing (WCSP)*, 2015, pp. 1–5.

[2] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, “A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017.

[3] S. Qureshi, S. A. Hassan, and D. N. K. Jayakody, “Divide-and-allocate: An uplink successive bandwidth division NOMA system,” *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 1, p. e3216, 2018.

[4] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, “Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[5] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, “Communicating while computing: Distributed mobile cloud computing over 5g heterogeneous networks,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 45–55, 2014.

[6] Y. Wang, M. Sheng, X. Wang, L. Wang, and J. Li, “Mobile-edge computing: Partial computation offloading using dynamic voltage scaling,” *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4268–4282, 2016.

[7] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, “Mobile edge computinga key technology towards 5g,” *ETSI white paper*, vol. 11, no. 11, pp. 1–16, 2015.

[8] M. S. Omar, S. A. Hassan, H. Pervaiz, Q. Ni, L. Musavian, S. Mumtaz, and O. A. Dobre, “Multiobjective optimization in 5g hybrid networks,” *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1588–1597, 2017.

[9] R. I. Ansari, C. Chrysostomou, S. A. Hassan, M. Guizani, S. Mumtaz, J. Rodriguez, and J. J. Rodrigues, “5g d2d networks: Techniques, challenges, and future prospects,” *IEEE Systems Journal*, vol. 12, no. 4, pp. 3970–3984, 2017.

[10] U. Saleem, S. Jangsher, H. K. Qureshi, and S. A. Hassan, “Joint subcarrier and power allocation in the energy-harvesting-aided d2d communication,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2608–2617, 2018.

[11] S. A. R. Naqvi, H. Pervaiz, S. A. Hassan, L. Musavian, Q. Ni, M. A. Imran, X. Ge, and R. Tafazolli, "Energy-aware radio resource management in d2d-enabled multi-tier hetnets," *IEEE Access*, vol. 6, pp. 16 610–16 622, 2018.

[12] R. I. Ansari, S. A. Hassan, and C. Chrysostomou, "A swipt-based device-to-device cooperative network," in *2017 24th International Conference on Telecommunications (ICT)*. IEEE, 2017, pp. 1–5.

[13] R. I. Ansari, S. A. Hassan, S. Ali, C. Chrysostomou, and M. Lestas, "On the outage analysis of a d2d network with uniform node distribution in a circular region," *Physical Communication*, vol. 25, pp. 277–283, 2017.

[14] A. Riaz, S. Saleem, and S. A. Hassan, "Energy efficient neighbor discovery for mmwave d2d networks using polya's necklaces," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.

[15] S. A. R. Naqvi, S. A. Hassan, H. Pervaiz, Q. Ni, and L. Musavian, "Self-adaptive power control mechanism in d2d enabled hybrid cellular network with mmwave small cells: An optimization approach," in *2016 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2016, pp. 1–6.

[16] R. I. Ansari, S. A. Hassan, and C. Chrysostomou, "Ranc: Relay-aided network-coded d2d network," in *2015 10th International Conference on Information, Communications and Signal Processing (ICICS)*. IEEE, 2015, pp. 1–5.

[17] ——, “Energy efficient relay selection in multi-hop d2d networks,” in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2016, pp. 620–625.

[18] H. Munir, S. A. Hassan, H. Pervaiz, Q. Ni, and L. Musavian, “Energy efficient resource allocation in 5g hybrid heterogeneous networks: A game theoretic approach,” in *2016 IEEE 84th vehicular technology conference (VTC-Fall)*. IEEE, 2016, pp. 1–5.

[19] A. Mudassir, S. A. Hassan, H. Pervaiz, S. Akhtar, H. Kamel, and R. Tafazolli, “Game theoretic efficient radio resource allocation in 5g resilient networks: A data driven approach,” *Transactions on Emerging Telecommunications Technologies*, p. e3582, 2019.

[20] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, M. Elkashlan, and J. Alonso-Zarate, “Noma-based d2d communications: Towards 5g,” in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, pp. 1–6.

[21] S. Singh, Y.-C. Chiu, Y.-H. Tsai, and J.-S. Yang, “Mobile edge fog computing in 5g era: Architecture and implementation,” in *2016 International Computer Symposium (ICS)*. IEEE, 2016, pp. 731–735.

[22] R. I. Ansari, H. Pervaiz, S. A. Hassan, C. Chrysostomou, M. A. Imran, S. Mumtaz, and R. Tafazolli, “A new dimension to spectrum management in iot empowered 5g networks,” *IEEE Network*, 2019.

[23] M. S. Omar, M. A. Anjum, S. A. Hassan, H. Pervaiz, and Q. Niv, “Performance analysis of hybrid 5g cellular networks exploiting mmwave

capabilities in suburban areas," in *2016 IEEE International Conference on Communications (ICC).* IEEE, 2016, pp. 1–6.

[24] S. Islam, M. Zeng, and O. A. Dobre, "Noma in 5g systems: Exciting possibilities for enhancing spectral efficiency," *arXiv preprint arXiv:1706.08215*, 2017.

[25] V. W. Wong, *Key technologies for 5G wireless systems.* Cambridge university press, 2017.

[26] H. K. Bizaki, *Towards 5G Wireless Networks: A Physical Layer Perspective.* BoD–Books on Demand, 2016.

[27] W. Xiang, K. Zheng, and X. S. Shen, *5G mobile communications.* Springer, 2016.

[28] D. Sabella, A. Vaillant, P. Kuure, U. Rauschenbach, and F. Giust, "Mobile-edge computing architecture: The role of mec in the internet of things," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 84–91, 2016.

[29] M. T. Beck, M. Werner, S. Feld, and S. Schimper, "Mobile edge computing: A taxonomy," in *Proc. of the Sixth International Conference on Advances in Future Internet.* Citeseer, 2014, pp. 48–55.

[30] R. Razavi, M. Dianati, and M. A. Imran, "Non-orthogonal multiple access (noma) for future radio access," in *5G Mobile Communications.* Springer, 2017, pp. 135–163.

[31] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (noma) for cellular future radio access," in *2013 IEEE 77th vehicular technology conference (VTC Spring)*. IEEE, 2013, pp. 1–5.

[32] S. Qureshi, S. A. Hassan, and D. N. K. Jayakody, "Successive bandwidth division noma systems: Uplink power allocation with proportional fairness," in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2017, pp. 998–1003.

[33] S. Qureshi and S. A. Hassan, "Mimo uplink noma with successive bandwidth division," in *2016 IEEE Wireless Communications and Networking Conference*. IEEE, 2016, pp. 1–6.

[34] S. A. R. Naqvi and S. A. Hassan, "Combining noma and mmwave technology for cellular communication," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2016, pp. 1–5.

[35] S. Zeb, Q. Abbas, S. A. Hassan, A. Mahmood, R. Mumtaz, S. Zaidi, S. A. R. Zaidi, and M. Gidlund, "Noma enhanced backscatter communication for green iot networks," in *16th International Symposium on Wireless Communication Systems (ISWCS)*, 2019.

[36] M. N. Jamal, S. A. Hassan, and D. N. K. Jayakody, "A new approach to cooperative noma using distributed space time block coding," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017, pp. 1–5.

[37] M. N. Jamal, S. A. Hassan, D. N. K. Jayakody, and J. J. Rodrigues, "Efficient nonorthogonal multiple access: Cooperative use of distributed space-time block coding," *IEEE Vehicular Technology Magazine*, vol. 13, no. 4, pp. 70–77, 2018.

[38] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5g systems with randomly deployed users," *IEEE signal processing letters*, vol. 21, no. 12, pp. 1501–1505, 2014.

[39] L. Gupta, R. Jain, and H. A. Chan, "Mobile edge computing–an important ingredient of 5g networks," *IEEE Software Defined Networks Newsletter*, 2016.

[40] B. Blanco, J. O. Fajardo, I. Giannoulakis, E. Kafetzakis, S. Peng, J. Pérez-Romero, I. Trajkovska, P. S. Khodashenas, L. Goratti, M. Paolino *et al.*, "Technology pillars in the architecture of future 5g mobile networks: Nfv, mec and sdn," *Computer Standards & Interfaces*, vol. 54, pp. 216–228, 2017.

[41] K. Zhang, Y. Mao, S. Leng, Q. Zhao, L. Li, X. Peng, L. Pan, S. Maharjan, and Y. Zhang, "Energy-efficient offloading for mobile edge computing in 5g heterogeneous networks," *IEEE access*, vol. 4, pp. 5896–5907, 2016.

[42] S. Nunna, A. Kousaridas, M. Ibrahim, M. Dillinger, C. Thuemmler, H. Feussner, and A. Schneider, "Enabling real-time context-aware collaboration through 5g and mobile edge computing," in *2015 12th In-*

ternational Conference on Information Technology-New Generations. IEEE, 2015, pp. 601–605.

[43] Q.-V. Pham, F. Fang, V. N. Ha, M. Le, Z. Ding, L. B. Le, and W.-J. Hwang, "A survey of multi-access edge computing in 5g and beyond: Fundamentals, technology integration, and state-of-the-art," *arXiv preprint arXiv:1906.08452*, 2019.

[44] Z. Ding, D. W. K. Ng, R. Schober, and H. V. Poor, "Delay minimization for NOMA-MEC offloading," *IEEE Signal Processing Letters*, vol. 25, no. 12, pp. 1875–1879, 2018.

[45] A. Kiani and N. Ansari, "Edge computing aware noma for 5g networks," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1299–1306, 2018.

[46] Z. Yang, J. Hou, and M. Shikh-Bahaei, "Energy efficient resource allocation for mobile-edge computation networks with noma," in *2018 IEEE Globecom Workshops (GC Wkshps).* IEEE, 2018, pp. 1–7.

[47] Y. Pan, M. Chen, Z. Yang, N. Huang, and M. Shikh-Bahaei, "Energy-efficient noma-based mobile edge computing offloading," *IEEE Communications Letters*, vol. 23, no. 2, pp. 310–313, 2019.

[48] Z. Ding, J. Xu, O. A. Dobre, and V. Poor, "Joint power and time allocation for noma-mec offloading," *IEEE Transactions on Vehicular Technology*, 2019.

[49] Z. Yang, J. Hou, and M. Shikh-Bahaei, "Resource allocation in full-duplex mobile-edge computing systems with noma and energy harvesting," *arXiv preprint arXiv:1807.11846*, 2018.

[50] Y. Ye, G. Lu, R. Q. Hu, and L. Shi, "On the performance and optimization for mec networks using uplink noma," *arXiv preprint arXiv:1901.04149*, 2019.

[51] Y. Wu, K. Ni, C. Zhang, L. P. Qian, and D. H. Tsang, "Noma-assisted multi-access mobile edge computing: A joint optimization of computation offloading and time allocation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12 244–12 258, 2018.

[52] F. Fang, Y. Xu, Z. Ding, C. Shen, M. Peng, and G. K. Karagiannidis, "Optimal task assignment and power allocation for noma mobile-edge computing networks," *arXiv preprint arXiv:1904.12389*, 2019.