



User Mobility-Aware Device-to-Device and Aerial Communication Networks

By

Mansi Peer

under the guidance of

Dr. Vivek Ashok Bohara and Prof. Anand Srivastava

**Department of Electronics and Communication Engineering,
Indraprastha Institute of Information Technology, Delhi
April, 2022**



User Mobility-Aware Device-to-Device and Aerial Communication Networks

By

Mansi Peer

PhD16117

under the guidance of

Dr. Vivek Ashok Bohara and Prof. Anand Srivastava

submitted

**in partial fulfillment of the requirements for the degree of
Doctor of Philosophy**

to

Indraprastha Institute of Information Technology, Delhi

April, 2022

Certificate

This is to certify that the thesis titled **User Mobility-Aware Device-to-Device and Aerial Communication Networks** being submitted by **Mansi Peer** to the Indraprastha Institute of Information Technology-Delhi, for the award of the degree of **Doctor of Philosophy**, is an original research work carried out by her under our supervision. In my opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.



April 2022

Dr. Vivek Ashok Bohara

Prof. Anand Srivastava



Indraprastha Institute of Information Technology-Delhi
New Delhi-110020

Abstract

Due to the enormous number of cellular users, the traffic load on the core network is increasing rapidly. Hence, it is difficult to maintain the required quality of service in the cellular network. Further, the normal operations of a cellular network is also hampered in wake of a natural or man-made disaster. To support cellular network in such scenarios, recent studies have proposed the use of device-to-device (D2D) and unmanned aerial vehicle (UAV) assisted communication networks. On one hand, in D2D, the information is not routed through the base station (BS) and hence, helps in offloading cellular traffic, lowering the delay, and improving the spectral efficiency. On the other hand, UAVs offer the opportunity to realize UAV-mounted BSs and UAV relays that can dynamically reposition themselves to boost coverage, spectral efficiency, and user quality of experience. Furthermore, D2D and UAVs are also useful in areas with limited cellular connectivity. However, the user devices in D2D and UAV networks are hand-held, which necessitates the knowledge of user mobility behavior for determining and optimizing the network performance. Hence, the goal of this thesis is to optimize the D2D and UAV-assisted communication network performance based on user mobility behavior for different applications such as content caching as well as disaster-affected scenarios.

Studies reveal that the user mobility has certain spatio-temporal correlations. Hence, in our work, we assess the viability of D2D and D2D multicast (D2MD) networks based on real-world mobility traces of users. Specifically, in case of D2MD networks, we propose two novel methods for exploiting spatio-temporal correlations present in the user mobility. We also demonstrate the relationship between spatio-temporal user locations and D2D/D2MD network performance. Based on the above observations, D2D/D2MD networks must carry out mobility aware network optimizations. Hence, we propose a real-world spatio-temporal user behavior aware optimal cache selection in the D2MD network. Content caching allows certain popular multimedia content to be cached at some user devices in the network in order to reduce the content delivery time. The proposed optimization framework minimizes the number of caches while achieving a desired load on the cellular network.

Next, in order to overcome the problem of limited access to cellular network, we investigate multi-hop D2D. Specifically, we utilize multi-hop D2D to establish a disaster-resilient communication network to serve users in the disaster-struck area. We solve the problem of coverage maximization, i.e., maximizing the number of end-to-end connection from user to an active BS, with joint routing, scheduling and source-destination pairing which outperforms shortest path routing based solutions.

Due to UAVs' mobility and ease of deployment, we explore UAVs to facilitate a disaster-resilient communication network. However, the optimal placement of UAVs over a disaster-affected area must be aware of the mobility of the ground users. Hence, we propose a ground user mobility aware multi-UAV placement strategy for disaster-resilient communication network where we maximize the number of ground users covered by the UAVs, while also taking into account the UAV flight time constraint.

The above work in multi-UAV network does not consider user coverage fairness and resource allocation. It also considered disaster-specific user mobility model and users' mobility is assumed to be synchronous. Further, the UAV-BS placement update interval is also a crucial parameter that impacts the network performance; hence, it needs to be optimized. Motivated by the above, we consider a more practical mobility model while jointly optimizing the UAV-BS placement, resource allocation, user association and update interval.

To conclude, we demonstrate the impact of user mobility on the D2D and UAV-assisted communication networks. We propose various network optimization frameworks while accounting for user mobility, which will be helpful in network planning and deployment. Specifically, in our work, we target the content caching, disaster-specific and remote connectivity applications.

To my loving parents . . .

Acknowledgements

I am grateful to the almighty God for giving me the grace I needed to pursue my PhD journey. I would like to express my deepest thanks to my advisors, Dr. Vivek Ashok Bohara and Prof. Anand Srivastava. This thesis would not have been possible without their guidance, support and encouragement. Their enthusiasm and devotion towards research kept me motivated and persistent with my work. Their profound knowledge and experience helped me throughout my journey. I am forever indebted to them for all my learnings about research and technical writing. I would also like to thank my collaborator Dr. Gourab Ghatak. My thoughtful discussions with him further enhanced the rigor of this thesis.

I am thankful to my monitoring committee members, Dr. Sumit Darak and Dr. Mohammad S. Hashmi, for their insightful feedback. It helped me a lot to improve my work. I would also like to extend my heartfelt thanks to Dr. Sanjit Kaul for taking time out for my random queries and providing his expert comments. My sincere thanks to Dr. Guan Yong Liang for offering me an internship at Temasek Lab, Nanyang Technological University (NTU), Singapore. Working with him and his team was an enriching experience for me. His immense knowledge and thought-provoking attitude are commendable. During my internship, I was fortunate enough to work on a Defence Science Organization (DSO) National Laboratories, Singapore project. In just a few months at NTU, I also made a wonderful friend, Dr. Deng Li. I am thankful to her for being so caring and making my stay so comfortable.

I will always cherish the time spent with my seniors and colleagues at Wirocomm Lab. I will always be grateful to Dr. Naveen Gupta, Dr. Parag Aggarwal, and Dr. Neha Jain for their constant guidance. I am thankful to Anand, Hamza, Rizwana, and Dilnashin for creating a learning environment in the lab. I am also grateful to Mitali for being a helpful friend. I will take this opportunity to thank the Admin Staff at IIIT-Delhi, especially Ms. Priti Patel, for resolving all my PhD related queries. I am grateful to the Visvesvaraya PhD Scheme of Ministry of Electronics Information Technology, Government of India, for providing me with the financial support for my work.

This roller coaster of a journey would not have been possible without few amazing people in my life. I cannot thank enough to my parents for their unwavering support. I dedicate this thesis to my very loving and caring parents. I am forever thankful to my confidant and my husband, Rahul, for always having faith in me and making my life a lot easier by his mere presence. My sincere gratitude to my parents-in-law for being very understanding and supportive. I am deeply thankful to my dearest friend, Suman, for being a patient listener to all my woes and always looking out for me.

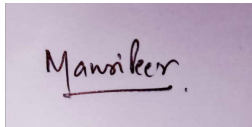
A small rectangular image with a purple background. It contains the handwritten name 'Mansikar' in black ink. The name is written in a cursive style and is underlined with a single horizontal line.

Table of Contents

Abstract	v
Dedication	vii
Acknowledgement	ix
List of Figures	xv
List of Tables	xix
Nomenclature	xxi
1 Introduction	1
1.1 Background and Motivation	1
1.2 Goal and Objectives	4
1.3 Related Work	4
1.4 Contributions	10
1.5 Organization	12
2 Assessing D2D/D2MD Networks Using Real-World User Mobility	15
2.1 Proposed Model for Analysis	16
2.2 Cellular V/s D2D+Cellular Framework	19
2.3 Results and Discussion	22
2.4 Detailed Set-Up and Data Collection for D2MD	25
2.5 Proposed Methods For Joint Spatio-Temporal Behavior Extraction	27
2.6 Interference Model	30
2.7 D2MD Networks And Achievable Rates	31
2.8 Simulation Results and Discussion	36

2.9	Conclusion	42
3	Cache Selection in D2MD Networks	45
3.1	System Model	48
3.2	Problem Formulation and Solution	53
3.3	Result and Discussion	63
3.4	Conclusion	67
4	Enabling Disaster Resilient Communication Using Multi-Hop D2D	69
4.1	Disaster-Resilient Network Model	71
4.2	Problem Formulation	73
4.3	Scheduling Constraint Aware Routing and Pairing (SCARP) Algorithm	76
4.4	Results and Discussion	79
4.5	Conclusion	85
5	UAV-BS Placement in Aerial Communication Networks	87
5.1	Network Model	89
5.2	Problem Formulation	91
5.3	Simulation Results	94
5.4	Network Model and Problem Formulation	98
5.5	Problem Solution	103
5.6	Results and Discussion	104
5.7	Conclusion	108
6	UAV-BS Placement Update with Resource Allocation	111
6.1	Network Model	113
6.2	Problem Formulation and Proposed Solution	114
6.3	Results and Discussion	124
6.4	Conclusion	131
7	Conclusion and Future Work	135
7.1	Future Work	137
8	Publications	139
	Appendix A PDF of Minimum of Channel Gains	141
	Appendix B Computational Complexity	143

Appendix C	Load Values for $K = 10$	145
Appendix D	Proof for Equation (5.16)	149
Appendix E	First-Order Taylor Expansion of Rate	151
References		153

List of Figures

2.1	Layout of IIIT-Delhi Campus.	16
2.2	Floor layout of hostel building (building H).	18
2.3	Floor layout of academic building (building A).	18
2.4	Floor layout of student center building (building S).	19
2.5	Spectral efficiency on weekdays and weekends.	23
2.6	Energy saving on weekdays and weekend.	23
2.7	Impact of n_{fl} on spectral efficiency of D2D and cellular communication.	24
2.8	Illustration of Conventional Transmission Schemes for $U = 7$	25
2.9	TDD frame structure.	26
2.10	Illustration of D2MD-U and D2MD-M networks for an observed occupancy state [3 3 1].	32
2.11	Box plots for occupancy of building H for $n = 1, 10, 14$ and $U = 6$	37
2.12	Box plots for occupancy of building A for $n = 1, 10, 14$ and $U = 6$	37
2.13	Box plots for occupancy of building S for $n = 1, 10, 14$ and $U = 6$	38
2.14	Comparison of NoTF required for D2MD-U and cellular unicast when $s = 18$ m.	39
2.15	Comparison of NoTF required for D2MD-M and cellular multicast when $s = 18$ m.	39
2.16	Comparison of OME and EME.	40
2.17	Data offload ratio of D2MD-U and D2MD-M at $U = 7$	41
2.18	Impact of s on D2MD-U at $U = 7$	41
2.19	Impact of s on D2MD-M.	42
3.1	Illustration of $K = 8$ users belonging to a social group spread across a geographical area consisting of two BSs.	48
3.2	Illustration of dependency of transition matrices on slot boundaries.	51

3.3	Contact graphs for illustrating the user proximity in a real campus set-up where $K = 5, m = 3$	52
3.4	Sum-rate of non-caching users w.r.t. time slot when $\gamma_{cell} = 20$ dB.	64
3.5	Number of Caches w.r.t. number of users to achieve $L_1 = 2$	64
3.6	User load on cellular network w.r.t. time slot for $K = 5, 7, 10$	65
3.7	Minimum user load on cellular network w.r.t. number of caches for $K = 20$	66
4.1	Disaster-resilient network model	71
4.2	Network example for showing the route selection in SCARP and SPR.	79
4.3	Network graph for a specific network realization.	80
4.4	Processing time w.r.t. density of D2D relays	80
4.5	Users covered w.r.t. number of D2D sources	81
4.6	Box plot with $ DS = 10, T_s = 4$ and density of D2D relays (a) $40/\text{km}^2$ (b) $50/\text{km}^2$ and (c) $60/\text{km}^2$	82
4.7	Box plot with $ DS = 10$ at $T_s = 5$ and density of D2D relays = $50/\text{km}^2$	84
4.8	Average energy per active D2D relay node	85
5.1	Network Model.	89
5.2	Zone-wise representation of Disaster Area	90
5.3	Illustration of UAV Placement timeline.	91
5.4	2-D illustration of the ground users along with a UAV's current and new placement.	92
5.5	Performance of proposed UAV Placement Strategy with 10 users and single UAV, $t_{pause} = 60$ seconds.	95
5.6	Performance of proposed UAV Placement Strategy with 20 users and single UAV, $t_{pause} = 60$ seconds.	95
5.7	Performance of proposed multi-UAV placement strategy with 10 users, $t_{pause} = 60$ seconds, 2 UAVs.	96
5.8	Performance of proposed multi-UAV placement strategy with 20 users, $t_{pause} = 60$ seconds, 2 UAVs.	96
5.9	Performance of proposed multi-UAV placement strategy with 10 users, $t_{pause} = 60$ seconds, 3 UAVs.	96
5.10	Performance of proposed multi-UAV Placement Strategy with 10 users, $t_{pause} = 60$ seconds, single UAV and v_{uav} ranges from 14 m/s to 22 m/s.	97
5.11	Performance of proposed multi-UAV Placement Strategy with 10 users, $t_{pause} = 60$ seconds, single UAV and v_{uav} ranges from 14 m/s to 22 m/s.	97
5.12	Network Model	99
5.13	UAV-BS Placement Timeline	99

5.14	Initial Deployment	105
5.15	Total UAV-BS Flight time	105
5.16	Coverage probability of users	106
5.17	Average Number of update instants during $T = 900$ seconds	106
5.18	Average update interval at $T = 900$ seconds	107
5.19	Average number of users covered at $T = 900$ seconds at $U = 20$	107
5.20	Average service time at $T = 900$ seconds	108
6.1	Network Model	113
6.2	Network Optimization Timeline	113
6.3	Flow diagram showing the Sequential Approach.	123
6.4	Sum Rate in Max Sum Rate Vs Max Min Rate Framework	125
6.5	Convergence of sum rate Vs Algorithm 9 iterations	125
6.6	Bandwidth allocation per user in logarithmic scale for max sum rate framework	126
6.7	Total UAV-BS Flight Time for Max Sum Rate framework	126
6.8	Coverage Probability for Max Sum Rate framework at $R_{th} = 190$ Mbps . . .	127
6.9	Minimum user rate in Max Min Rate framework	128
6.10	Minimum user rate Vs Algorithm 9 Iterations	128
6.11	Bandwidth allocation in Max Min Rate framework	129
6.12	Total UAV-BS Flight time for Max Min Rate framework	129
6.13	Coverage Probability for Max Min Rate framework at $R_{th} = 3.6$ Mbps	130
6.14	Comparison of Analytical and Simulated coverage probability at $R_{th} = 3.5$ Mbps	130
6.15	Average Update Interval Vs number of UAV-BSs	132
6.16	Comparison of service time with benchmark approach in Max sum rate framework	132
6.17	Comparison of coverage probability with benchmark approach in Max sum rate framework	132
6.18	Average Update Interval Vs Number of UAV-BSs for $\alpha = \{0.2, 0.5, 0.8\}$. . .	133
6.19	Comparison of service time with benchmark approach in max min rate framework	133
6.20	Comparison of coverage probability with benchmark approach in max min rate framework	133

List of Tables

2.1	Probability for Weekdays and Weekends	17
2.2	Probability of D2D communication on Weekdays and Weekends	20
2.3	p-Values	28
2.4	r^2 values and coefficients of fitted curve	30
2.5	Simulation Parameters	36
2.6	Probability of C in each building	41
3.1	Feasibility sets for each building	50
3.2	Table for $ C_1 =1$ and $K=6$	57
3.3	Table for $ C_1 =2$ and $K=6$	57
3.4	Table for $ C_1 =1$ and $K=8$	58
3.5	Table for $ C_1 =2$ and $K=8$	58
3.6	Table for $ C_1 =3$ and $K=8$	59
3.7	Results for $K=20$	60
3.8	Table for $ C_2 =3$ and $K=6$ in case of re-optimization at 10^{th} slot	63
3.9	Table showing impact of γ_{cell}	64
5.1	User Placement in Disaster-Affected Area	94
5.2	Simulation Parameters	94
5.3	Simulation Parameters	104
5.4	Standard Deviation	107
6.1	Simulation Parameters	125
C.1	Table for $ C_1 =1$ and $K=10$	145
C.2	Table for $ C_1 =2$ and $K=10$	145
C.3	Table for $ C_1 =3$ and $K=10$	146
C.4	Table for $ C_1 =4$ and $K=10$	147

Nomenclature

Abbreviations

2-D	Two Dimensional
3-D	Three Dimensional
3GPP	Third Generation Partnership Project
BS	Base Station
D2D	Device-to-Device
D2I	Device-to-Infrastructure
D2MD	D2D Multicasting
DL	Downlink
EFR	Emergency First Responder
EME	Expected Mobility Exploitation
I2D	Infrastructure-to-Device
LoS	Line-of-sight
LP	Linear Programming
LTE	Long Term Evolution
MIQCP	Mixed Integer Quadratically Constrained Problem
NLoS	Non-Line-of-Sight

OME	Observed Mobility Exploitation
PDF	Probability Density Function
PL	Path-Loss
QoS	Quality of Service
RB	Resource Block
ROI	Region-of-Interest
SINR	Signal-to-Interference and Noise Ratio
SNR	Signal-to-Noise Ratio
SPPP	Spatial Poisson Point Process
SPR	Shortest Path Routing
TDD	Time Division Duplexing
UAV	Unmanned Aerial Vehicle
UE	User Equipment
UL	Uplink

Introduction

In this chapter, we discuss the motivation, objectives, and contributions of this thesis.

1.1 Background and Motivation

Seamless wireless connectivity has led to unprecedented growth in mobile devices. Consequently, the mobile data traffic on the core network is increasing rapidly. It has been reported that between 2016 to 2022, the total mobile data traffic will experience a compound annual growth rate (CAGR) of nearly 45% [1]. The exponential increase in data traffic has led to considerable load on existing cellular networks. Further, it is also challenging to provide wireless connectivity in remote areas as well as during emergency scenarios. For instance, approximately 29,000 BSs were damaged during the Japan tsunami in 2011 [2]. However, in the aftermath of disasters, man-made or natural, it is imperative to carry out search and rescue to save as many lives as possible. Presently, in a disaster scenario, the search and rescue teams depend on the legacy public safety communication networks like terrestrial trunked radio (TETRA) in Europe and the association of public-safety communications officials (APCO 25) in US that support only voice services [3]. It has been suggested that the public safety communication networks need to be upgraded with emerging technologies that can support both voice and multimedia broadband services. The third generation partnership project (3GPP) Releases 13-15 recommend the adoption of long term evolution (LTE) technology for the public safety communication [4].

To cater to the problem of growing mobile traffic, device-to-device (D2D) communication is incorporated as a part of LTE Release 12-15 to enable the devices in proximity to communicate directly with each other rather than routing the information through the core network [5]. It has been reported in [6] that there exists an 8:2 or 9:1 traffic ratio between downlink (DL) and uplink (UL). The overburden on the DL resources can be lowered by allowing the proximate

cellular users to communicate using UL resources in D2D mode. Along with the offloading capability, D2D also helps in reducing the latency and increasing the throughput between the devices, thanks to its shorter communication range. Although mobility is integral to the analysis of D2D networks, the existing works on D2D networks are limited to the simulated mobility behavior of users and fail to incorporate the real-world spatio-temporal behavior of users. Recently, there has been a considerable amount of research on D2D multicasting (D2MD) networks that leverage the multicast nature of cellular communication. Similar to the work on D2D networks, the prior work on D2MD networks also failed to consider user mobility. However, it is necessary to take into account user mobility while evaluating such networks. For instance, work in [7] also highlights the necessity to have more accurate D2D mobility behavior. Although randomly generated user locations may provide a general overview of the impact of mobility on the D2MD performance, the spatio-temporal correlations in the user mobility are critical to analyzing the viability of D2D cluster/group formation in a real-world scenario.

In addition to the above, recently, it has been shown in [8] that the benefits of D2D can be complemented by incorporating content caching feature. Content caching at the network edge refers to temporary storage of popular multimedia content at user devices to alleviate duplication of content requests at the core network [9–11]. In order to reduce the burden on the core network, promising data offloading solutions such as D2D communication and content caching at the network edge have been recently proposed. It has been shown in [8] that the benefits of D2D can be complemented by incorporating content caching feature. However, caching at the user devices on the network edge requires the information of the mobility behavior of the users [12, 13]. The recent literature on mobility-aware caching strategy for D2D at user devices is based on the inter-contact time model [12, 14] for a pair of users. However, they do not take into account the joint spatio-temporal behavior of the users and fail to analyze the real-world interactions among multiple users for dynamic D2MD networks. Further, the mobility-aware caching strategies that are discussed in [12, 14, 15] are primarily content placement strategies. They consider a large library of files and utilize file segmentation to store segments of files in a distributed manner. However, in scenarios where the library of files to be cached is small, for example, within a social group of users (explained later in detail), cache selection can be utilized where the complete file is stored on the selected caches. Optimal cache selection can alleviate the burden of content caching on the cellular network. Moreover, unlike the file segmentation approach, caching the complete file at each cache does not require the proximity of all caches to the requesting users.

In addition to traffic overload, limited access to a cellular network in a remote area or during a disaster is a key challenge. Releases 13-15 of 3GPP, also propose a D2D communication framework wherein a device that is out of cellular network coverage can be assisted by another

nearby device (as a relay) to communicate with the cellular network [16]. Moreover, recent studies have shown that multi-hop D2D is a viable technique for alleviating the damage caused to the cellular network by reinforcing the cellular connectivity [17]. Multi-hop D2D ensures the extension of network coverage, improvement in battery life, and reducing end-to-end delay, which are all crucial parameters in a disaster-affected scenario. The performance of multi-hop networks generally relies on the routing decisions as routing is responsible for establishing inter-network communications. Therefore, routing in multi-hop D2D networks has received considerable attention from researchers everywhere. Multi-hop D2D networks have broadly two categories of routing: multi-hop D2D routing and multi-hop device-to-infrastructure/infrastructure-to-device (D2I/I2D) routing [5]. In multi-hop D2D routing, the source and destination are two user devices which route the information through intermediate D2D relays. In multi-hop D2I/I2D routing, a route is selected between a user device and a base station (BS). The route consists of intermediate D2D relays, also known as network relays. Further, in multi-hop D2I/I2D routing BS acts as the source and destination in downlink and uplink, respectively.

Multi-hop D2D routing has been shown to be useful for offloading traffic from the cellular network and also enhancing spectral efficiency. Unlike the conventional cellular scenario, a user in the disaster-struck area (or dead spot) would like to immediately communicate with close relatives, remote emergency control room, etc. This communication can only happen by first routing the information to a BS. Since most of the BSs in the dead spot are already destroyed or non-active, it is desirable to use the multi-hop D2I/I2D routing to connect to a far-off BS which is still active. The prior works on multi-hop D2I/I2D routing for public safety and disaster scenarios result in inefficient use of wireless radio resources in the presence of interference or contention among different wireless links [18]. This is due to the fact that they have overlooked the half-duplex nature of the D2D relays, i.e., a D2D relay can either transmit or receive so that the transmitted and received data flows do not interfere while making the routing decisions. Further, it is important to restrict the data flows that can be transmitted or received simultaneously by a D2D relay to limit the contention among different flows/users [19]. Hence, it is imperative to jointly address the problem of routing and scheduling in the multi-hop D2D network. In addition, specifically for a disaster scenario, there is also a need to provide coverage to a maximum number of users present in the dead spot via active BS.

Further, due to the flexibility and maneuverability of unmanned aerial vehicles (UAVs) [20], UAV-assisted networks can also support emergency communication during disasters or provide on-demand wireless connectivity in remote areas. In addition, UAVs are also capable of offering low-latency communication services. With the ongoing third generation partnership project (3GPP) standardization for UAV-assisted communication networks, researchers are actively working on optimizing the performance of such networks [21]. In UAV-assisted

communication networks, UAVs can either be deployed as aerial relays or as aerial base stations (BSs). These aerial BSs are often referred to as UAV-BSs. Air-to-ground and air-to-air channel modeling, resource allocation, interference management, UAV-BS placement, or UAV-BS trajectory design are some of the critical challenges in UAV-assisted communication networks [22]. Recently, research in UAV-BS placement optimization and resource allocation in UAV networks has also gained a lot of momentum [23–27]. Moreover, the recurring placement optimization of UAV-BSs becomes challenging when the ground users are mobile. The ground user mobility can profoundly impact the network performance if the placement of UAVs is not optimized with the change in the ground user locations.

1.2 Goal and Objectives

This thesis deals with device-to-device and aerial communication networks that have been proposed as a solution to overcome the challenges faced in conventional cellular networks. Broadly, the goal of this thesis is to characterize the user mobility for determining and optimizing the D2D and UAV-assisted network performance for different applications like content caching and disaster-resilient communication. The main objectives of this thesis are as follows:

- To identify the spatio-temporal correlations in real-world mobility of users to assess the viability of D2D/D2MD networks.
- To leverage content caching in D2MD networks while accounting for the user mobility.
- To explore multi-hop D2D for establishing time-bounded disaster-resilient communication.
- To optimize placement of UAV-BSs in a disaster-struck area while accounting for disaster-specific user mobility.
- To jointly optimize UAV-BS placement, resource allocation, and update interval for a multi-UAV network while accounting for user mobility.

1.3 Related Work

In this section, we discuss the existing literature in four broad areas as follows:

1.3.1 D2D/D2MD networks

A number of approaches have been proposed in the literature to analyze a D2D communication underlying a cellular network. For instance, in [28, 29] an outage analysis of these networks is

performed, whereas in [30] spectral efficiency is evaluated. In all of the above, it is assumed that the users follow a spatial Poisson Point Process (PPP) with distance-dependent path loss. Although distance-dependent path loss models can characterize an outdoor propagation environment, however, they are fairly inaccurate when considering the indoor propagation environment [31]. In [32, 33] indoor D2D scenarios are investigated with the aim of channel measurement and modeling. The authors perform measurements in different experimental set-ups over a wide range of frequencies. They derived the channel parameters that better characterize an indoor D2D channel. Specifically, [32] compared the measurement results with the existing standard indoor channel models, and [33] demonstrated the susceptibility of D2D communication channel to shadowing caused by the human body. Nevertheless, the above-cited works lack in providing an insight into the performance of D2D underlying a cellular network. The work in [34] proposes an indoor D2D scenario where D2D links share uplink resources with a cellular network. However, the D2D pairs are generated and restricted to be present within a single room.

Further, there has been a considerable amount of research on D2MD networks that leverage the multicast nature of communication. For instance, in [35], authors presented a cluster-based multicast transmission method for D2D communication with an objective to decrease the data distribution time. They derived the outage capacity of a single cluster and proposed a game-theoretic strategy for an otherwise non-convex D2MD grouping/clustering problem. Authors in [36] proposed an efficient delivery of multicast services in LTE-Advanced (LTE-A) networks by exploiting D2D links between user equipments within a multicast group. Further, the literature provided insights into the mechanisms for D2MD group formation, which are influenced by the physical factor, i.e., the proximity of the users and the social factor that quantifies the social bonding of the users [37–39]. Authors in [40] proposed a socially aware D2D communication framework to develop a spectrum and energy-efficient cellular network. They presented a cluster formation scheme for segregating a set of users into different D2D clusters. Further, they investigated half-duplex and full-duplex schemes for coordinated channel sharing between the cellular links and the D2D links. Authors in [41] presented a content sharing D2MD framework for cellular networks based on physical and social factors of D2MD clustering. Further, they also optimized the power and channel allocation among the D2MD clusters. Although the prior work has studied various aspects of the D2MD networks, they failed to consider user mobility. For instance, work in [7] also highlights the necessity to have more accurate D2D mobility behavior. Authors in [42] studied the effect of mobility on various multicast metrics. However, user locations were randomly generated.

1.3.2 Content Caching in D2D/D2MD Networks

Caching at the user devices on the network edge requires the information of the mobility behavior of the users, and the mobility behavior is characterized by the spatio-temporal behavior of the user [12, 13]. The recent literature on mobility-aware caching strategy for D2D network is based on the inter-contact time model [12, 14] for a pair of users. Specifically, authors in [12] solved a caching placement problem that maximizes the data offloading ratio through a greedy algorithm. They extended this work in [14] by modeling the pairwise contact pattern as an alternating renewal process and provided an improved cache placement strategy. An optimal caching policy based on user preferences was proposed in [43]. Similar to [12, 14], in [43] the problem of maximizing offloading probability is solved using a greedy algorithm. In [15], a caching strategy to minimize the cost is presented where caching took place at both SBSs and user devices. Further, a closed-form expression for the average system cost is also derived. The work in [41] dealt with resource allocation for D2D and D2MD content sharing and utilized social and physical domain knowledge for D2D cluster formation for a given time instant. It is shown in [41] that exploiting D2MD at the user devices resulted in less consumption of resources at the content caching device. In [44], the authors proposed a novel approach to minimize the downloading latency and maximize the social welfare simultaneously for a socially aware D2D network. For achieving the above objectives, they efficiently selected the important users and matched the contents with users in a joint manner.

However, as evident from above, the existing works on caching strategies for D2D underlying cellular networks do not take into account the joint spatio-temporal behavior of the users and fail to analyze the real-world interactions among multiple users for dynamic D2MD networks. Further, the mobility-aware caching strategies that are discussed in [12, 14, 15] are primarily content placement strategies. They consider a large library of files and utilize file segmentation to store segments of files in a distributed manner. However, in scenarios where the library of files to be cached is small, for example, within a social group of users (explained in detail in Chapter 3), cache selection can be utilized where the complete file is stored on the selected caches. Optimal cache selection can alleviate the burden of content caching on the cellular network. Moreover, unlike the file segmentation approach, caching the complete file at each cache does not require the proximity of all caches to the requesting users. In addition to the above, the authors in [12, 14] assumed the transmission rates to be the same for all D2D pairs. This is because the inter-contact time models used in these works only looked at the contact time duration and neglected the spatial locations and spatio-temporal correlations in user mobility. However, spatial preferences and temporal dependencies are needed to evaluate the holistic performance of content caching D2MD underlaid cellular networks.

1.3.3 Multi-hop D2D Networks

The performance of multi-hop networks generally relies on the routing decisions as routing is responsible for establishing inter-network communications. Therefore, routing in multi-hop D2D networks has received considerable attention from researchers everywhere. Multi-hop D2D networks have broadly two categories of routing: multi-hop D2D routing and multi-hop device-to-infrastructure/infrastructure-to-device (D2I/I2D) routing [5]. In multi-hop D2D routing, the source and destination are two user devices that route the information through intermediate D2D relays. In multi-hop D2I/I2D routing, a route is selected between a user device and a BS. Further, in multi-hop D2I/I2D routing BS acts as the source and destination in downlink and uplink, respectively.

Multi-hop D2D routing has been shown to be useful for offloading traffic from the cellular network and also enhancing spectral efficiency. For instance, authors in [45] proposed an interference-aware routing algorithm with an objective to minimize the hop counts for a multi-hop D2D network. Similar to the work in [45], an emergency route selection framework during an urban terrorist attack is proposed for multi-hop D2D networks co-existing with a cellular network [46]. The closed-form expressions for the number of hops and outage probability are presented in [47] for a tractable theoretical framework of multi-hop D2D underlying cellular network. Moreover, the multi-hop D2D framework in [47] used the shortest-path routing (SPR) algorithm.

Since most of the BSs in the disaster-struck area are already destroyed or non-active, it is desirable to use the multi-hop D2I/I2D routing to connect to a far-off BS which is still active. For instance, in [48], authors studied the benefits of multi-hop D2D in extending the coverage area of BS in public safety scenarios. It is shown that the average energy and spectral efficiencies due to multi-hop D2D are enhanced when the number of hops is increased. Like [47], they utilized the SPR algorithm. Authors in [49] proposed a routing scheme utilizing the ant colony optimization to maximize the end-to-end throughput for all the data flows originating from the area without cellular network coverage. However, the aforementioned works of multi-hop D2I/I2D routing for public safety and disaster scenarios result in inefficient use of wireless radio resources in the presence of interference or contention among different wireless links[18]. This is due to the fact that they have overlooked the half-duplex nature of the D2D relays, i.e., a D2D relay can either transmit or receive so that the transmitted and received data flows do not interfere while making the routing decisions. Further, it is important to restrict the data flows that can be transmitted or received simultaneously by a D2D relay to limit the contention among different flows/users [19]. Hence, it is imperative to jointly address the problem of routing and scheduling in the multi-hop D2D network. In addition, specifically for a disaster scenario, there is also a need to provide coverage to a maximum number of users present in the dead spot via active BS.

Joint routing and scheduling algorithms have been widely discussed in the literature in the context of ad-hoc wireless networks [18, 50–52]. These works generally involve a single destination or fixed source-destination pairs. Further, routing and scheduling decisions are carried out in a distributed manner, and they primarily deal with network lifetime and throughput maximization problems. However, in a disaster scenario, the source-destination pairs, i.e., the information about which user in the dead spot will connect to which BS is not known beforehand. Further, as suggested in [49], a disaster-resilient network will require a centralized approach for routing to reduce energy consumption. To the best of our knowledge, the existing literature has not dealt with the coverage maximization problem using joint source-destination pairing, routing, and scheduling for a multi-hop D2D network in a disaster scenario.

1.3.4 UAV-BS Placement in Aerial Communication Network

With the recent technological advances, it is feasible to deploy UAVs for the roles of flying/aerial base stations (BSs) as well as aerial user equipments (UEs) [20, 53]. Specifically, on using UAVs as flying BSs, the key challenge is the optimal 3-D placement of the UAVs for efficient network performance [54]. The 3-D placement of UAVs has drawn a lot of attention from researchers. For instance, authors in [55] have dealt with maximization of covered users via optimal 3-D placement of a single UAV. In [56], the optimal 3-D deployment of multiple UAVs is investigated to maximize the downlink coverage performance with minimum transmit power. Similarly, in [57], an efficient UAV-BS 3-D placement method is proposed for maximizing the number of covered users while minimizing the required transmit power where the UAV placement in the vertical dimension is decoupled from the horizontal dimension. Further, a geometric disk cover problem has been formulated in [58] to optimally place the UAVs in order to minimize the number of UAVs required. In [59] and [60], 3-D placement UAV-BSs is studied for maximizing the sum logarithmic rate of the users and effectively prolonging the lifetime of the network, respectively. Further, they analyzed the network performance for different user distributions like the Poisson point process and clustered user distribution.

In the existing literature, placement optimization has been studied for both static and mobile UAV-BS [61]. In the case of static deployment, UAV-BSs are placed at a hovering location, and they remain static throughout their mission duration. For instance, [62] has considered a heterogeneous network consisting of both macro BSs and UAV-BSs. The placement of UAV-BSs is optimized to maximize the downlink received signal strength. Further, [63] jointly optimized UAV-BS placement and user association in a heterogeneous network to maximize the spectral efficiency of a hotspot area. The spectral efficiency for both the wireless access as well as the wireless backhaul links is considered. The minimum number of UAV-BSs required

to provide coverage to a set of ground users in the absence of fixed infrastructure has been determined in [58]. However, networks with the static deployment of UAV-BSs cannot adapt to the varying user locations and demands. In the case of mobile UAV-BSs, the placement of UAV-BSs changes over time. For instance, in order to collect data from the sparsely deployed sensor nodes in a wireless sensor network (WSN) framework, UAV-BSs have to move around the region of interest [64]. Similarly, in [65], a single UAV's trajectory is optimized to serve some randomly distributed sensor nodes when the data volume of the sensors is unknown.

In [66] a framework is presented to optimize UAV trajectory, power allocation, and user scheduling in order to maximize the minimum rate of users. It considers a multi-UAV network where each UAV serves the users using a time division multiple access (TDMA) protocol. In [67], a single UAV network is considered, which optimizes the UAV trajectory, user association, bandwidth, and power allocation to maximize the overall network energy efficiency. However, unlike [66], it combines TDMA with frequency division multiple access (FDMA) for serving users. It may be noted that in [66] and [67], the user locations are assumed to be fixed. The optimization is carried out in an offline manner for all time slots at once. Consequently, UAV follows the same trajectory in a periodic manner. Hence, the frameworks proposed in [66] and [67] will not be applicable when users are mobile. In [68], a multi-agent reinforcement learning (RL) framework is proposed to perform distributed intelligent resource management in a UAV-assisted communication network when the UAVs only have individual local information. In order to maintain fairness among the ground users, the authors in [69] proposed a 3-D UAV scheduling with energy replenishment RL framework. With a limited number of UAVs, the work in [69] tries to efficiently perform the 3-D UAV scheduling so that users are served fairly while maintaining reliable communication. In [70], UAV-BSs are deployed to maximize the coverage area subject to the energy, user coverage fairness, and inter-UAV collision avoidance constraints. In [71], a swarm of UAV-BSs overlaid over macrocells is considered. The 3D placement of these UAV-BSs is optimized in order to maximize the total data transmitted by all the UAV-BSs in the limited network lifetime.

The recurring placement optimization of UAV-BSs becomes challenging when the ground users are mobile. For instance, in [72], UAV-BS locations are updated in accordance with the user locations over time in order to maximize the network throughput. Further, the authors in [73] propose a RL framework to efficiently update the placement of UAV-BS in a dynamic heterogeneous network while maintaining the desired quality-of-service (QoS). Authors in [74] proposed a RL framework for 3-D movement design of multiple UAVs in order to maximize the sum mean opinion score of users. Specifically, the area of interest is divided into clusters, and it is assumed that each cluster is served by one UAV. The users are supposed to be confined within the cluster and cannot roam outside the cluster. Hence, [74] fails to consider a more realistic user mobility model. In [75], an echo state network-based prediction algorithm is

used to predict the user positions, and then a multi-agent Q-learning based algorithm is used to design UAV-BS trajectory in advance.

It may be noted that, in most of the prior works, the placement optimization of mobile UAV-BSs is generally carried out at specific time instants. Moreover, the user mobility behavior will play a key role in determining the time separation between two consecutive UAV-BS placement updates (or update interval). However, the prior works fail to characterize the relationship between user mobility and update interval. For example, [72] and [75] assume a fixed update interval whereas [73, 74] lack in quantifying such an update interval.

1.4 Contributions

The major contributions of this thesis can be summarized as follows:

- The viability of D2D communication within the indoor environment is explored. For indoor propagation, the impact of additional factors such as floor and wall attenuation is considered. Instead of randomly generating the user location information, the real-world location information of a user is leveraged. A probability model is developed to analyze the trade-offs of D2D communication vis-a-vis cellular communication. Further, the realistic performance of D2MD networks, by utilizing the joint spatio-temporal mobility behavior of the users in terms of content delivery time and data offloading ratio, is investigated. To obtain the required joint spatio-temporal behavior of users, two novel methods, namely observed mobility exploitation (OME) and expected mobility exploitation (EME) are proposed. OME is applied when past location information of all the users is available, whereas EME is utilized when the location information of a set (i.e., the training set) of users, following an average spatial-temporal behavior similar to the intended users, is available. The relationship between spatio-temporally correlated user mobility and D2MD network performance is demonstrated.
- A user spatio-temporal behavior aware cache selection framework for D2MD networks is proposed. An optimization problem is formulated to minimize the number of caches required to achieve a desired user load on the cellular network. The minimization of caches is helpful in alleviating the caching load on the cellular network. However, frequent optimizations due to the unexpected occurrences of network congestion will result in frequent changes in selected caches. Consequently, the proposed framework does not discard the previously selected caches, which further reduced the caching load. The formulated problem has an exponential search space. Hence, a greedy algorithm for cache selection with lower complexity is proposed. An inhomogeneous discrete-time

Markov chain model based on real-world location information of users is presented to predict the spatio-temporal behavior.

- A joint source-destination pairing, routing, and scheduling framework is proposed for multi-hop D2D communication in a disaster-struck area (or dead spot). The maximization of covered users in the dead spot is formulated as an ILP. Hence, a low complexity scheduling constraint aware routing and pairing (SCARP) algorithm is proposed to perform joint pairing, routing, and scheduling, unlike the existing algorithms.
- A UAV based disaster-resilient communication network is proposed. The 3-D placement of multiple UAVs is optimized in order to maximize the number of ground users covered in a disaster-affected area while taking into account the UAV flight time constraint. In the proposed strategy, a disaster-specific ground user mobility model is considered. The results for single-UAV and multi-UAV scenarios are demonstrated in terms of the number of covered users and average coverage time in a disaster-resilient network. However, in the above work, the update interval is not optimized. Subsequently, joint optimization of the UAV-BS placement and the placement update interval is proposed for a single UAV network. Later, a multi-UAV network is considered where the UAV-BS placement, resource allocation, user association, and update interval are jointly optimized. Further, the analytical expression for user coverage probability in terms of user mobility and update interval are derived.

1.4.1 List of Publications

The following papers are the outcome of the work done towards this thesis:

Journals

1. **M. Peer**, V. A. Bohara, A. Srivastava, G. Ghatak, "User Mobility-Aware UAV-BS Placement Update with Optimal Resource Allocation," submitted to *IEEE Transactions on Communications*.
2. D. N. Anwar, **M. Peer**, K. Lata, A. Srivastava, V. A. Bohara, "3-D Deployment of VLC Enabled UAV Networks with Energy and User Mobility Awareness," in *IEEE Transactions on Green Communications and Networking*, 2022.
3. **M. Peer**, V. A. Bohara and A. Srivastava, "Enabling Disaster-Resilient Communication Using Multi-Hop Device-to-Device Framework," in *Wireless Networks (Springer)*, 2020.

4. **M. Peer**, V. A. Bohara and A. Srivastava, "Cache Selection in Dynamic D2D Multicast Networks Using Inhomogeneous Markov Model," in *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3235-3245, 1 Oct.-Dec. 2020.
5. **M. Peer**, V. A. Bohara and A. Srivastava, "Real-World Spatio-Temporal Behavior Aware D2D Multicast Networks," in *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1675-1686, 1 July-Sept. 2020.

Conferences

1. S. Dohadwalla, **M. Peer**, V. A. Bohara, "Fair Multiple Subchannel Assignment and 3-D UAV-BS Placement in UAV-Enabled Networks," accepted in *IEEE VTC Spring*, 2022.
2. Y. Gupta, **M. Peer**, V. A. Bohara, "Performance Analysis of RF/VLC Enabled UAV Base Station in Heterogeneous Network," accepted in *IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021.
3. **M. Peer**, V. A. Bohara, A. Srivastava and G. Ghatak, "User Mobility-Aware Time Stamp for UAV-BS Placement," 2021 *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2021, pp. 1-6, doi: 10.1109/WCNCW49093.2021.9420032.
4. **M. Peer**, V. A. Bohara and A. Srivastava, "Multi-UAV Placement Strategy for Disaster-Resilient Communication Network," 2020 *IEEE 92nd Vehicular Technology Conference (VTC-Fall)*, 2020, pp. 1-7, doi: 10.1109/VTC2020-Fall49728.2020.9348687.
5. **M. Peer**, V. A. Bohara and A. Srivastava, "Multi-Hop D2D Framework for Disaster-Resilient Communication Network," 2020 *IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 584-589, doi: 10.1109/5GWF49715.2020.9221256.
6. **M. Peer**, V. A. Bohara and A. Srivastava, "On the performance of network-assisted indoor device-to-device communication using location awareness and realistic path loss models," 2017 *IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1-7, doi: 10.1109/PIMRC.2017.8292583.

1.5 Organization

The rest of the thesis is organized as follows.

- Chapter 2 presents our work on the assessment of the viability of D2D and D2MD networks. The chapter is divided into two parts. In part I, real-world mobility based

model for analysis of D2D-enabled cellular and cellular frameworks is proposed. The simulation results demonstrate the diurnal variation in spectral efficiency and energy saving on weekdays and weekend. In part II, for analyzing a D2MD network, the detailed set-up and data collection procedure for obtaining the user mobility is presented. We propose methods for joint spatio-temporal behavior extraction from the collected data. The simulation results exhibit the data offloading performance of D2MD networks.

- Chapter 3 presents a real-world user mobility aware cache selection framework for D2MD networks. We present the system model for a D2MD underlaid cellular network and the details of the inhomogeneous Markov chain. An optimization problem is formulated and an algorithm is proposed. We demonstrate the simulation results with real data traces.
- In Chapter 4, we investigate the applicability of multi-hop D2D to establish a disaster-resilient communication network. An optimization problem to maximize the number of users covered in the dead spot is presented. We explain the working of the proposed SCARP algorithm for solving the formulated problem. We discuss the results obtained on applying the optimal solution (obtained using exhaustive search) and the proposed SCARP algorithm on a disaster-resilient communication network. The performance of SCARP is also compared with the SPR based scheduling and pairing.
- Chapter 5 investigates UAV-BS placement in a UAV-assisted communication network. The part I of the chapter is focused on setting up a disaster-resilient network utilizing UAV-BSs while taking into account the mobility of EFRs. Unlike part I, in part II of the chapter, we maintain the user coverage fairness and optimize the update interval.
- Chapter 6 extends the work done in part II of Chapter 5 to multi UAV-BS network while also accounting for resource allocation. We discuss two different optimization frameworks. Further, we propose algorithms to solve the proposed optimization problems.
- Finally, Chapter 7 concludes the work done towards this thesis. Further, we also highlight the possible future research directions.

Notations: $\mathcal{N}(\mu, \sigma^2)$ is used to denote a Gaussian distribution with mean μ and variance σ^2 . $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian distribution with mean μ and variance σ^2 . $\mathbb{E}[\cdot]$ is the expectation operator. $\mathbb{E}_x[\cdot]$ denotes the expectation with respect to x . \exp denotes an exponential distribution. $(\cdot)^T$ and $(\cdot)^{-1}$ denote the transpose and inverse operator respectively. $\|a\|$ denotes the norm of vector a . $\lceil \cdot \rceil$ denotes the ceil operator.

Assessing D2D/D2MD Networks Using Real-World User Mobility

In this chapter, we utilize the real-world location information of users to assess the viability of D2D/D2MD communication in indoor propagation environment. In part I of this chapter, we utilize the real-world location information of users to assess D2D in a conventional cellular network. We consider a heterogeneous framework where user devices are capable of supporting cellular as well as D2D communication. Further, based on real-world location information probability of D2D communication at a given time of the day is evaluated. This probability information is utilized to quantify the performance of the proposed D2D-enabled cellular framework and compared with the conventional cellular framework. This work is discussed in Sections 2.1 to 2.3.

Next, in part II of this chapter, we focus on the realistic performance evaluation of the D2MD network where the joint spatio-temporal user mobility behavior is taken into account. Further, in order to extract the required joint spatio-temporal behavior from the past real-world location information, we propose two novel methods, namely observed mobility exploitation (OME) and expected mobility exploitation (EME). OME is applicable when past location information of all the users is available. On the other hand, EME can be utilized when the location information of a set (i.e., the training set) of users, following an average spatial-temporal behavior similar to the intended users, is available. The realistic performance of two D2MD networks is evaluated for a tractable campus set-up in terms of content delivery time and data offload ratio. Sections 2.4 to 2.8 present the details of the work done on D2MD networks.

Now, we present our work on the performance analysis of D2D enabled cellular networks based on real-world user mobility.

2.1 Proposed Model for Analysis

In the proposed analysis, two users denoted as User 1 and User 2, are considered. The users can opt for either cellular or D2D communication to communicate with each other. It has been assumed that users are allocated orthogonal resource blocks¹ (RBs) for communication. So, a specific RB will be accessed by the cellular transmitter or by the D2D pair.

The users are assumed to be the students present in our IIIT-Delhi campus. The mobility of the users was studied every day from 9:00 a.m. till 7:00 p.m. for a period of one academic semester. During this specific time frame the campus is vibrant with students present in one building or another. The location data of the users is extracted from the Wi-Fi access points that are distributed all over the campus. Fig. 2.1 shows the building locations within the campus. On a typical weekday, the users are generally present in the hostel building (Building H), academic building (Building A), or student center (Building S). Whereas, on the weekends, the users can be present in any of the three buildings or can be outside the campus (O). Using the past data, we generate a probability for the two users of interest to be at a specific indoor site, as discussed in the next subsection.

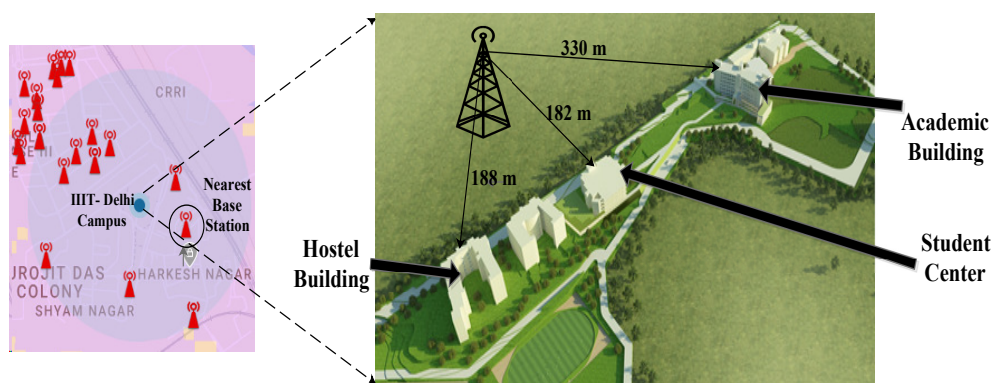


Figure 2.1: Layout of IIIT-Delhi Campus.

2.1.1 Determining the Diurnal User Probability

The fact that human mobility generally follows a regular pattern encouraged us to use the past data for determining the probability with which a user can be found in a building in a given time slot. The time from 9:00 a.m. to 7:00 p.m. was divided into 20 time slots of 30 minutes each and the location of the users for each slot was recorded. In the analysis, it is assumed that on an average a user tends to stay in a building for a minimum of 30 minutes, which is good enough considering the low mobility of the user across the buildings. The location of users can

¹The orthogonal resource blocks can be time, frequency, codes or a combination of two.

Table 2.1: Probability for Weekdays and Weekends

Slot Number	Time of the Day	Weekdays								Weekends							
		User 1				User 2				User 1				User 2			
		$P_{\mathbf{H}}^{Bx}$	$P_{\mathbf{A}}^{Bx}$	$P_{\mathbf{S}}^{Bx}$	$P_{\mathbf{O}}^{Bx}$	$P_{\mathbf{H}}^{Bx}$	$P_{\mathbf{A}}^{Bx}$	$P_{\mathbf{S}}^{Bx}$	$P_{\mathbf{O}}^{Bx}$	$P_{\mathbf{H}}^{Bx}$	$P_{\mathbf{A}}^{Bx}$	$P_{\mathbf{S}}^{Bx}$	$P_{\mathbf{O}}^{Bx}$	$P_{\mathbf{H}}^{Bx}$	$P_{\mathbf{A}}^{Bx}$	$P_{\mathbf{S}}^{Bx}$	$P_{\mathbf{O}}^{Bx}$
1	9-9:30	1	0	0	0	1	0	0	0	0.7	0	0	0.3	0.7	0	0	0.3
2	9:30-10	0.59	0	0.41	0	1	0	0	0	0.6	0	0.1	0.3	0.7	0	0	0.3
3	10-10:30	0.9	0.1	0	0	0.82	0	0.18	0	0.4	0	0.3	0.3	0.6	0	0.1	0.3
4	10:30-11	0.72	0.24	0.04	0	0.88	0.06	0.06	0	0.5	0.2	0	0.3	0.6	0	0.1	0.3
5	11-11:30	0.59	0.41	0	0	0.94	0.06	0	0	0.5	0.2	0	0.3	0.7	0	0	0.3
6	11:30-12	0.54	0.46	0	0	0.88	0.12	0	0	0.4	0.2	0	0.4	0.7	0	0	0.3
7	12-12:30	0.45	0.55	0	0	0.88	0.12	0	0	0.3	0.2	0	0.5	0.7	0	0	0.3
8	12:30-13	0.32	0.63	0.05	0	0.88	0.12	0	0	0.2	0.3	0	0.5	0.7	0	0	0.3
9	13-13:30	0.23	0.64	0.13	0	0.76	0.17	0.07	0	0.2	0.3	0	0.5	0.6	0.1	0	0.3
10	13:30-14	0.18	0.45	0.37	0	0.47	0.29	0.24	0	0.2	0.2	0.1	0.5	0.5	0.1	0.1	0.3
11	14-14:30	0.23	0.55	0.22	0	0.35	0.36	0.29	0	0.2	0.2	0.1	0.5	0.3	0.1	0.2	0.4
12	14:30-15	0.41	0.59	0	0	0.12	0.7	0.18	0	0.2	0.3	0	0.5	0.2	0.2	0.2	0.4
13	15-15:30	0.28	0.72	0	0	0.17	0.7	0.13	0	0.1	0.4	0	0.5	0.3	0.2	0	0.5
14	15:30-16	0.363	0.63	0	0	0	1	0	0	0.2	0.4	0	0.4	0.2	0.2	0	0.6
15	16-16:30	0.13	0.86	0	0	0.06	0.88	0.06	0	0.1	0.5	0	0.4	0.2	0.2	0	0.6
16	16:30-17	0.09	0.9	0	0	0.06	0.88	0.06	0	0.1	0.5	0	0.4	0.2	0.2	0	0.6
17	17-17:30	0.09	0.86	0.04	0	0	0.94	0.06	0	0.1	0.5	0	0.4	0.1	0.2	0.1	0.6
18	17:30-18	0.04	0.95	0	0	0.07	0.76	0.17	0	0.1	0.5	0	0.4	0.2	0.2	0	0.6
19	18-18:30	0.18	0.77	0.04	0	0.06	0.7	0.24	0	0.1	0.5	0	0.4	0.2	0.2	0	0.6
20	18:30-19	0.18	0.81	0	0	0.12	0.88	0	0	0.1	0.5	0	0.4	0.2	0.2	0	0.6

be in buildings **H**, **A** or **S** on weekdays or buildings **H**, **A**, **S** or **O** on weekends. Consequently, based on the readings for each slot a probability term associated with each building for each user is determined as shown in Table 2.1.

2.1.2 Propagation Channel Modeling

We have adopted WINNER II channel models for modeling the D2D and cellular links [31]. WINNER II specifies path loss models for different propagation environments. Out of the path loss models specified in [31], models that describe the indoor to outdoor scenario and indoor office scenario are of interest to us². The path loss models for these scenarios are given below:

Indoor to Outdoor (urban macro cell)

$$PL(\text{in dB}) = [44.9 - 6.55 \log_{10}(h_{BS})] \log_{10}(d) + 34.46 + 5.83 \log_{10}(h_{BS}) + 23 \log_{10}(f_c/5) + 17.4 + 0.5d_{in} - 0.8h_{MS}, \quad (2.1)$$

where base station height, $h_{BS} = 25$ m, mobile device height, $h_{MS} = 3n_{Fl} + 1.5$ m and n_{Fl} is the floor index (for ground floor $n_{Fl} = 1$). d_{in} is the distance from the indoor transmitter to the wall, d_{out} is the distance between the point on the wall that is nearest to the indoor transmitter and base station, and $d = d_{out} + d_{in}$.

Indoor Office

$$PL = 20 \log_{10}(d) + 46.4 + 20 \log_{10}(f_c/5) + WL + FL, \quad (2.2)$$

²Indoor to outdoor scenario to be same as outdoor to indoor (urban macro cell) scenario due to the principle of channel reciprocity.

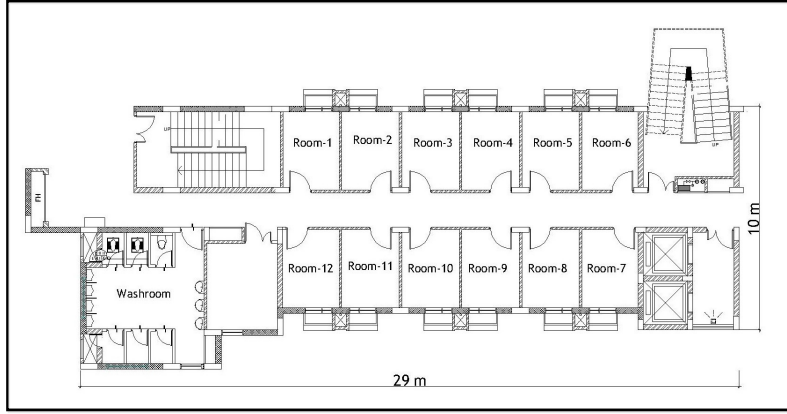


Figure 2.2: Floor layout of hostel building (building **H**).

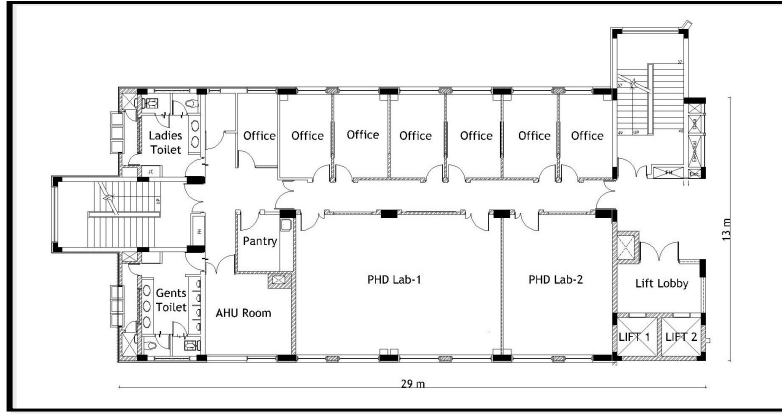


Figure 2.3: Floor layout of academic building (building **A**).

where d is the distance between the D2D transmitter and receiver in meters, f_c is the frequency³ of operation in GHz, $WL = 5n_w$ (for soft walls) or $12n_w$ (for hard walls) and n_w are the number of walls penetrated by the signal. The campus buildings in our case have hard wall partitions thus $WL = 12n_w$ has been considered. $FL = 17 + 4(n_{fl} - 1)$ and n_{fl} is the number of floors penetrated by the signal. In this work, (2.1) and (2.2) have been applied for modeling the cellular and D2D links respectively. The antenna gains at the base station and user devices are assumed to be 10 dBd and 0 dBd respectively.

Fig. 2.2, 2.3 and 2.4 show the floor layouts for the buildings **H**, **A** and **S**. All the floors of a building have a similar floor layout. In our analysis, a user in any given building at any point of time will most likely be present on a specific floor and room. This is because User 1 has a hostel Room-1 assigned in 1st floor and User 2 has a Room-2 assigned in 3rd floor. Similarly, in building **A**, User 1 has a cubicle in PhD Lab-1 on 2nd floor and User 2 has a cubicle in PhD Lab-2 on 4th floor. Hence, in the case of D2D communication, when both users are present

³WINNER II model can be applied to all LTE carrier frequencies falling in the range of 2-6 GHz and in our case we have $f_c = 2.3$ GHz.

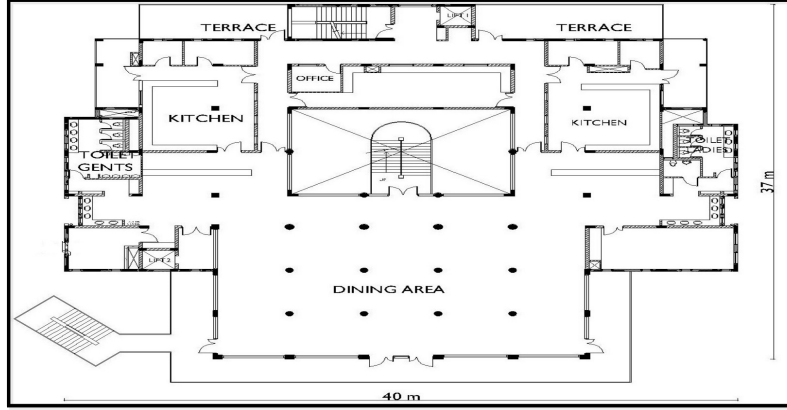


Figure 2.4: Floor layout of student center building (building S).

in building **H** or **A** the floor separation is two. When in building **S** the users typically are on the ground floor and 1st floor dining areas and therefore a maximum separation of one floor. So, $n_{fl} = 2$ for buildings **H** and **A** whereas $n_{fl} = 1$ for building **S**. Further, as evident from the floor layout, the horizontal separation between the users is limited to a single wall, i.e., $n_w = 1$.

The ceiling height in building **H** and **A** is 3 m. Hence, the maximum transmission distance⁴ between the two users present in building **H** and **A** will be 6 m. Further, the ceiling height of the building **S** is 7 m, and hence maximum transmission distance will be 7 m for the users located in this building. When User 1 or (and) User 2 are outside campus, D2D communication may not be possible, hence, not considered in the proposed work. Thus, in such a scenario, an average cellular transmission distance of 250 m is assumed⁵.

2.2 Cellular V/s D2D+Cellular Framework

For the performance analysis, two transmission frameworks are considered:

1. Cellular: In this framework, the two users always use the conventional cellular mode of communication to communicate with each other.
2. D2D+Cellular: In this framework, if the two users are in the same building, they will use D2D mode of communication; otherwise, they would communicate via cellular mode.

2.2.1 Probability of D2D Communication

First, we calculate the probability with which D2D can occur in the D2D+cellular framework. The probability of D2D communication in n^{th} time slot for both weekdays and weekends can

⁴Restricting the transmission distance to be equal to ceiling height times number of floors penetrated.

⁵Urban scenario with a high density of cellular towers and cellular range of 500 m [76].

Table 2.2: Probability of D2D communication on Weekdays and Weekends

Slot Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Weekdays	1	0.59	0.73	0.64	0.57	0.53	0.46	0.35	0.29	0.30	0.33	0.46	0.55	0.63	0.76	0.79	0.81	0.72	0.56	0.74
Weekend	0.49	0.42	0.27	0.30	0.35	0.28	0.21	0.14	0.15	0.13	0.10	0.10	0.11	0.12	0.12	0.12	0.11	0.12	0.12	0.12

be given by:

$$P_{D2D}^n = P_{\mathbf{H}}^{tx} P_{\mathbf{H}}^{rx} + P_{\mathbf{A}}^{tx} P_{\mathbf{A}}^{rx} + P_{\mathbf{S}}^{tx} P_{\mathbf{S}}^{rx}, \quad (2.3)$$

where $P_{\mathbf{H}}^{tx}$, $P_{\mathbf{A}}^{tx}$ and $P_{\mathbf{S}}^{tx}$ are the probabilities of the D2D transmitter (User 1) being in buildings **H**, **A** and **S** in a given time slot respectively. Similarly, $P_{\mathbf{H}}^{rx}$, $P_{\mathbf{A}}^{rx}$ and $P_{\mathbf{S}}^{rx}$ are the probabilities of the D2D receiver (User 2) being in buildings **H**, **A** and **S** in a given time slot respectively. Table 2.2 shows the per time slot probability of D2D communication for weekdays and weekend as given in (2.3). The higher D2D probability in a slot corresponds to a higher chance of User 1 and User 2 being in the same building in a given time slot. For example, in the 1st slot i.e. 9:00 a.m. to 9:30 a.m. both the users are present in hostel with probability 1. This gives rise to a D2D probability of 1 in Table 2.2. Similarly, in the evening hours there is higher probability that both the users are present in building **A**. Hence, in those time slots D2D probability will be higher.

2.2.2 Performance Metrics

The performance comparison of the cellular and D2D+cellular frameworks using the spectral efficiency per slot is performed for both weekdays and weekends. Further, the amount of energy saving that can be achieved in D2D+cellular mode for a predefined quality of service (QoS) has been evaluated.

2.2.2.1 Spectral efficiency

The spectral efficiency achievable corresponding to average signal-to-noise ratio (SNR) values of each slot is given by:

$$S = \log_2(1 + \gamma_{avg}) \text{ bits/s/Hz}, \quad (2.4)$$

where γ_{avg} is the average signal-to-noise ratio of each slot. $\gamma_{avg} = \gamma_{avg}^C$ for cellular framework and $\gamma_{avg} = \gamma_{avg}^D$ for D2D+cellular framework. We will have different SNR values depending on the location of the cellular transmitter in case of cellular communication and location of the D2D pair in the case of D2D communication. $\gamma_{\mathbf{B}}^C$, $\gamma_{\mathbf{B}}^D$ are the SNR values for the cellular and D2D communication respectively and here $\mathbf{B} \in \{\mathbf{H}, \mathbf{A}, \mathbf{S}\}$. γ_0 is the SNR value when the cellular transmitter is outside campus. In general, SNR (dB) is given by :

$$\gamma = P_r(\text{dBm}) - \sigma_o^2(\text{dBm}), \quad (2.5)$$

where P_r is the received power, which is given by $P_r = P_t - PL$. P_t is the transmitted power in dBm and σ_o^2 is the noise power at receiver.

In the cellular mode, the average SNR values per slot are evaluated as a weighted sum of the SNR values corresponding to \mathbf{H} , \mathbf{A} , \mathbf{S} and \mathbf{O} . These weights are the transmitter's probability of being present in each building per slot.

$$\gamma_{avg1}^C = P_{\mathbf{H}}^{tx} \gamma_{\mathbf{H}}^C + P_{\mathbf{A}}^{tx} \gamma_{\mathbf{A}}^C + P_{\mathbf{S}}^{tx} \gamma_{\mathbf{S}}^C, \quad (2.6)$$

$$\gamma_{avg2}^C = P_{\mathbf{H}}^{tx} \gamma_{\mathbf{H}}^C + P_{\mathbf{A}}^{tx} \gamma_{\mathbf{A}}^C + P_{\mathbf{S}}^{tx} \gamma_{\mathbf{S}}^C + P_{\mathbf{O}}^{tx} \gamma_{\mathbf{O}}, \quad (2.7)$$

where $P_{\mathbf{O}}^{tx}$ is the probability of user being outside campus in a given time slot on weekends. (2.6) and (2.7) show the average SNR values for cellular mode on weekdays and weekends respectively.

Similarly, the average SNR per slot in D2D+cellular framework will be a weighted sum of SNR values corresponding to D2D and cellular communication in each building. For D2D communication, SNR values are weighted by the joint probability of D2D pair in each building and for cellular communication the weights consider the probability when D2D is not feasible.

$$\begin{aligned} \gamma_{avg1}^D &= P_{\mathbf{H}}^{tx} P_{\mathbf{H}}^{rx} \gamma_{\mathbf{H}}^D + P_{\mathbf{A}}^{tx} P_{\mathbf{A}}^{rx} \gamma_{\mathbf{A}}^D + P_{\mathbf{S}}^{tx} P_{\mathbf{S}}^{rx} \gamma_{\mathbf{S}}^D \\ &\quad + P_{\mathbf{H}}^{tx} (1 - P_{\mathbf{H}}^{rx}) \gamma_{\mathbf{H}}^C + P_{\mathbf{A}}^{tx} (1 - P_{\mathbf{A}}^{rx}) \gamma_{\mathbf{A}}^C \\ &\quad + P_{\mathbf{S}}^{tx} (1 - P_{\mathbf{S}}^{rx}) \gamma_{\mathbf{S}}^C, \end{aligned} \quad (2.8)$$

$$\begin{aligned} \gamma_{avg2}^D &= P_{\mathbf{H}}^{tx} P_{\mathbf{H}}^{rx} \gamma_{\mathbf{H}}^D + P_{\mathbf{A}}^{tx} P_{\mathbf{A}}^{rx} \gamma_{\mathbf{A}}^D + P_{\mathbf{S}}^{tx} P_{\mathbf{S}}^{rx} \gamma_{\mathbf{S}}^D \\ &\quad + P_{\mathbf{H}}^{tx} (1 - P_{\mathbf{H}}^{rx}) \gamma_{\mathbf{H}}^C + P_{\mathbf{A}}^{tx} (1 - P_{\mathbf{A}}^{rx}) \gamma_{\mathbf{A}}^C \\ &\quad + P_{\mathbf{S}}^{tx} (1 - P_{\mathbf{S}}^{rx}) \gamma_{\mathbf{S}}^C + P_{\mathbf{O}}^{tx} \gamma_{\mathbf{O}}. \end{aligned} \quad (2.9)$$

The above equations (2.8) and (2.9) represent the average SNR values for D2D+cellular mode for weekdays and weekends, respectively.

2.2.2.2 Energy Saving

For a cellular network, there is always a trade-off between the transmission power and the achievable QoS. The achievable QoS can be measured in terms of received SNR, bit error rate (BER), throughput, etc. In order to achieve a better QoS, more transmission power is required; hence more energy is consumed. In the following, we show that opting for D2D+cellular framework can lead to considerable energy saving compared to the conventional cellular framework. The amount of energy saved depends on the time slot under consideration and also on the fact that whether it is a weekday or weekend.

Let us consider a scenario where a fixed QoS needs to be maintained in cellular and D2D+cellular frameworks. This implies same average SNR is to be achieved in the two modes. Hence, on equating (2.6) and (2.8), the reduced transmitted power, P' corresponding to the D2D link within the D2D+cellular framework⁶, per slot for weekdays, will be:

$$P' = \frac{\gamma_{avg}^C - \tau}{\alpha + \beta + \eta}, \quad (2.10)$$

where $\alpha = \frac{P_H^{tx} P_H^{rx}}{PL_{HN}^D}$, $\beta = \frac{P_A^{tx} P_A^{rx}}{PL_{AN}^D}$, $\eta = \frac{P_S^{tx} P_S^{rx}}{PL_{SN}^D}$,

$$\tau = P_H^{tx}(1 - P_H^{rx})\gamma_H^C + P_A^{tx}(1 - P_A^{rx})\gamma_A^C + P_S^{tx}(1 - P_S^{rx})\gamma_S^C,$$

and PL_B^D is the path loss in case of D2D. Similarly, the reduced transmitted power for weekend can be calculated. The two users are assumed to be communicating with each other for a duration of 60, 90 or 120 minutes from 9 a.m. to 7 p.m. on a weekday or weekend. This means a mobile usage of $k = 3, 4.5$ or 6 minutes in each slot. However, the usage time can be fixed to any value. The energy saved in each slot can be determined in the following manner:

$$\text{Energy saving} = [P' P_{D2D}^n + P_t(1 - P_{D2D}^n)] \times k \times 60. \quad (2.11)$$

2.3 Results and Discussion

This section presents the results for spectral efficiency for the cellular and D2D+cellular frameworks, and the amount of energy saving that can be obtained in D2D+cellular framework due to reduced transmitted power over the D2D link. In the following simulations, $P_t = 23$ dBm and $\sigma_o^2 = -100$ dBm has been considered. Fig. 2.5 shows the spectral efficiency achievable in cellular and D2D+cellular frameworks. It can be observed that slots with higher D2D probability (D2D probability can be seen in Table 2.2) have higher spectral efficiency. This is due to the fact that when D2D probability is high, there are more chances of D2D communication; hence average SNR improves. As a consequence, higher spectral efficiency is obtained.

It can also be observed that the spectral efficiency for D2D+cellular on weekdays is more than that on weekends because the D2D communication probability is significantly less on weekends. For instance, if we compare the 10th time slot for User 1 and User 2, we observe that the spectral efficiency values for weekdays and weekends are 11.68 bits/s/Hz and 10.3 bits/s/Hz respectively. It can be explained as follows. Using (2.5), (2.8) and (2.9), γ_{avg} comes out to be 35 dB for weekdays and 31 dB for weekends. Hence, we have $S = 11.68$ bits/s/Hz

⁶Only the transmitted power for D2D gets reduced, cellular transmitted power will be the same in both frameworks.

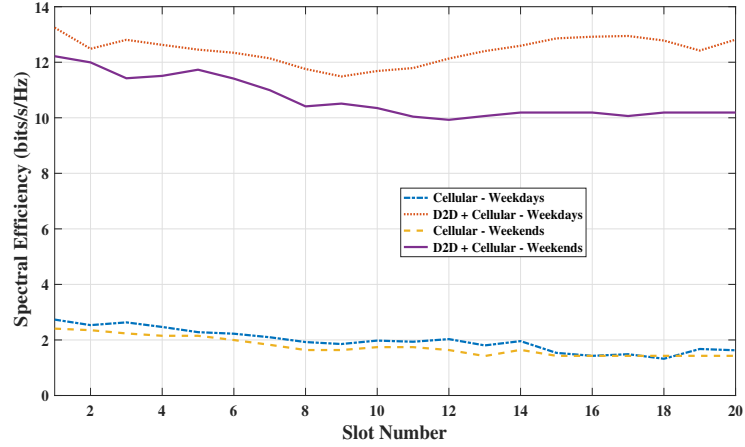


Figure 2.5: Spectral efficiency on weekdays and weekends.

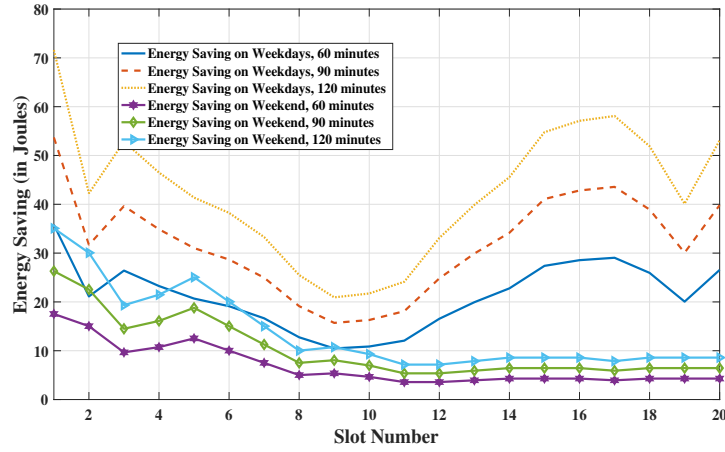


Figure 2.6: Energy saving on weekdays and weekend.

for weekdays and $S = 10.3$ bits/s/Hz for weekends. Using (2.5), (2.6), (2.7) and (2.4) spectral efficiency for cellular framework is obtained as 1.97 bits/s/Hz and 1.74 bits/s/Hz on weekdays and weekends respectively. This implies a spectral efficiency gain of 9.71 bits/s/Hz and 8.56 bits/s/Hz on weekdays and weekends, respectively, in the 10th slot. After a similar evaluation for the other time slots, the result is an average spectral efficiency gain of 10.45 bits/s/Hz and 8.93 bits/s/Hz on any given weekday and weekend, respectively.

It can be seen from Fig. 2.6, as mobile usage increases from 60 minutes to 120 minutes, the amount of energy saved also increases. This is because the time over which the transmitter is operating in the D2D mode increases. Also, while using D2D communication, the transmitted power can be reduced to achieve the fixed QoS. The higher the D2D probability, the higher are the chances of D2D communication. Hence, there will be an increase in the amount of energy saved. Due to this combined effect of mobile usage duration and probability of D2D communication, the energy saving pattern varies over different time slots of weekdays and

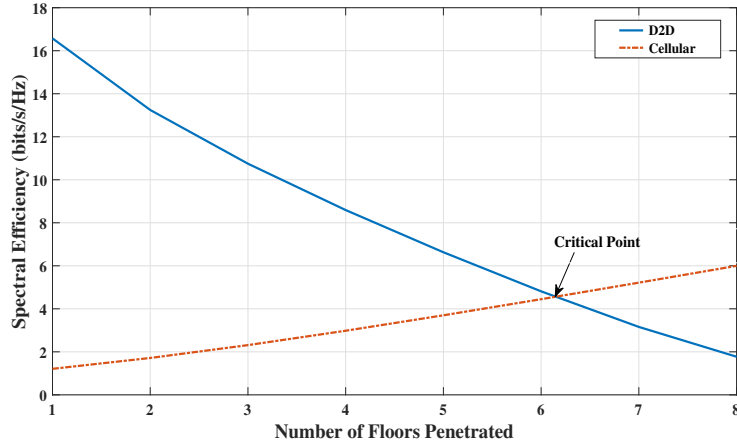


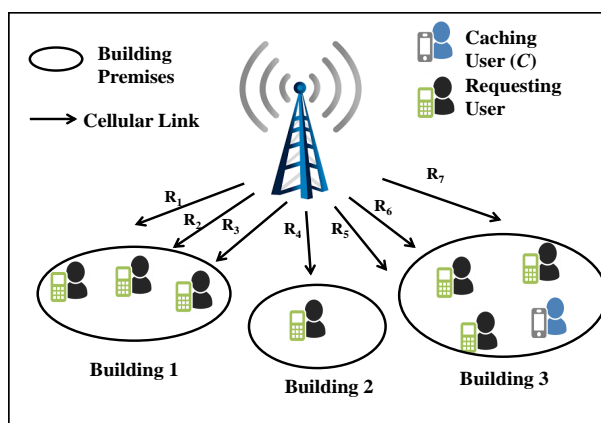
Figure 2.7: Impact of n_{fl} on spectral efficiency of D2D and cellular communication.

weekends. Let us illustrate through an example. In the 10^{th} slot, using (2.10) and (2.11), the amount of energy saved is 21.72 J and 9.3 J on weekdays and weekends respectively for $k = 6$. Similarly, on determining the energy saved for each of the time slots, on average, 42.60 J and 13.88 J of energy is saved on weekdays and weekends. This accounts for 59.64 % and 19.34 % average energy savings on weekdays and weekends, respectively.

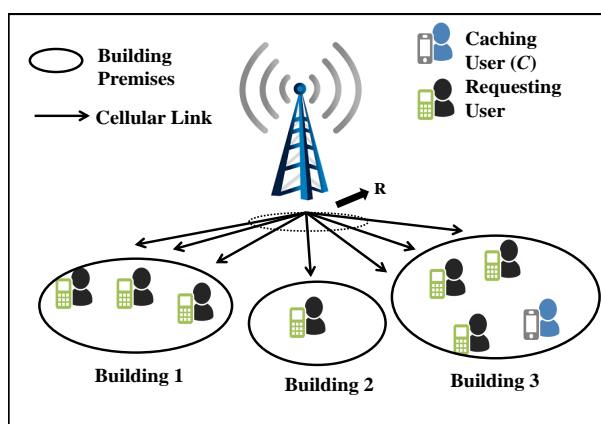
2.3.1 Effect of Floors

In this section, we investigate the impact of the number of floors on the performance of D2D communication vis-a-vis cellular communication. Specifically, we have restricted our analysis to two users present in the same building **A**. The receiver's location is fixed on the ground floor and the transmitter's location is varied from floor 1 to floor 8. In line with the previous D2D framework, we assume that if the two users are in the same building, they will establish a D2D communication link. We measured the spectral efficiency and compared it with the conventional cellular scenario. Fig. 2.7 shows the spectral efficiency for the D2D and cellular scenarios. It is evident that after a critical point (corresponding to a separation of 6 floors), even D2D communication link starts performing worse than the cellular link. Beyond the critical point, with a link distance ranging from 18-24 m (building **A**) D2D communication is unable to match up to the performance of cellular communication having a link distance of 330 m. Thus, validating the point that D2D communication cannot be opted solely on the basis of proximity as the impact of the number of floors in an indoor environment should also be considered.

Now, let us characterize the joint spatio-temporal user behavior for the D2MD network and assess the network performance.



(a) Conventional Unicast Scheme



(b) Conventional Multicast Scheme

Figure 2.8: Illustration of Conventional Transmission Schemes for $U = 7$.

2.4 Detailed Set-Up and Data Collection for D2MD

Let at any given point of time there be U users, constituting a set \mathcal{U} , spread across b regions-of-interest (ROIs) or buildings that request multimedia content from the core network. These users share common interests, and hence, they are part of a social group \mathcal{K} that consists of K ($U < K$) users [41]. A multimedia content request will be sent by the users to the BS. Conventionally, BS will review the request and either establish one link per user (unicast) or one link per multicast group formed by the U users and transmit the content. Figs. 2.8(a) and 2.8(b) illustrate the communication links established in cellular networks using unicast and multicast schemes for $U = 7$ and $b = 3$ respectively. Users in Fig. 2.8(a) can achieve rates $R_1, R_2 \dots R_7$ bits/sec whereas users in Fig. 2.8(b) can achieve R bits/sec where R is the rate supported by the user with the worst channel. In the unicast scenario, it is assumed one RB is assigned to each requesting user, whereas one RB is allocated per multicast group in cellular multicast.

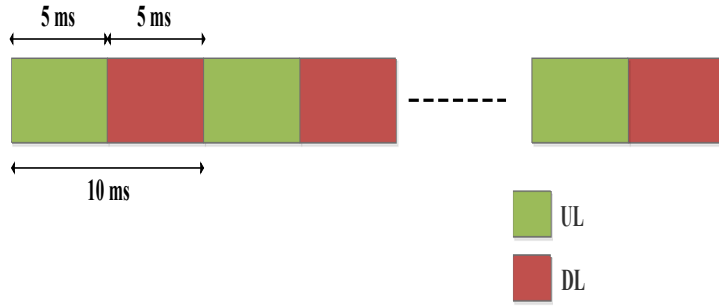


Figure 2.9: TDD frame structure.

For the case of network-assisted D2D, it is assumed that BS maintains a social directory [77]. The social directory is updated using the social structure information available with the BS. Further, the location information of the users can also be retrieved by the network, and hence network can discover devices in each others' proximity. For the cases where the location information is unavailable with the network, this proximity information can be determined locally by the users and sent to the BS [78]. The social information along with the proximity information will be utilized by the BS to choose between D2D and cellular modes of communication. The popular content requested by users is assumed to be cached with the user C , where $C \in \mathcal{H}$. As $C \in \mathcal{H}$, C will be willing to share its cached content with other users.

However, for realistic performance analysis of D2MD networks before actual deployment, the spatio-temporal behavior of the K users will be required. In the proposed work, we have utilized our institute's campus set-up and have collected the real-world location information of the users that belong to the same social group⁷. On a typical working day, users can be present in any of the three campus buildings, namely hostel (**H**), academic (**A**), or student center (**S**). The floor layouts of the buildings **H** and **A** are similar. However, building **S** has a different floor layout [79]. The campus layout with the cellular link distance for each building is shown in Fig. 2.1. As mentioned before, since the considered users are students, the location samples from a semester will be sufficient to fully capture the mobility information. The location data has been gathered from 9:00 to 19:00 hrs (20 slots of 30 minutes each) in an uncontrolled environment⁸. Wi-Fi signals have been used as location beacons to identify the location of the user. Please note, due to mobility and channel fluctuations; sometimes the users are "disconnected" from the Wi-Fi network. Hence, the data from the preceding slot has been extrapolated to the "disconnected" slot. However, the number of these slots where data is extrapolated is very few compared to the total number of slots for which data was collected.

⁷We ensure that the users belong to the same social group by considering only those users who belong to the same friend circle.

⁸During the data collection drive the users were requested to maintain their usual Wi-Fi usage.

Further, the cellular and D2D users operate in the time division duplexing (TDD) mode⁹ employing a frame structure of 10 ms similar to LTE-TDD Type-1 [81] frame structure with equal DL and UL time resources as shown in Fig. 2.9.

2.5 Proposed Methods For Joint Spatio-Temporal Behavior Extraction

In this section, we present the two proposed methods to acquire joint spatio-temporal behavior of users in \mathcal{K} . In OME, the joint spatio-temporal behavior of users is extracted from the past location information of the K users. On the other hand, EME utilizes a training dataset to estimate the joint spatio-temporal behavior of \mathcal{U} when the past location information of \mathcal{U} is not available¹⁰. The methods have been explained in detail in the following subsections.

2.5.1 Observed Mobility Exploitation (OME)

As mentioned above, this method is applied to extract the joint spatio-temporal behavior from the past location information of users in \mathcal{K} . From location information of users in \mathcal{U} for n^{th} time slot, observed occupancy states can be found where occupancy state denotes the number of users in each building. The occupancy states are studied to characterize the joint user spatio-temporal behavior. Each observed state is denoted as a vector $X^n = (x_1, x_2, \dots, x_b)$. Depending on the time of the day or time slot, the probability of each state will vary. In the considered campus set-up, $b = 3$ and x_1, x_2 and x_3 is denoted as $x_{\mathbf{H}}, x_{\mathbf{A}}$ and $x_{\mathbf{S}}$ respectively, where $x_{\mathbf{H}}, x_{\mathbf{A}}$ and $x_{\mathbf{S}}$ are the occupancies of buildings \mathbf{H}, \mathbf{A} and \mathbf{S} respectively, hence $X^n = (x_{\mathbf{H}}, x_{\mathbf{A}}, x_{\mathbf{S}})$. For instance, in slot $n = 10$ and for $U = 5$ following states were observed: [0 2 3], [0 3 2], [0 4 1], [1 1 3], [1 2 2], [1 3 1], [2 1 2], [2 2 1], [2 3 0]. It is also evident from above that the number of observed occupancy states, which in our example is 9, is significantly less than the total possible number of occupancy states, which in our example can be 32. This is due to the fact that rather than being randomly distributed across the three buildings, the users which are students in the proposed set-up are more likely to be present in buildings \mathbf{S} and \mathbf{A} in $n = 10$ (i.e., 13:30 to 14:00 hrs) not in building \mathbf{H} .

As the location of cache is crucial for D2MD group formation, the location information is jointly considered with X^n . Hence, we denote $\Omega_B^n = P[X^n = (x_1, x_2, \dots, x_b), C^B]$ as the joint probability of observed state (x_1, x_2, \dots, x_b) and event C^B , where $B \in \{1, 2, 3, \dots, b\}$. C^B is the event that C is in building B . These joint probability terms are evaluated for each observed state

⁹TDD is a preferred duplexing mode over FDD for D2D supporting cellular networks [80].

¹⁰The past location information of C is assumed to be always available. Consequently, the estimation of C 's behavior is not required while using EME.

Table 2.3: p-Values

Distributions	p-Value		
	H	A	S
Poisson	0.558	0.087	0.938
Geometric	7.75×10^{-5}	5.4×10^{-13}	0.6×10^{-3}
Discrete Uniform	0.1919	0.042	0.1769

and defined as the ratio of the number of days when (x_1, x_2, \dots, x_b) and event C^B are observed in n^{th} slot to the number of days in the location database. This time-slotted model captures the fluctuations in the building occupancy due to change in the time slot as well as change of day. Hence, the location database is utilized to determine the mobility statistics corresponding to each observed occupancy state that in turn define the joint spatio-temporal behavior of users. The above mobility statistics can be applied for a realistic evaluation of the D2MD networks, provided sufficient past location samples are available.

2.5.2 Expected Mobility Exploitation (EME)

Let us now assume that the location information of set of requesting users \mathcal{U} is unknown. A larger community of users consisting of \mathcal{K}^{train} and \mathcal{U} is considered where users typically have similar spatio-temporal behavior. However, the past location information of K^{train} users constituting the set \mathcal{K}^{train} is known. Therefore, utilizing the expected mobility exploitation (EME) method, the joint spatio-temporal behavior of \mathcal{U} can be estimated using \mathcal{K}^{train} . For estimating the joint spatio-temporal behavior of \mathcal{U} , first, the parameters defining the joint spatio-temporal behavior of \mathcal{K}^{train} are evaluated.

We define the joint spatio-temporal behavior of \mathcal{K}^{train} using the distribution of occupancy of each building in each slot. The occupancy of each building in a slot is modeled as the number of users arriving in a bounded region of space in a specific time interval which is stochastic in nature. For the considered campus set-up, we utilized graphical methods and chi-squared test for finding the distribution for the observed building occupancies. The p-values [82] of few standard discrete distributions, in contention for being a fit for the observed building occupancies, were determined. Let $\mathcal{U}^{train} \subseteq \mathcal{K}^{train}$ and consist of U^{train} users. Table 2.3 demonstrates the p-values obtained for standard discrete distributions when $U^{train} = 15$ and $n = 10$. To illustrate, let us consider building **S**, p-values for Poisson, geometric distribution and discrete uniform is 0.938, 0.0006 and 0.1769 respectively. However, for Poisson and discrete uniform distribution p-values being greater than the significance value of 0.05 implies that the null hypothesis (i.e., the location data having a Poisson or discrete uniform distribution) cannot be rejected. Further, the distribution with a higher p-value, i.e., Poisson fits the location data better. Hence, it can be inferred that the building occupancy, for a fixed U^{train} users,

in each of the buildings in each slot is Poisson distributed, i.e., $\text{Poisson}(\lambda_B^{n,U^{train}})$, where $\lambda_B^{n,U^{train}}$ is the expected occupancy of slot n , building B and U^{train} . On performing the chi-squared test for different values of $U^{train} \leq K^{train}$, buildings and slots, the p-values for Poisson distribution came out to be the highest and above significance level of 0.05. On generalizing the above finding, the users in each building or ROI are assumed to be distributed according to a spatial Poisson point process (SPPPs). Consequently, the building occupancies of \mathcal{U} can be assumed to be Poisson distributed, and users will follow SPPP. In general, these SPPPs are time-independent as they fail to relate the user locations to the time of the day [83]. In EME, K^{train} is used to estimate the time-dependent intensity (or expected/average occupancy) of SPPP followed by \mathcal{U} , unlike the work done in [28–30]. In other words, $\bar{\lambda}_B^{n,U^{train}}$ is a reasonable parameter to obtain $\hat{\lambda}_B^{n,U}$, i.e., the estimated average occupancy of the requesting users where $\bar{\lambda}_B^{n,U^{train}}$ is averaged over all possible \mathcal{U}^{train} and \mathcal{U}^{train} is the unique set of U^{train} . It may be noted that as U^{train} increases, it's quite intuitive that building occupancies will increase or will be the same as the previous value. The same has been observed while analyzing the data.

As $\bar{\lambda}_B^{n,U^{train}}$ is linearly dependent on U^{train} , linear regression is utilized to determine the relationship between $\bar{\lambda}_B^{n,U^{train}}$ and U^{train} for each building B and slot n . The relationship is modeled using simple linear regression as follows [84]:

$$\mathbf{y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.12)$$

where

$$\mathbf{y} = \begin{bmatrix} \bar{\lambda}_B^{n,1} \\ \bar{\lambda}_B^{n,2} \\ \cdot \\ \cdot \\ \bar{\lambda}_B^{n,N} \end{bmatrix}, \quad (2.13)$$

$$\mathbf{F} = \begin{bmatrix} 1 & U_1^{train} \\ 1 & U_2^{train} \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & U_N^{train} \end{bmatrix}, \quad (2.14)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \quad (2.15)$$

Table 2.4: r^2 values and coefficients of fitted curve

	r^2			β_2			β_1		
	H	A	S	H	A	S	H	A	S
Slot 1	0.9980	0.9221	0.9557	0.7630	0.0580	0.1790	0.5353	-0.1320	-0.4033
Slot 10	0.9805	0.9912	0.9782	0.3213	0.4271	0.2516	-0.4289	0.1366	0.2923

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_1 \\ \cdot \\ \cdot \\ \varepsilon_N \end{bmatrix}. \quad (2.16)$$

\mathbf{y} is the vector containing the average occupancies of building B and \mathbf{F} contains the data points U_1^{train} to U_N^{train} corresponding to the independent variable U^{train} . $\boldsymbol{\beta}$ is the parameter vector, $\boldsymbol{\varepsilon}$ is an unobserved random vector that adds “noise” to the linear relationship and $N = K^{train}$. To estimate $\boldsymbol{\beta}$, we have utilized the least squares estimation. The least squares solution is as follows:

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y} \quad (2.17)$$

The estimated linear curve coefficients or parameter vector $\hat{\boldsymbol{\beta}}$ will determine $\hat{\lambda}_B^{n,U}$. For example, Table 2.4 shows the coefficient of determination (r^2) values for the campus set-up which is a statistical measure of how well the given data fits the curve and coefficients of the fitted linear curves (β_1 and β_2) [85]. Here, $K^{train} = 40$ which is a sufficient dataset to capture the mobility behavior of the community of users having similar spatio-temporal behavior. r^2 ranges from 0 to 1, where a higher value depicts how well the curve fits the observed outcomes. The value $(\beta_2 \times U) + \beta_1$ per slot per building will give an estimate of building occupancy for a given number of requesting users $U \in \mathcal{K}$. Similarly, the estimates can be obtained for any network set-up and applied to \mathcal{U} to evaluate its D2MD performance. Further, the probability of C being in building B is given as P_B^n , which is calculated from the caching user’s past location information.

2.6 Interference Model

The location of cache C , the building occupancies, and the cellular channel conditions are the factors governing the formation of the D2MD group. If at a given time instant, C is in building B and there are one or more requesting users with link to cache C better than the cellular link

then a multicast group can be formed in building B . Therefore, as there is one cache¹¹; at most one multicast group will exist at a time instant. More details of the above factors have been provided in Section 2.7.

This D2MD group will reuse the one RB assigned to some cellular user C_R in the UL phase; however, this will result in interference at the requesting users, who are part of the D2MD group, due to transmission from C_R as shown in Fig. 2.10. Similarly, C will create interference at BS. C_R can be any user randomly located within the cellular coverage area of the BS. Consequently, the interference I_{d2d_i} at the i^{th} requesting user, which is the part of the D2MD group, can be modeled as below:

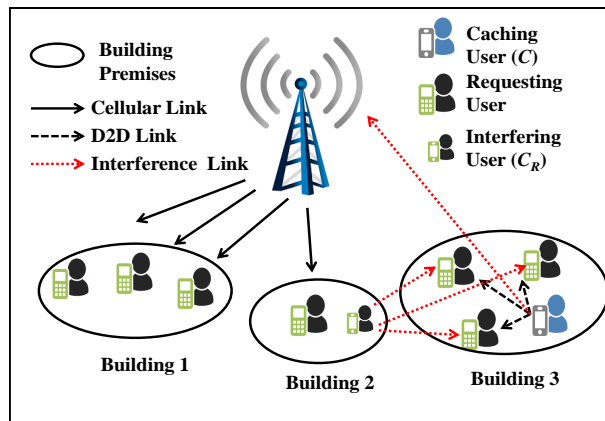
$$I_{d2d_i} = P_t^{UL} |h|^2, \quad (2.18)$$

where $h \sim \mathcal{CN}(0, (L^{C_R, i})^{-1})$ and $L^{C_R, i}$ is the path loss between C_R and i^{th} user calculated using the WINNER II model for cellular mode, and P_t^{UL} is the transmitted power of mobile device. The distance between the i^{th} user in the D2MD group and C_R is randomly selected from [10, 250] m.

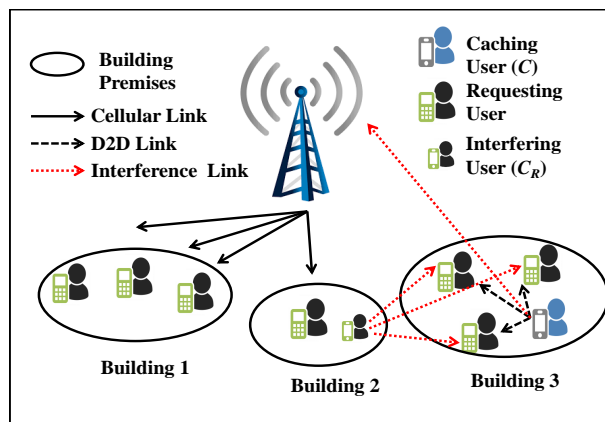
2.7 D2MD Networks And Achievable Rates

In the proposed work, we have considered two types of D2MD networks, D2MD-U and D2MD-M. In D2MD-U, a user can opt for either D2MD or cellular unicast to access the requested content. On the other hand, in D2MD-M, a user can opt for either D2MD or cellular multicast for accessing the requested content. D2MD or cellular unicast/multicast access will depend on the location of the user. Fig. 2.10 exhibits the communication scenario for an occupancy state of [3 3 1] in the two D2MD networks when $U = 7$ and $b = 3$. It can be observed that user C is present in building 3 and $x_3 \geq 1$. Therefore, a D2MD group will be formed in building 3. The rest of the users will either be served by the cellular unicast or multicast links depending on the type of D2MD network, D2MD-U, or D2MD-M. The number of users who are part of the D2MD group will vary with the time of the day. Further, the location of the D2MD group will also vary with time. Let there be W ($W \leq U$) users denoted by set \mathcal{W} in the building B where C is also present. Only $\mathcal{W}' \subseteq \mathcal{W}$ will be directed by the BS to form a D2MD group. In the following subsections, we will analyze the achievable rates for the two D2MD networks using OME and EME methods.

¹¹One cache has been considered to maintain the analytical tractability of rate expressions. However, the proposed work can be extended to the case wherein multiple caches are selected.



(a) D2MD-U



(b) D2MD-M

Figure 2.10: Illustration of D2MD-U and D2MD-M networks for an observed occupancy state [3 3 1].

2.7.1 Achievable Rate in D2MD-U

As mentioned above, the D2MD group formation in D2MD-U will also depend on the instantaneous channel conditions of cellular link. Let us first determine the SNR expressions for the conventional cellular unicast scheme. The instantaneous SNR of the cellular link corresponding to building $B \in \{1, 2, \dots, b\}$ will be:

$$\gamma_B^{uni} = \frac{P_t^{DL} (|g^B|^2)}{N_o BW}, \quad (2.19)$$

where $|g^B|^2$ is the channel gain of cellular link corresponding to building B calculated by averaging over the individual cellular channel gains of users in B in DL, N_o is the additive white Gaussian noise (AWGN) spectral density and BW is the bandwidth of one RB. The i^{th} user with instantaneous SINR, $\gamma_B^i = \frac{P_t^{UL} (|f_i^B|^2)}{I_{d2d_i} + N_o(BW)} \geq \gamma_B^{uni}$ is included in \mathcal{W}' . $f_i^B \sim \mathcal{CN}(0, (L^{C,i})^{-1})$ is the coefficient of Rayleigh fading channel between i^{th} user and C , and $L^{C,i}$ is the path loss. Now, the content will be disseminated to the W' users by C using one RB of C_R in UL phase. $(U - W')$ users will get the content in DL phase using one RB per cellular unicast link.

The instantaneous signal-to-noise and interference ratio (SINR) at i^{th} requesting user when it is served via D2MD in building B will be:

$$\gamma_B = \frac{P_t^{UL} (\min(|f_1^B|^2, |f_2^B|^2, \dots, |f_i^B|^2, \dots, |f_{W'}^B|^2))}{\bar{I}_{d2d} + N_o(BW)}, \quad (2.20)$$

where \bar{I}_{d2d} denotes the interference averaged over all the W' requesting users in the D2MD group. The min operator is used as the D2MD will operate at a rate supported by the user with the worst SINR. Hence, the instantaneous rate achievable at the i^{th} user when served via D2MD in building B is given as:

$$R_B = BW \log_2(1 + \gamma_B). \quad (2.21)$$

2.7.1.1 Achievable Rate using OME

The building occupancies will vary depending on the time slot. Consequently, a user will either go for unicast or D2MD implying distinct performance in each slot. Let R_B^{uni} denote the unicast rates achievable for the users in building B , hence

$$R_B^{uni} = BW \log_2(1 + \gamma_B^{uni}), \quad (2.22)$$

The achievable rate per user for D2MD-U in the n^{th} slot is given as follows:

$$R_{D2MD-U}^n = \sum_{X^n} \sum_B \underbrace{\left(\frac{W'R_B + (x_B - W')R_B^{uni} + \sum_{D,D \neq B}^b R_D^{uni} x_D}{U} \right)}_{\mathcal{I}} \Omega_B^n, \quad (2.23)$$

for $X^n : \sum_B x_B = U, B, D \in \{1, 2, \dots, b\}$.

For an observed state X^n and building B in which the cache C is present with probability $\Omega_B^n = P[X^n = (x_1, x_2, \dots, x_b), C^B]$, the achievable rate for D2MD-U, denoted by \mathcal{I} , will depend on W' and the instantaneous rates derived in (2.21) and (2.22). This is because only W' users will receive the content with rate R_B . Rest of the users will receive the content at the building dependent unicast rates as given in (2.22). Further, the achievable rates are summed over all buildings in $\{1, 2, \dots, b\}$ and all the observed occupancy state vectors in the n^{th} slot.

2.7.1.2 Achievable Rate using EME

Unlike OME, EME provides a single averaged occupancy state vector for each time slot. The achievable rate per user for D2MD-U network in the n^{th} slot will be as follows:

$$\mathcal{R}_{D2MD-U}^n = \sum_{B \in \mathcal{S}_1} \left(\frac{W'R_B + (\hat{\lambda}_B^{n,U} - W')R_B^{uni} + \sum_{D,D \neq B}^b R_D^{uni} \hat{\lambda}_D^{n,U}}{U} \right) P_B^n + \phi \sum_{B \in \mathcal{S}_2} P_B^n, \quad (2.24)$$

where $\hat{\lambda}_B^{n,U}$ is the estimated expected occupancy of building B in n^{th} slot. If $\hat{\lambda}_B^{n,U} \geq 1$ only then a D2MD group will be formed in building B . The set $\{1, 2, \dots, b\}$ of buildings is divided into two subsets \mathcal{S}_1 and \mathcal{S}_2 based on whether D2MD group formation is supported or not. \mathcal{S}_1 contains those buildings which have $\hat{\lambda}_B^{n,U} \geq 1$ and $\mathcal{S}_2 = \{1, 2, \dots, b\} - \mathcal{S}_1$. For a given building $B \in \mathcal{S}_1$ in which the cache C is present with probability P_B^n , the achievable rate is similar to (2.23) with $x_B = \hat{\lambda}_B^{n,U}$. When the cache C is in building $B \in \mathcal{S}_2$, the achievable rate per user will be $\phi = \frac{\sum_B \hat{\lambda}_B^{n,U} R_B^{uni}}{U}$.

2.7.2 Achievable Rate in D2MD-M

In D2MD-M, similar to D2MD-U, there is a formation of a D2MD group consisting of W' users reusing one RB in UL phase. Let γ_{mul} be the instantaneous cellular multicast SNR whose expression is given below:

$$\gamma_{mul} = \frac{P_t^{DL} \left[\min \left(\{ |g_j^B|^2 : B \in \{1, \dots, b\}, j \in [1, x_B] \} \right) \right]}{N_o BW} \quad (2.25)$$

where $g_j^B \sim \mathcal{CN}(0, L_B^{-1})$ is the channel coefficient of links between j^{th} user in buildings B and BS. L_B is the pathloss between users in building B and BS. The i^{th} user in \mathcal{W} is included in \mathcal{W}' if $\gamma_B^i \geq \gamma_{mul}$.

However, $(U - W')$ users who are not part of the D2D group use the cellular multicast feature for content download from BS in the DL phase using one RB per cellular multicast group.

2.7.2.1 Achievable Rate using OME

Similar to D2MD-U, in D2MD-M, users not part of the D2MD group will receive the requested content via multicast. The achievable rate in cellular multicast can be written as follows:

$$R_{mul} = BW \log_2(1 + \gamma_{mul}). \quad (2.26)$$

However, for the case of D2MD the expression of instantaneous SNR of cellular multicast for remaining $(U - W')$ users will get modified as given follows:

$$\gamma_B^{mul} = \frac{P_t^{DL} \left[\min \left(\left\{ |g_1^B|^2, \dots, |g_{x_B - W'}^B|^2 \right\}, \left\{ |g_j^D|^2 : D \in \{1, \dots, b\}, j \in [1, x_D], D \neq B \right\} \right) \right]}{N_o BW}. \quad (2.27)$$

γ_B^{mul} will be utilized as cellular multicast SNR if the D2MD group is formed in building B . The instantaneous rate achievable per user for D2MD-M in the n^{th} slot is given as follows:

$$R_{D2MD-M}^n = \sum_{X^n} \sum_B \left(\frac{W' R_B + (U - W') R_B^{mul}}{U} \right) \Omega_B^n, \text{ for } X^n : \sum_B x_B = U, B \in \{1, 2, \dots, b\}. \quad (2.28)$$

where

$$R_B^{mul} = BW \log_2(1 + \gamma_B^{mul}) \quad (2.29)$$

Table 2.5: Simulation Parameters

Parameter	Value
Bandwidth of one RB	180 kHz
Size of content, Q	1 MB
Number of users in social group \mathcal{K}, K	10
BS transmission power, P_t^{DL}	46 dBm [36]
Device transmission power, P_t^{UL}	23 dBm [28]
Distance between BS and buildings \mathbf{H}, \mathbf{A} and \mathbf{S}	188 m, 330 m and 182 m respectively
Maximum distance of a requesting user from C in D2MD group, s	18 m
Path loss model for cellular mode	$[44.9 - 6.55 \log_{10}(h_{BS})] \log_{10}(d) + 34.46 + 5.83 \log_{10}(h_{BS}) + 23 \log_{10}(f_c/5) + 17.4 + 0.5d_{in} - 0.8h_{MS}$
Path loss model for D2D mode	$20 \log_{10}(d) + 46.4 + 20 \log_{10}(f_c/5) + WL + FL$
Carrier frequency	2.3 GHz

For an observed state X^n and building B in which the cache C is present with probability $\Omega_B^n = P[X^n = (x_1, x_2, \dots, x_b), C^B]$, the achievable rate for D2MD-M will depend on W' , (2.21) and (2.29). Further, the achievable rates are summed over all buildings in $\{1, 2, \dots, b\}$ and all the observed occupancy state vectors in the n^{th} slot.

2.7.2.2 Achievable Rate using EME

Similar to (2.24), the instantaneous rate achievable per user for D2MD-M in the n^{th} slot using EME will be as given as follows:

$$\mathcal{R}_{D2MD-M}^n = \sum_{B \in \mathcal{S}_1} \left(\frac{W' R_B + (U - W') R_B^{mul}}{U} \right) P_B^n + R_{mul} \sum_{B \in \mathcal{S}_2} P_B^n. \quad (2.30)$$

When the cache C is in building $B \in \mathcal{S}_2$, the achievable rate per user will be as given in (2.26).

2.8 Simulation Results and Discussion

This section compares the performance of D2MD-U and D2MD-M with the conventional cellular networks using unicast and cellular multicast schemes for the campus set-up with

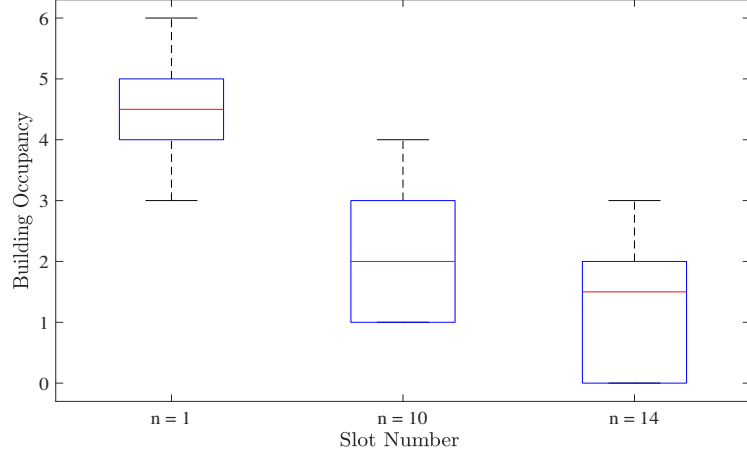


Figure 2.11: Box plots for occupancy of building **H** for $n = 1, 10, 14$ and $U = 6$.

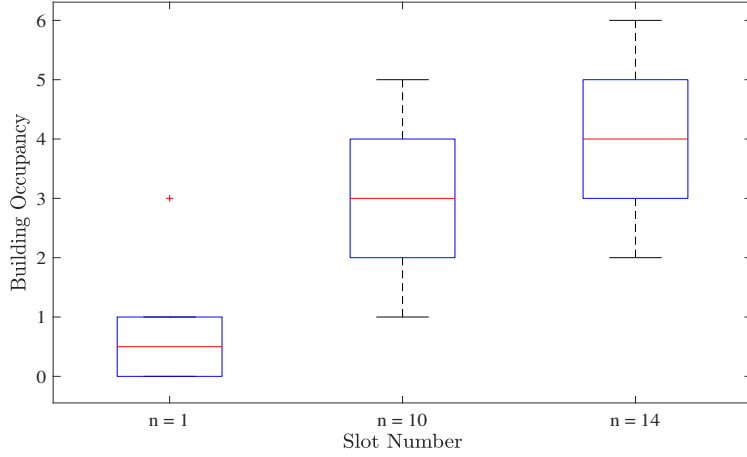


Figure 2.12: Box plots for occupancy of building **A** for $n = 1, 10, 14$ and $U = 6$.

$b = 3$ discussed in Section 2.4 It has been assumed that when the requesting users are in the same building as C , they are distributed uniformly within a radius of s m with respect to C . This distribution will hold even when the network performance is evaluated using EME as then the users in \mathcal{U} are distributed using time-dependent SPPP.

The simulation parameters are provided in Table 2.5. The bandwidth of each RB is 180 kHz, the carrier frequency f_c is taken as 2.3 GHz, and the size of multimedia content to be delivered is set as 1 MB. The transmission powers of BS, P_t^{DL} and mobile device, P_t^{UL} are set to 46 dBm and 23 dBm respectively. The path loss equations for cellular and D2D links [79] are also given in Table 2.5 where base station height, $h_{BS} = 25$ m, mobile device height, $h_{MS} = 3n_{Fl} + 1.5$ m and n_{Fl} is the floor index (for ground floor $n_{Fl} = 1$). d_{in} is the distance from the indoor transmitter to the wall, d_{out} is the distance between the point on the wall that is nearest to the indoor transmitter and base station. d is the distance between each pair of transmitter and receiver in meters. $WL = 5n_w$ (for soft walls) or $12n_w$ (for hard walls) and n_w are the number

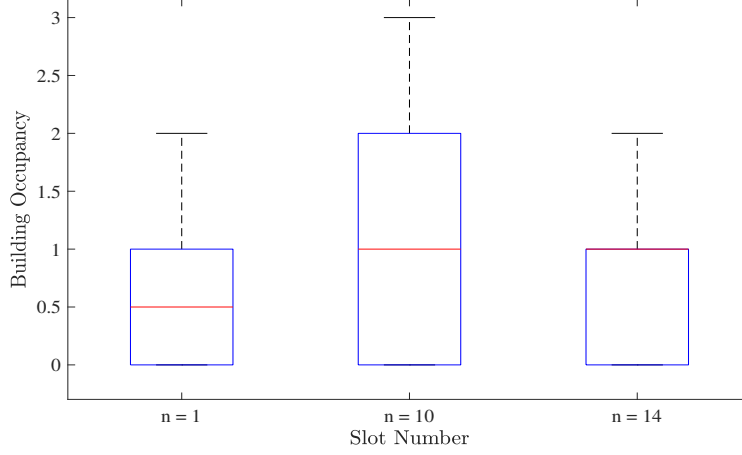


Figure 2.13: Box plots for occupancy of building **S** for $n = 1, 10, 14$ and $U = 6$.

of walls penetrated by the signal. The campus buildings in our case have hard wall partitions; thus, $WL = 12n_w$ has been considered. The value of n_w depends on the distance between the requesting user and C and the building layout. $FL = 17 + 4(n_{fl} - 1)$ is the floor loss and n_{fl} is the number of floors penetrated by the signal. The layout for all the three buildings has been given in [79]. The small-scale Rayleigh fading channel has been kept static over one subframe of the TDD frame structure. The achievable rates have been averaged over 10^4 iterations of the Rayleigh fading channel, and in the case of OME, the achievable rates are also averaged over $\binom{K}{U}$ unique sets of U users. For any value of the average achievable rate \bar{R} per user, the number of time frames (NoTF) is calculated as follows:

$$\text{NoTF} = \left(\frac{Q \times 1000}{\bar{R} \times 5} \right) \quad (2.31)$$

where Q/\bar{R} is the average time taken for content delivery to a requesting user. For determining NoTF, the time taken has been divided by 5 ms because effectively each user will have 5 ms of each time frame to receive the content. Figs. 2.11-2.13 demonstrate the impact of temporal behavior of the users on the building occupancies for $U = 6$ and time slots, $n = 1, 10$ and 14 with the help of box plots. The box plots illustrate the spread of the building occupancies, and the red line represents the median value. It can be observed that as the day progresses, the occupancy of **H** decreases whereas the occupancy of **A** increases. This is because more users (or students in the present case) move from a residential hostel to an academic building. However, building **S** which hosts the dining center has higher occupancy during lunch hours, i.e., 13:30- 14:00 hrs. Similar variations in building occupancies were observed for higher values of U .

Figs. 2.14 and 2.15 show the plot of NoTF required to deliver 1 MB content to U users with conventional cellular unicast, cellular multicast, D2MD-U, and D2MD-M evaluated using

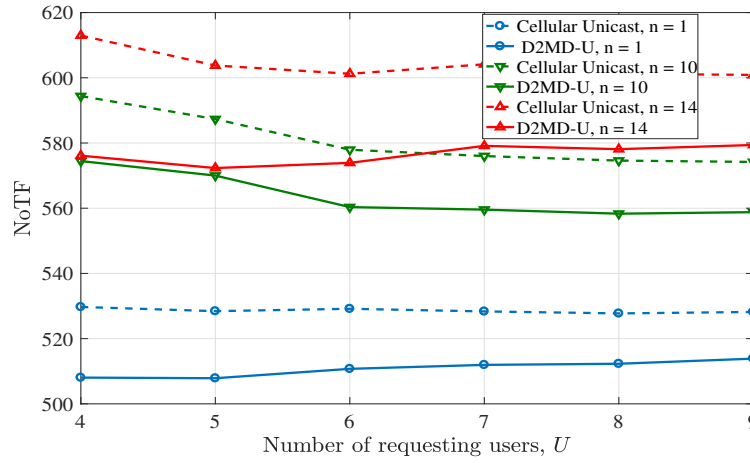


Figure 2.14: Comparison of NoTF required for D2MD-U and cellular unicast when $s = 18$ m.

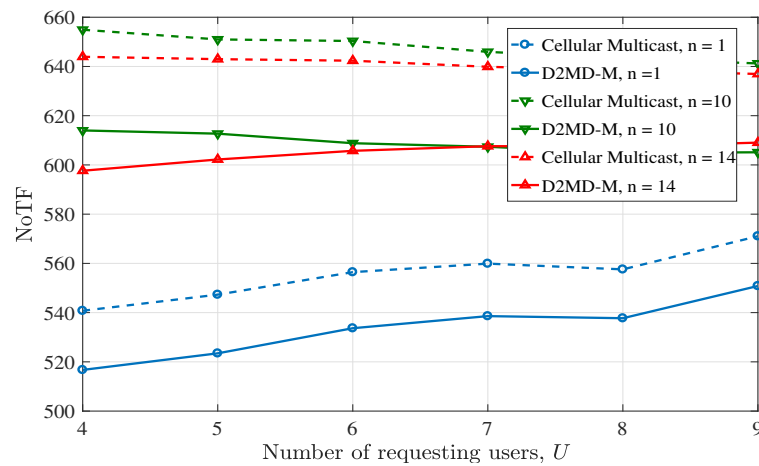


Figure 2.15: Comparison of NoTF required for D2MD-M and cellular multicast when $s = 18$ m.

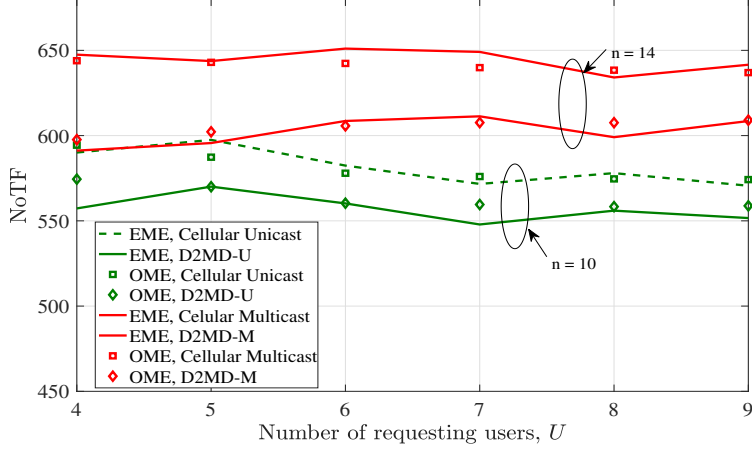


Figure 2.16: Comparison of OME and EME.

OME for $n = 1, 10$ and 14 . The performance of D2MD networks depends on the following factors: 1) Floor layout of each building 2) Cellular link length from each building to BS, 3) Probability of C in each building, 4) Observed states at each value of U and the probability of each observed state, and 5) Time of the day (temporal behavior). It can be observed that for a fixed U , NoTF is higher for D2MD-U as compared to D2MD-M because supporting the rate of worst channel user deteriorates D2MD-M's performance. Further, in case of D2MD-U and cellular unicast while progressing from $n = 1$ to $n = 14$ NoTF increases. This is because more requesting users start occupying the building **A** which has longer cellular link length as compared to building **H** and **S**. Hence, cellular unicast and D2MD-U performance deteriorates. On the other hand, NoTF at $n = 14$ is lower than the NoTF at $n = 10$ for D2MD-M and cellular multicast. This is because at $n = 14$ there are some observed occupancy states for which $x_{\mathbf{H}} = 0$ and this changes the mean of the distribution of the minimum of channel gains¹². Consequently, γ_{mul} in $n = 14$ is less worse than in $n = 10$. Further, it can be seen that for all the time slots, D2MD-U and D2MD-M are performing better than the conventional cellular unicast and cellular multicast, respectively.

Fig. 2.16 presents the plot of NoTF for D2MD-U and D2MD-M evaluated using EME and compared to the performance evaluated using OME. The probability of cache C in each slot and $n = 1, 10, 14$, obtained from its past location information, is given in Table 2.6. As mentioned before, EME extracts the estimated expected building occupancies. From Fig. 2.16 it can be observed that the NoTF values obtained using EME are in good agreement with the values obtained using OME. The slight deviation of EME plots from the OME plots is due to rounding off of the non-integral estimated expected building occupancies.

¹²For more details, interested readers may refer to the Appendix A.

Table 2.6: Probability of C in each building

Slot No.	H	A	S
$n = 1$	0.8	0.04	0.16
$n = 10$	0.08	0.04	0.88
$n = 14$	0.4	0.6	0

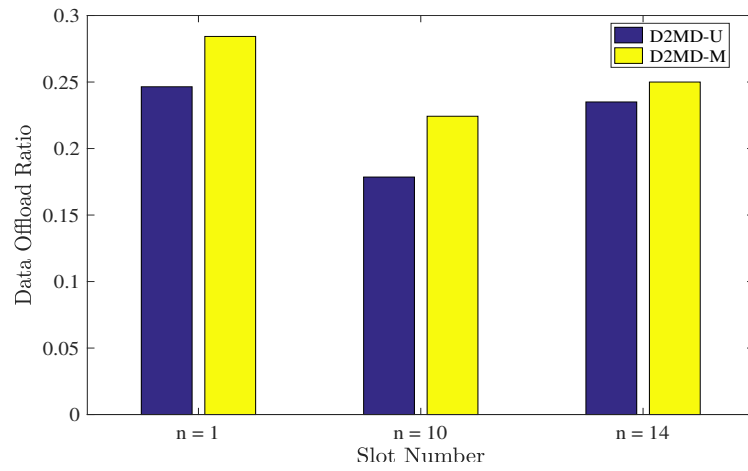


Figure 2.17: Data offload ratio of D2MD-U and D2MD-M at $U = 7$.

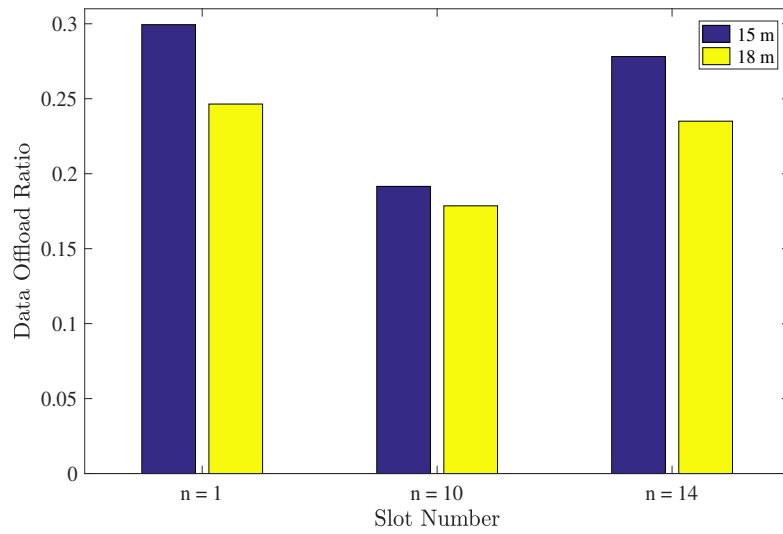


Figure 2.18: Impact of s on D2MD-U at $U = 7$.

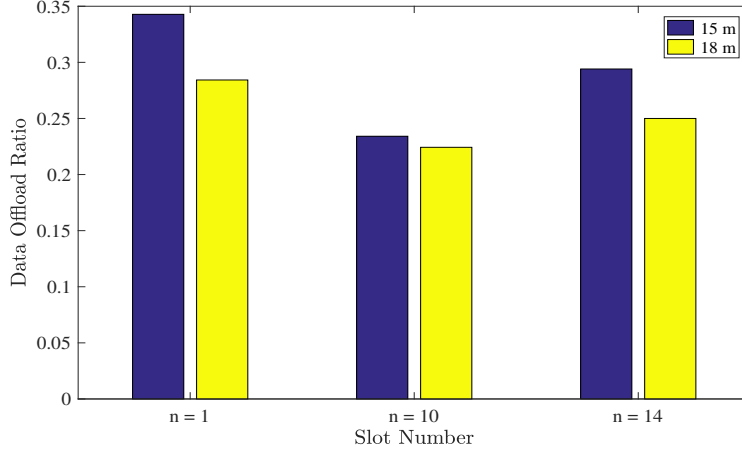


Figure 2.19: Impact of s on D2MD-M.

Fig. 2.17 shows the data offloading ratio for D2MD-U and D2MD-M at $U = 7$. The data offload ratio is the ratio of the amount of data offloaded to the maximum amount of data that can be offloaded. At $U = 7$, the maximum amount of data that can be offloaded is $1 \text{ MB} \times 7 = 7 \text{ MB}$. It may be noted that unlike the conventional cellular unicast and multicast schemes, D2MD-M and D2MD-U inherently help in data offloading.

It can be seen from Fig. 2.17 that the data offload ratio of D2MD-M is higher than D2MD-U. The criteria for D2MD group formation in D2MD-M involves a threshold value of γ_{mul} which is lower than γ_B^{mi} and results in D2D groups where a higher number of users get their content via D2D links. Further, the fluctuations in the data offload ratio can be observed with the change in the time of the day. Peak data offloading of 28% is observed in $n = 1$ because this slot witnesses the highest probability of D2D.

The impact of the spatial spread of the D2MD group within a building on the data offload ratio was studied. The simulations were carried out by varying the maximum distance of a requesting user from C in the D2MD group. The number of requesting users, $U = 7$, and users were assumed to be within a radius of $s = 15 \text{ m}$, 18 m from C . Figs. 2.18 and 2.19 demonstrate the impact of s on D2MD-U and D2MD-M respectively. The data offloading capability corresponding to 15 m for both D2MD-U and D2MD-M is higher than the values at 18 m where a maximum of 35% data offloading (i.e., 0.35 data offloading ratio) can be achieved at 15 m .

2.9 Conclusion

This work analyzed the D2D+cellular framework for an indoor propagation environment by making use of real world user location data. The performance of the proposed framework is

evaluated in terms of spectral efficiency and amount of energy saving. The real-world data helped to exploit the spatio-temporal correlations present in user mobility. Subsequently, a realistic performance evaluation of two D2MD networks, namely D2MD-U and D2MD-M utilizing the joint spatio-temporal behavior of the users, is presented. To obtain the joint spatio-temporal behavior of the users, two novel methods, OME and EME, are proposed. In OME, the joint spatio-temporal behavior is extracted from the past location information of the requesting users, whereas in EME, the location information of the requesting users is unknown, and the expected occupancies were estimated using a training dataset of users. The results demonstrated the impact of spatiotemporally correlated user mobility on the data offloading capability of D2MD networks.

Cache Selection in D2MD Networks

In the previous chapter, we observed that the user mobility results in a diurnal variation in the performance of D2D and D2MD networks. Motivated by the above and the need to offload traffic from the cellular network, we propose a user spatio-temporal behavior aware cache selection framework for D2MD networks. This framework minimizes the number of selected caches while achieving a desired user load on the cellular network. The selection of the minimum number of caches is formulated as an optimization problem. Further, minimizing the subset of users selected as caches can alleviate the burden of content caching on the cellular network. However, due to the unexpected occurrences of network congestion, frequent optimizations will be required [86]. Assuming that the popular content needs to be updated on a daily basis¹, an optimization is carried out at the beginning of each day to select the caches optimal on that day. In a realistic setting, the core network can equip the caches with the popular multimedia content, and on a daily basis, update the content based on demand and popularity. For instance, a viral news article can be cached at selected users by the core network, which can be downloaded by other requesting or non-caching users sometime later in the day. As the day progresses, more optimizations will be performed to manage the sudden variation in the user load constraint preset by the core network. However, frequent optimizations should not discard the previously selected caches. Otherwise, the cellular network will be burdened with content caching at the newly selected caching users. This will lead to a further increase in the caching load on the cellular network. In the formulated optimization problem, we take into consideration the above constraint and thus do not discard the previously selected caches. The problem is combinatorial in nature, and the complexity of the problem increases with an increase in the number of users. Hence, a greedy algorithm for cache selection is proposed that exploits the problem structure to reduce the search space. It is shown that the complexity of

¹The cached content is flushed out at the end of the day, and optimization is again carried out at the beginning of the next day[87].

the greedy algorithm for K users is $\mathcal{O}(K^2)$ whereas the complexity of the widely applicable exhaustive search is $\mathcal{O}(2^{(K-1)})$. This has been discussed in detail in Section 3.2. To the best of our knowledge, the proposed work is the first of its kind that finds an optimal set of caches for dynamic D2MD networks that alleviates the caching load on the cellular network.

Leveraging the fact that a user prefers or requests a few content files more [43], the caching can be performed for a group of users who share common interests, i.e., the users are part of a social group and tend to be in each others proximity [88]. Consequently, the users in this social group may generate requests for common content files and it will be easier to accommodate a few commonly requested files at each selected cache. Hence, the proposed work also performs social tagging, i.e., a set of caches is selected from a social group and is assigned the task to serve the other members of the social group via D2MD communication. Base stations (BSs) that cover the spatial spread of the group of users can be informed a priori about the selection decision for that group. Hence, if a user of a particular social group requests for popular content file, BS will associate the user to one of the caches tagged to the user's social group. The above will be useful in a network-assisted D2D peer discovery scenario as this will lead to fast peer (or cache) discovery since the search space for establishing a D2D connection is now limited to the optimal set of caches [88].

In order to capture the spatio-temporal behavior of the users, the proposed framework requires the joint mobility pattern of users. Consequently, an inhomogeneous discrete-time Markov chain [89, 90] is presented that utilizes the real-world location information to model the joint mobility pattern of users, wherein the time of the day is chosen as the line of reference. The location data was gathered from 9:00 to 19:00 hrs for one semester (only working days) in a campus set-up where the user locations were broadly categorized into three buildings. Since the considered users are students, the location samples from a semester were collected to fully capture the mobility information. The duration of 9:00 to 19:00 hrs was divided into 20 slots of 30 minutes each. The gathered real-world data is used as the training dataset to determine the transition probabilities of the Markov chain. The spatio-temporal correlations of each group of users are derived from the joint mobility pattern. This is explained in detail in Section 3.1.

Moreover, as mentioned before, the prior works [12, 14] assumed the transmission rates to be the same for all D2D pairs. However, generally, the transmission rates depend on the channel conditions of the users, which may be present in different spatial locations. As 90% of the time, users stay in an indoor environment [91], we use the indoor path loss models to evaluate the varying transmission rates.

3.0.1 Main Contributions

In this work, a novel cache selection framework is proposed that exploits the spatio-temporal behavior of the users to facilitate D2MD communication in a cellular network. The major contributions of this chapter are as follows:

- The proposed framework minimizes the number of caches required to achieve a desired user load on the cellular network, thus alleviating the caching load on the cellular network.
- The selection of the minimum number of caches is formulated as an optimization problem. However, frequent optimizations due to the unexpected occurrences of network congestion result in frequent changes in selected caches. Consequently, the proposed framework does not discard the previously selected caches which further reduces the caching load.
- The formulated problem has an exponential search space. Hence, a greedy algorithm for cache selection with complexity $\mathcal{O}(K^2)$ is proposed that exploits the problem structure.
- An inhomogeneous discrete-time Markov chain model based on real-world location information² of users is presented to predict the spatio-temporal behavior.
- The proposed framework is compared to mobility-unaware cache selection. Our work exhibits that our framework outperforms mobility-unaware cache selection in terms of achievable sum-rate of non-caching users and user load on core network. It is shown that to achieve the desired user load with mobility-unaware cache selection the caching load on the cellular network should be increased.

3.0.2 Terminologies

Below we have defined the key terms used throughout this chapter:

- Cache: Cache or caching user refers to a user in the network that temporarily stores a multimedia file or content.
- Cache Selection: Cache selection refers to the process of selecting a set of caches where the multimedia content may be stored.

²The dataset is available at https://www.iiitd.edu.in/~wirocomm/resources/Social_Group_data.rar.

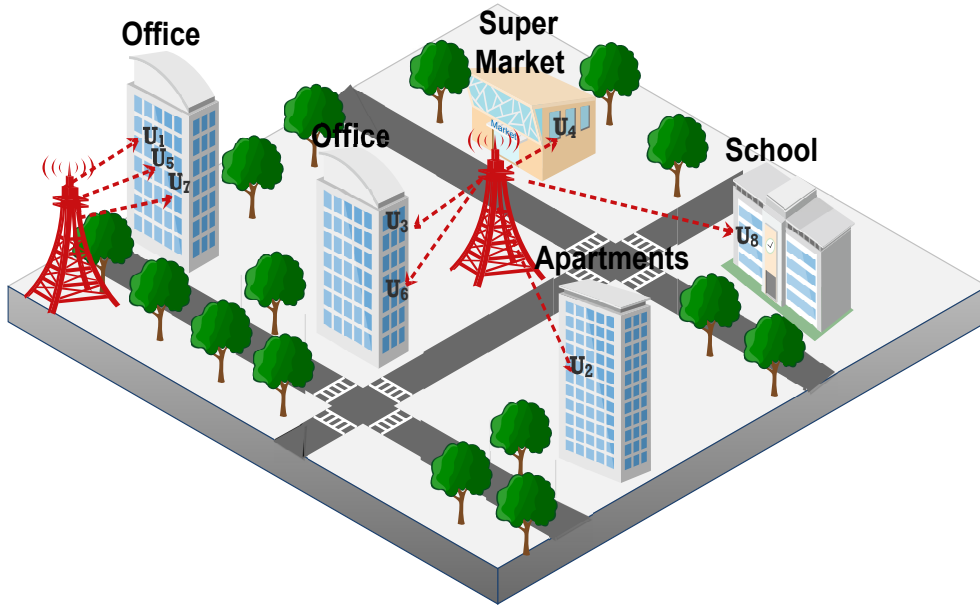


Figure 3.1: Illustration of $K = 8$ users belonging to a social group spread across a geographical area consisting of two BSs.

- User Load: It refers to the number of users that are served by the cellular network directly.
- Mobility Unaware Cache Selection: It is the cache selection process where the user mobility is not taken into account.

3.1 System Model

The system model considers a total of K users in a social group denoted by set \mathcal{K} and each user is represented as U_i where $i \in [1, K]$. These users are distributed across m buildings namely $B_j, j \in [1, m]$. This scenario can depict either a single cell (BS) or a multi-cell (BSs) network depending on the geographical span of the user mobility. In this chapter, the terms users supported by the cellular network and user load are used interchangeably. Fig. 3.1 demonstrates a scenario where $K = 8$ users, belonging to a social group and accessing a popular multimedia file via unicast cellular links, are spread across a geographical area consisting of five buildings, i.e., $m = 5$ at a specific time of the day. Assuming the users supported by the cellular network that request the same popular file is equivalent to the user load, the user load for the above case will be 8. However, a more efficient approach would be to cache the popular multimedia content at some specific user (or users) such that the user load is less than 8.

In the proposed work, a set of caches is selected for a social group so that each cache is willing to share the cached content with others via D2D multicast. It is assumed that the

users request the content at any time of the day but only once. Let L_p be the desired load of requesting users on the core network averaged over a day and the mobility pattern of the users in \mathcal{H} . p is the instant of optimization and $p \in \{1, 2, 3 \dots\}$. The set of caches has to be optimally selected to reduce the average user load to L_p . As illustrated in Fig. 3.1, it is considered that the members of a social group request for a popular multimedia file at the same time. In a realistic scenario, the number of users requesting the same content at a given time will be less than the total users in the group. However, the optimal selection assuming all users are requesting the same content ensures that an average user load of L_p or below is achieved irrespective of the number of requesting users. The information of the set of cache selected for the network will be shared among the BSs that span the m buildings by a central controller where the central controller is the entity where the cache selection decision is being taken.

Let us denote the average signal-to-interference and noise ratio (SINR) value as $SINR_{ik}^j$ for users U_i and U_k present in building B_j . The SINR values will be characterized by the layout of building B_j , users' relative positions, and interference. The interference will be due to the cellular user³ C_R whose uplink (UL) RB is being reused by a D2D multicast group. $SINR_{ik}^j$ can be defined as follows:

$$SINR_{ik}^j = \frac{P_t \Gamma}{I_{avg} + N_o BW}, \quad (3.1)$$

where P_t is the transmit power of U_i and $\Gamma = \mathbb{E} |h_{ik}|^2$. $h_{ik} \sim \mathcal{CN}(0, \beta_{ik}^{-1})$, β_{ik}^{-1} is the pathloss between U_i and U_k and $|h_{ik}|^2$ is the channel gain. Further, N_o is the noise spectral density and BW is the bandwidth of one RB. $I_{avg} = \mathbb{E}_d [P_t |g|^2]$ is the average interference at U_k due to transmission from C_R located at a distance d from user k . $g \sim \mathcal{CN}(0, \beta_{C_R,k}^{-1})$ where $\beta_{C_R,k}$ is the path loss between C_R and U_k . The pathloss values have been determined using WINNER II models [79]. Let us assume that the cellular network guarantees an average signal-to-noise ratio (SNR), γ_{cell} to its users in order to maintain a pre-defined quality of service (QoS). Now, user U_i which is present in building B_j will be receiving content from U_k via D2MD only if $SINR_{ik}^j > \gamma_{cell}$, i.e., QoS obtained through D2MD will be better than that obtained through cellular. Hence, in other words, we can state that if U_i belongs to the feasibility set of U_k , U_i can be served by U_k via D2MD. The users who are not served by D2MD will receive the content from the cellular network. It may be noted that each building can have a maximum of K feasibility sets depending on the geographical distribution of the K users in them.

For instance, the role of γ_{cell} in the generation of these feasibility sets has been demonstrated for a real campus set-up where $m = 3$ and $K = 5$. The matrix $SINR^j$ contains the average pairwise SINRs of $K = 5$ for the j^{th} building based on $P_t = 23$ dBm, $N_o = -173$ dBm/Hz, $BW = 180$ kHz and d which is uniformly distributed in the range (10, 250) m. The SINR matrices are

³It is assumed that the RBs assigned to a cellular user in UL is reused by one D2D group.

Table 3.1: Feasibility sets for each building

γ_{cell}	B_1	B_2	B_3
20 dB	$U_1 - \{U_5\}$	$U_1 - \{U_2, U_3, U_4, U_5\}$	$U_1 - \{U_2, U_3, U_5\}$
	$U_2 - \{U_5\}$	$U_2 - \{U_1, U_3, U_4, U_5\}$	$U_2 - \{U_1, U_3, U_4, U_5\}$
	$U_3 - \{U_4, U_5\}$	$U_3 - \{U_1, U_2, U_4, U_5\}$	$U_3 - \{U_1, U_2, U_4, U_5\}$
	$U_4 - \{U_3, U_5\}$	$U_4 - \{U_1, U_2, U_3, U_5\}$	$U_4 - \{U_2, U_3, U_5\}$
	$U_5 - \{U_1, U_2, U_3, U_4\}$	$U_5 - \{U_1, U_2, U_3, U_4\}$	$U_5 - \{U_1, U_2, U_3, U_4\}$
25 dB	$U_1 - \{U_5\}$	$U_1 - \{U_2, U_3, U_4, U_5\}$	$U_1 - \{U_2, U_5\}$
	$U_2 - \{\}$	$U_2 - \{U_1, U_3, U_4, U_5\}$	$U_2 - \{U_1, U_3, U_5\}$
	$U_3 - \{U_5\}$	$U_3 - \{U_1, U_2, U_4, U_5\}$	$U_3 - \{U_2, U_4\}$
	$U_4 - \{U_5\}$	$U_4 - \{U_1, U_2, U_3, U_5\}$	$U_4 - \{U_3, U_5\}$
	$U_5 - \{U_1, U_3, U_4\}$	$U_5 - \{U_1, U_2, U_3, U_4\}$	$U_5 - \{U_1, U_2, U_4\}$

given below where $SINR_{ik}^j$ is the $(i, k)^{th}$ element of $SINR^j$:

$$SINR^1 = \begin{bmatrix} \text{NA} & 16.72 & 13.23 & 19.35 & 25.03 \\ 16.72 & \text{NA} & 15.46 & 15.36 & 22.95 \\ 13.23 & 15.46 & \text{NA} & 20.16 & 39.92 \\ 19.35 & 15.36 & 20.16 & \text{NA} & 34.34 \\ 25.03 & 22.95 & 39.92 & 34.34 & \text{NA} \end{bmatrix}, \quad (3.2)$$

$$SINR^2 = \begin{bmatrix} \text{NA} & 31.27 & 33.06 & 27.98 & 36.66 \\ 31.27 & \text{NA} & 25.96 & 40.78 & 29.45 \\ 33.06 & 25.96 & \text{NA} & 28.21 & 33.23 \\ 27.98 & 40.78 & 28.21 & \text{NA} & 40.70 \\ 36.66 & 29.45 & 33.23 & 40.70 & \text{NA} \end{bmatrix}, \quad (3.3)$$

$$SINR^3 = \begin{bmatrix} \text{NA} & 33 & 24 & 17.73 & 31.72 \\ 33 & \text{NA} & 26.35 & 24.74 & 27.50 \\ 24 & 26.35 & \text{NA} & 29.77 & 20.44 \\ 17.73 & 24.74 & 29.77 & \text{NA} & 29.29 \\ 31.72 & 27.50 & 20.44 & 29.29 & \text{NA} \end{bmatrix}, \quad (3.4)$$

where NA is used to denote the SINR value for the link from user U_i to itself. For instance, the SINR corresponding to the link between U_1 and U_5 in building B_1 is $SINR_{15}^1 = SINR_{51}^1 = 25.03$ dB. Using (3.2)-(3.4), the feasibility sets observed for each building at $\gamma_{cell} = 20$ dB, 25 dB for the campus set-up are presented in Table 3.1. It can be seen with the increase in γ_{cell} , the number of elements in feasibility sets reduces. This is because as the threshold γ_{cell} increases, some pairwise SINR values will fail to meet the increased threshold.

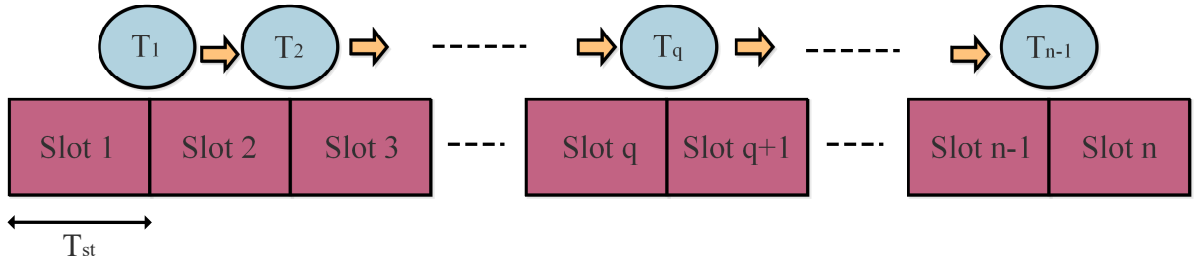


Figure 3.2: Illustration of dependency of transition matrices on slot boundaries.

3.1.1 Inhomogeneous Discrete-time Markov Chain

As stated before, for optimal cache selection, one of the key elements is to deduce the mobility pattern of the users. However, as the proposed work involves D2MD where multiple users are involved, there is a need to analyze the joint mobility pattern of the users. This will help in analyzing the interaction of multiple users at the same time. We present a discrete-time Markov chain to model the joint mobility pattern of the users. The state of the network is denoted by a vector of size $1 \times K$ where element i of the state vector is the building visited by U_i and can have m possible values, i.e., B_1, \dots, B_m . Due to the mobility of the users, the network will transit from one state to another. The network is assumed to stay in any state ‘ a ’ for time μT_{st} , where T_{st} is the minimum time the network will stay in any state ‘ a ’ and μ is an integer that depends on the user mobility. As the mobility pattern is modeled using a Markov chain, the next state will only depend on the current state of the network. A complete day is divided into n slots of T_{st} duration each. The transition probability from one state to another has been derived from the observed joint mobility pattern of the users. There are inherent spatio-temporal correlations in the observed joint mobility pattern, i.e., given a time slot, certain buildings will be more preferred by users than others. Since the building preferences are time-dependent, each slot will observe a different set of states, and each slot boundary will have a different transition matrix. Consequently, the Markov chain is inhomogeneous [90]. The transition matrix corresponding to q^{th} slot boundary is denoted as $T_q, q \in [1, n-1]$ and $(n-1)$ are the number of slot boundaries as illustrated in Fig. 3.2. Further, each element $p_{a,b}$ in a transition matrix denotes the transition probability from state ‘ a ’ in q^{th} slot to ‘ b ’ in $(q+1)^{th}$ slot. The joint location information must be gathered for a large number of days to capture mobility information completely. The state vectors are obtained for each time slot. For instance, if there are K users in the network and m buildings, it means in a given slot there are m^K possible state vectors per slot. However, because of the regularity in the human mobility pattern the set of observed state vectors in q^{th} slot will be considerably small and will depend on the time of the day. Let us denote the observed state vectors in the q^{th} slot as X_q . To better illustrate the user proximity with the change in state vectors, each observed state vector can be represented as m contact graphs where each graph corresponds to each building. The connection between one

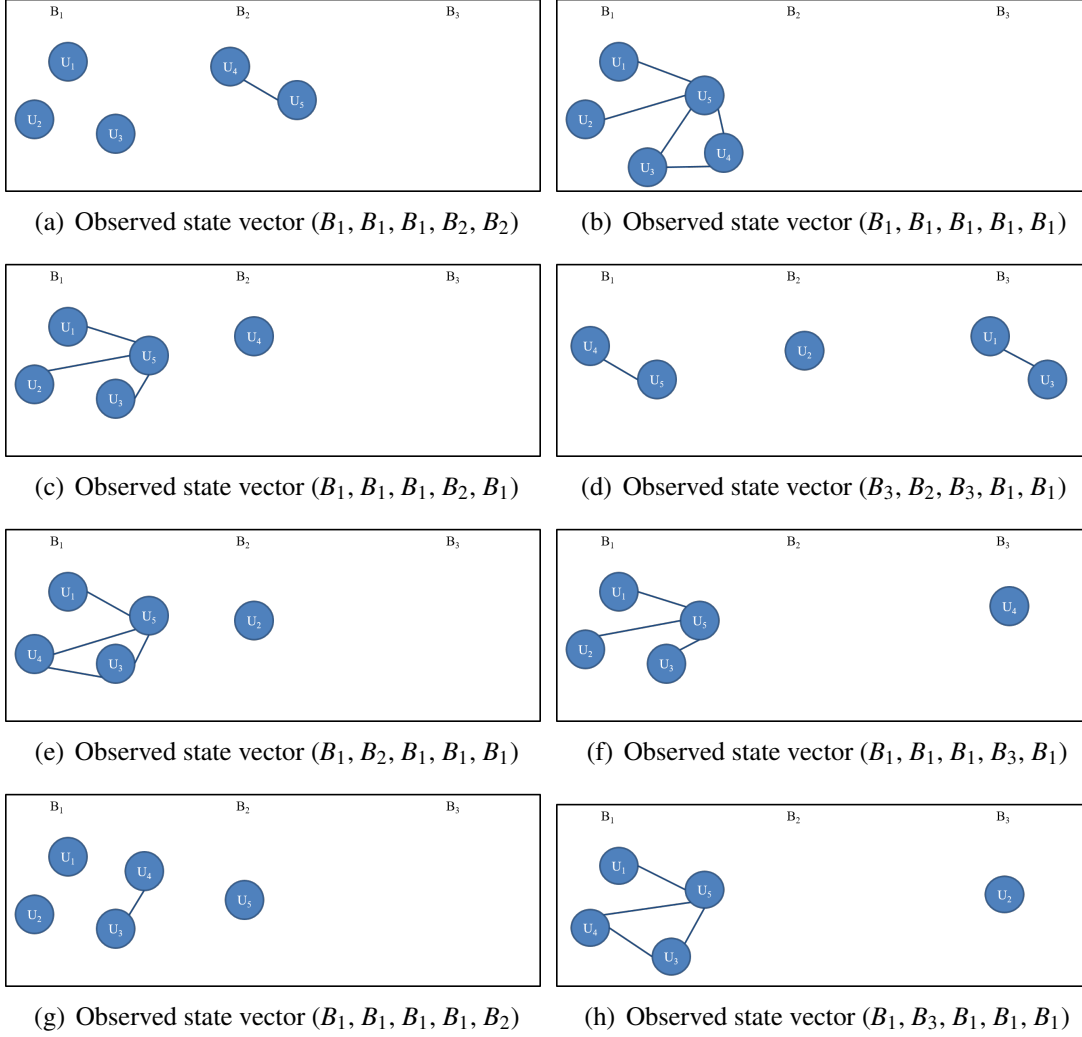


Figure 3.3: Contact graphs for illustrating the user proximity in a real campus set-up where $K = 5$, $m = 3$.

user to another will be decided by observing its feasibility set in that building. If user U_i lies in the feasibility set of U_k for building B_j there will be an edge in the graph of B_j connecting nodes U_i and U_k . These graphs will be dynamic in nature depending on the time of the day. Fig. 3.3 plots the graphs corresponding to the unique observed state vectors for slot 1 with $T_{st} = 30$ minutes, $K = 5$, $m = 3$ with $\gamma_{cell} = 20$ dB for the real-world campus set-up considered in this work. It can be seen that $|X_1| = 8$, which is significantly less than $3^5 = 243$. Due to the small value of $|X_n|$ for any social group of users, the number of transitions possible at a slot boundary will also be low. As the states in the Markov chain increase, complexity increases. Therefore, to reduce the complexity, we have restricted the size of transition matrix T_q to $M_q \times M_q$ where M_q is the number of states observed in $X_q \cup X_{q+1}$ [92]. The complexity can be reduced further by the methods suggested in [90], which is out of the scope of this work. Further, it may be

noted that certain entries in the transition matrices will be zero due to the spatio-temporal preferences of the users.

Algorithm 1 describes the steps for constructing the transition matrix T_q . At first, the observed state vectors in the q^{th} and $(q+1)^{th}$ slots are recorded. The state vectors in $X_q \cup X_{q+1}$ are the observable states of the Markov chain at the q^{th} boundary. This is followed by an evaluation of transition probabilities from one state to another. The steps are repeated for all the slot boundaries.

Algorithm 1: Constructing Transition Matrix, T_q

- 1 **Step 1:** First record the observed sets X_q and X_{q+1} .
 - 2 **Step 2:** Let $Y_q = (X_q \cup X_{q+1})$ denote the set of states at q^{th} slot boundary and $M_q = |Y_q|$.
 - 3 **Step 3:** Determine the transition probability of each state a to b , i.e., $p_{a,b} \forall a, b \in Y_q$.
These transition probabilities will constitute T_q .
-

3.2 Problem Formulation and Solution

Once the transition matrices have been constructed, the next task is the optimal selection of a set of caches taking into account the spatio-temporal behavior of the users. The optimal selection minimizes the number of selected caches subject to user load L_p or below on the core cellular network and must not discard the previous cache selections. Let v_0 be the state of the network at the instant of optimization and $v = (v_1, v_2, \dots, v_r, \dots, v_{n_0})$ be the observable sequence of states that occurs after T_{st} minutes of v_0 , where a state vector in v can be present more than once and n_0 is the number of slots remaining that day (i.e., future time slots). For instance, when the optimization is carried out in slot 1, v_0 is the state of network in slot 1, $n_0 = (n-1)$ and all observable sequences v can be obtained from Y_1, \dots, Y_{n-1} . Let $\rho(v|v_0)$ denote the probability of sequence v given v_0 has occurred and can be given as:

$$\begin{aligned}
\rho(v|v_0) &= \rho[(v_1, v_2, \dots, v_{n_0})|v_0], \\
&= P[v_1|v_0]P[v_2|(v_0, v_1)] \dots P[v_r|(v_0, v_1, v_{r-1})] \\
&\dots P[v_{n_0}|(v_0, v_1, v_{n_0-1})],
\end{aligned} \tag{3.5}$$

where $P[v_r|(v_0, v_1, \dots, v_{r-1})]$ is the probability of state v_r given sequence $(v_0, v_1, \dots, v_{r-1})$. As Markov chain model is assumed, $\rho(v|v_0)$ can be written as follows:

$$\begin{aligned}\rho(v|v_0) &= P[v_1|v_0]P[v_2|v_1]\dots P[v_{n_0}|(v_{n_0-1})], \\ &= \prod_{r=1}^{n_0} p_{v_{r-1}, v_r}.\end{aligned}\quad (3.6)$$

The optimization problem formulated to minimize the number of selected caches at p^{th} ($p \in \{1, 2, 3 \dots\}$) instant of optimization as follows:

$$\min_{C_p} |C_p|, \quad (3.7a)$$

$$\text{subject to} \quad \mathbb{E}_v \left[\mathbb{E}_{v_l} \left[(K - |C_p|) - U_{C_p}^{v_l} \right] \right] \leq L_p, \quad (3.7b)$$

$$C_{p-1}^* \subseteq C_p, C_0 = \emptyset \quad (3.7c)$$

$$|C_p| \leq \lfloor K/2 \rfloor \quad (3.7d)$$

where C_p is the set of selected caches and $|C_p|$ denotes cardinality of C_p . Let state $v_l \in \{v_0\} \cup v$ where l ranges from 0 to n_0 . $U_{C_p}^{v_l}$ are the non-caching users in v_l state served by users in C_p via D2D multicast. $U_{C_p}^{v_l}$ will be determined by the feasibility sets of the users in C_p . For $K = 5$, $C_p = \{U_1, U_2\}$, $v_l = (B_1, B_1, B_1, B_1, B_1)$ and feasibility sets as given in Table 3.1, $U_{C_p}^{v_l} = 1$. L.H.S of (3.7b) is the expected user load when set C_p is selected. As the content request at the cache can be generated at any time of the given day, while selecting the set of caches, we are averaging over the user load corresponding to current time slot and future time slots. Given a observable sequence v , first the user load is averaged over each state $v_l \in \{v_0\} \cup v$. Then, the user load is averaged over all the observable sequences. Constraint (3.7b) assures that the expected user load is less than L_p on the core network. As mentioned previously, frequent optimizations are required in the network to tackle with the sudden occurrences of network congestion. In such scenarios, to assure that the previously selected caches are not discarded, a constraint needs to be applied. In other words, for a given K , the optimal solution with cardinality, let us say, $|C_{p-1}^*|$ at desired user load L_{p-1} must be a subset of optimal solution with cardinality $|C_p^*|$ at desired user load L_p where $L_{p-1} > L_p$. Constraint (3.7c) takes care of the above requirement. It may be noted that the cardinality of optimal solution, C_p^* cannot go beyond $\lfloor K/2 \rfloor$ because at a given time no more than $\lfloor K/2 \rfloor$ D2MD groups can exist and has been accounted for using (3.7d). After slot 1 of the day, more optimizations will only be required when there is a decrease in the load constraint, because only then more caches will have to be selected.

The problem in (3.7a)-(3.7d) qualifies to be a combinatorial problem. The optimization will be done in two stages: (1) Stage 1: optimization at the beginning of each day, i.e., $p =$

1, (2) Stage 2: optimization due to sudden network congestion during the day, i.e., $p > 1$. For stage 1, constraint (3.7c) will not hold as there will be no previously selected caches at the beginning of the day. One approach to solve the optimization problem is to perform an exhaustive search. However, exhaustive search has a complexity of $\mathcal{O}(2^{(K-1)})$ (see Appendix B) due to the exponential search space. As a consequence, in the proposed work, a greedy algorithm for cache selection is proposed that exploits the problem structure and reduces the search space.

3.2.1 Greedy Algorithm for Cache Selection

As the cardinality of the selected set of caches needs to be minimized such that the constraints are met, the search will start from the lowest cardinality, i.e., 1. If the load constraint (3.7b) is not met at a given cardinality of set C_p then the cardinality of the set C_p has to be increased to search for the optimal solution. Exploiting this behavior, we have proposed a greedy algorithm for cache selection that takes into account only those candidate sets with cardinality c which have \mathcal{Q} , with cardinality $c - 1$, as a subset. \mathcal{Q} is the set that fails to be optimal but has a minimum expected user load among the candidate sets at cardinality $c - 1$. The greedy algorithm has a search space of the order $\mathcal{O}(K^2)$ (see Appendix B), whereas the exhaustive search is of the order $\mathcal{O}(2^{(K-1)})$. Hence, the greedy algorithm reduces the search space of the formulated problem.

3.2.2 Illustration of Exhaustive Search and Greedy Algorithm with Real Data

Algorithm 2: Stage 1: Exhaustive Search

```

1 Step 1: Initialize  $|C_1| = 0$ .
2 Step 2: Increment value of  $|C_1|$ .
3 if  $|C_1| > \lfloor K/2 \rfloor$  then
4   | Exit search. No optimal solution.
5 else
6   | Calculate observable user load for all  $\binom{K}{|C_1|}$  sets of  $|C_1|$  caches.
7 end if
8 if Observable Load  $\leq L_1$  for at least one set then
9   | Optimal solution obtained
10 else
11   | Go to Step 2
12 end if

```

In this subsection, we have compared the performance of the exhaustive search with the greedy algorithm. We have considered a campus set-up with $m = 3$ and K users in a social group. The joint mobility pattern of these K users is modeled as a discrete-time inhomogeneous Markov chain. As discussed in Section 3.1, transition matrices are constructed using the real-world location data of the users in the campus set-up.

Algorithm 3: Stage 1: Greedy Algorithm

```

1 Step 1: Initialize  $\mathcal{Q} = \emptyset$  and  $|C_1| = 0$ .
2 Step 2: Increment value of  $|C_1|$ .
3 if  $|C_1| > \lfloor K/2 \rfloor$  then
4   | Exit search.
5 else
6   | Consider only those candidates in Step 3 which have  $\mathcal{Q}$  as a subset.
7 end if
8 Step 3: Calculate observable user load for candidate sets of  $|C_1|$  caches.
9 if Observable Load  $\leq L_1$  for at least one set then
10  | Solution obtained
11 else
12  | Update  $\mathcal{Q}$  with the candidate having minimum user load. Go to Step 2.
13 end if

```

3.2.2.1 Stage 1 - Optimization at the Beginning of Each Day

Algorithm 2 presents the exhaustive search to compute the optimal set of caches. Let $|C_1|$ be the cardinality of the selected set of caches. At cardinality $|C_1|$, there will be $\binom{K}{|C_1|}$ candidate sets. Algorithm 2 evaluates the load corresponding to all the candidate sets. Then, it is checked whether there is a candidate set that has a load less than the load requirement, i.e., L_1 . If there is no such candidate set then the value of $|C_1|$ is incremented by 1, and the above process is repeated exhaustively. However, if there are candidate sets that meet the load requirement, then the optimal solution is obtained by selecting that candidate set that has the minimum load, and the algorithm stops. Further, while incrementing $|C_1|$, $|C_1| \leq \lfloor K/2 \rfloor$ condition should always be met. If the load requirement is not met until $|C_1| = \lfloor K/2 \rfloor$, then it implies that no optimal solution is feasible and the algorithm stops. Further, Algorithm 3 presents the proposed greedy algorithm for cache selection. Similar to Algorithm 2, $|C_1|$ is the cardinality of the selected set of caches. If a solution is not obtained at cardinality $|C_1|$ then \mathcal{Q} will be updated with the candidate set having minimum user load, and $|C_1|$ will be incremented by 1. The above process will be repeated until $|C_1| = \lfloor K/2 \rfloor$ or optimal solution is obtained.

Let $K = 6$, $p = 1$, $L_1 = 2$, $|C_0| = 0$ and $C_0 = \emptyset$. Using Algorithm 2, C_1 is initialized as 0. At step 2, $|C_1|$ is incremented to 1. Using Table 3.2, load corresponding to all the candidate sets is

Table 3.2: Table for $|C_1|= 1$ and $K = 6$

Candidate Sets	User Load
$\{U_1\}$	3.476
$\{U_2\}$	3.618
$\{U_3\}$	3.644
$\{U_4\}$	3.518
$\{U_5\}$	2.784
$\{U_6\}$	2.726

Table 3.3: Table for $|C_1| = 2$ and $K = 6$

Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2\}$	2.037	$\{U_2, U_6\}$	1.590
$\{U_1, U_3\}$	1.845	$\{U_3, U_4\}$	1.996
$\{U_1, U_4\}$	1.809	$\{U_3, U_5\}$	1.553
$\{U_1, U_5\}$	1.765	$\{U_3, U_6\}$	1.568
$\{U_1, U_6\}$	1.671	$\{U_4, U_5\}$	1.584
$\{U_2, U_3\}$	1.921	$\{U_4, U_6\}$	1.506
$\{U_2, U_4\}$	1.696	$\{U_5, U_6\}$	1.265
$\{U_2, U_5\}$	1.671		

evaluated. However, all candidate sets have load greater than L_1 . Hence, $|C_1|$ is incremented to 2. It can be observed from Table 3.3 that there are multiple candidate sets that have load lower than L_1 . Consequently, we select the candidate set with minimum load equal to 1.265 as the optimal set of caches, i.e., $C_1^* = \{U_5, U_6\}$.

Further, on applying the greedy algorithm, $|C_1|$ is initialized as 0 and $\mathcal{Q} = \emptyset$. At step 2 of Algorithm 3, $|C_1| = 1$. On incrementing $|C_1|$, only candidate sets with \mathcal{Q} as a subset will be considered. It can be observed from Table 3.2, no candidate set has load lower than L_1 . Hence, C_1 is incremented to 2. It is shown in Table 3.2, set $\{U_6\}$ has minimum value of user load. Therefore, $\mathcal{Q} = \{U_6\}$. Considering only those candidate sets in Table 3.3 that have \mathcal{Q} as a subset, there are multiple sets that have load less than L_1 . The candidate set with the minimum load equal to 1.265 is the solution of the greedy algorithm, $C_1^{prop} = \{U_5, U_6\}$. In this example, $C_1^* = C_1^{prop} = \{U_5, U_6\}$.

Similarly, for $K = 8$ and $L_1 = 2$, the optimal solution is evaluated as $\{U_5, U_6, U_7\}$, using Tables 3.4, 3.5 and 3.6, for exhaustive search as well as greedy algorithm. Further, for $K = 10$ and $L_1 = 1$, the solutions for the exhaustive search and greedy algorithm are obtained as $C_1^* = \{U_3, U_6, U_7, U_9\}$ and $C_1^{prop} = \{U_5, U_6, U_7, U_9\}$ ⁴. Hence, the cardinality of the solutions obtained using exhaustive search and greedy algorithm are same, however, the set of selected caches is different. For $K = 10$ and $L_1 = 1.8$, the optimal solution using exhaustive search is

⁴See Appendix C for the load values of the candidate sets.

Table 3.4: Table for $|C_1|=1$ and $K=8$

Candidate Sets	User Load
$\{U_1\}$	4.943
$\{U_2\}$	4.906
$\{U_3\}$	4.812
$\{U_4\}$	5.325
$\{U_5\}$	4.128
$\{U_6\}$	4.537
$\{U_7\}$	5.203
$\{U_8\}$	4.275

Table 3.5: Table for $|C_1|=2$ and $K=8$

Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2\}$	2.981	$\{U_3, U_5\}$	2.315
$\{U_1, U_3\}$	2.946	$\{U_3, U_6\}$	2.328
$\{U_1, U_4\}$	2.540	$\{U_3, U_7\}$	2.625
$\{U_1, U_5\}$	2.596	$\{U_3, U_8\}$	2.609
$\{U_1, U_6\}$	2.446	$\{U_4, U_5\}$	2.490
$\{U_1, U_7\}$	2.890	$\{U_4, U_6\}$	2.787
$\{U_1, U_8\}$	3.131	$\{U_4, U_7\}$	3.534
$\{U_2, U_3\}$	2.553	$\{U_4, U_8\}$	2.334
$\{U_2, U_4\}$	2.681	$\{U_5, U_6\}$	2.175
$\{U_2, U_5\}$	2.506	$\{U_5, U_7\}$	2.465
$\{U_2, U_6\}$	2.634	$\{U_5, U_8\}$	2.296
$\{U_2, U_7\}$	2.706	$\{U_6, U_7\}$	2.753
$\{U_2, U_8\}$	2.690	$\{U_6, U_8\}$	2.484
$\{U_3, U_4\}$	2.693	$\{U_7, U_8\}$	2.440

obtained as $C_1^* = \{U_1, U_4, U_6\}$ with load 1.77. However, the minimum load at cardinality three using greedy algorithm is 1.82. As the gap between the minimum load at cardinality three using exhaustive search and greedy algorithm is small, the solution using greedy algorithm is obtained at cardinality four with $C_1^{prop} = \{U_5, U_6, U_7, U_9\}$. However, the exhaustive algorithm has a search space of $\binom{K}{1} + \binom{K}{2} + \binom{K}{3} = 175$ candidate sets whereas the greedy algorithm has $(K + (K-1) + (K-2) + (K-3)) = 34$ candidate sets. The gap between the search spaces of the exhaustive search and greedy algorithm widens with K .

In order to validate the efficacy of the proposed algorithm for even higher number of users, we have done a detailed comparison of exhaustive search and proposed cache selection algorithm for $K=20$. The comparison is also presented for different load constraints. We

Table 3.6: Table for $|C_1|= 3$ and $K = 8$

Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2, U_3\}$	1.715	$\{U_1, U_6, U_8\}$	1.787	$\{U_3, U_4, U_7\}$	1.956
$\{U_1, U_2, U_4\}$	1.421	$\{U_1, U_7, U_8\}$	1.712	$\{U_3, U_4, U_8\}$	1.493
$\{U_1, U_2, U_5\}$	1.828	$\{U_2, U_3, U_4\}$	1.578	$\{U_3, U_5, U_6\}$	1.350
$\{U_1, U_2, U_6\}$	1.693	$\{U_2, U_3, U_5\}$	1.640	$\{U_3, U_5, U_7\}$	1.509
$\{U_1, U_2, U_7\}$	1.596	$\{U_2, U_3, U_6\}$	1.575	$\{U_3, U_5, U_8\}$	1.550
$\{U_1, U_2, U_8\}$	2.081	$\{U_2, U_3, U_7\}$	1.403	$\{U_3, U_6, U_7\}$	1.456
$\{U_1, U_3, U_4\}$	1.668	$\{U_2, U_3, U_8\}$	1.712	$\{U_3, U_6, U_8\}$	1.618
$\{U_1, U_3, U_5\}$	1.734	$\{U_2, U_4, U_5\}$	1.609	$\{U_3, U_7, U_8\}$	1.531
$\{U_1, U_3, U_6\}$	1.606	$\{U_2, U_4, U_6\}$	1.784	$\{U_4, U_5, U_6\}$	1.559
$\{U_1, U_3, U_7\}$	1.793	$\{U_2, U_4, U_7\}$	1.831	$\{U_4, U_5, U_7\}$	1.778
$\{U_1, U_3, U_8\}$	2.028	$\{U_2, U_4, U_8\}$	1.453	$\{U_4, U_5, U_8\}$	1.446
$\{U_1, U_4, U_5\}$	1.562	$\{U_2, U_5, U_6\}$	1.637	$\{U_4, U_6, U_7\}$	2.084
$\{U_1, U_4, U_6\}$	1.393	$\{U_2, U_5, U_7\}$	1.459	$\{U_4, U_6, U_8\}$	1.565
$\{U_1, U_4, U_7\}$	1.818	$\{U_2, U_5, U_8\}$	1.678	$\{U_4, U_7, U_8\}$	1.593
$\{U_1, U_4, U_8\}$	1.465	$\{U_2, U_6, U_7\}$	1.534	$\{U_5, U_6, U_7\}$	1.293
$\{U_1, U_5, U_6\}$	1.453	$\{U_2, U_6, U_8\}$	1.703	$\{U_5, U_6, U_8\}$	1.471
$\{U_1, U_5, U_7\}$	1.631	$\{U_2, U_7, U_8\}$	1.453	$\{U_5, U_7, U_8\}$	1.393
$\{U_1, U_5, U_8\}$	1.753	$\{U_3, U_4, U_5\}$	1.618	$\{U_6, U_7, U_8\}$	1.503
$\{U_1, U_6, U_7\}$	1.431	$\{U_3, U_4, U_6\}$	1.596		

have tabulated the solutions⁵ for $K = 20$ using exhaustive search and proposed algorithm in Table 3.7. Let us analyze one of the scenarios given in Table 3.7. Suppose at Stage 1 at a given day, $L_1 = 8$. On one hand, using the exhaustive search, $C_1^* = \{U_{16}, U_{19}\}$ and $|C_1^*| = 2$. On the other hand, proposed algorithm provides $C_1^{prop} = \{U_{15}, U_{16}\}$ and $|C_1^{prop}| = 2$. Hence, for this scenario, the cardinality of the selected set of caches is equal for both approaches. Similarly, observations were also made for other scenarios in Table 3.7. It exhibits that the load requirement also has an impact on the efficacy of the proposed algorithm.

In the exhaustive search, at a cardinality c of candidate sets, the candidate set with minimum load that meets the load constraint is selected as optimal solution. We observed that the minimum load at cardinality c obtained using the greedy algorithm is sometimes equal to that of exhaustive search (e.g., $K = 6, L_1 = 2$ and $K = 8, L_1 = 2$) or slightly higher (e.g., $K = 10, L_1 = 1, 1.8$). In cases where the minimum load is same, exhaustive and greedy algorithm will give the same solution. However, when there is a small difference in the minimum load values, the cardinality in greedy algorithm needs at most to be incremented by 1. Hence, based on our observations, the solution using the greedy algorithm will be $|C_p^{prop}| \in \{|C_p^*|, |C_p^*| + 1\}$.

⁵User load values corresponding to each of the candidate sets for Stage 1 are available at https://www.iiitd.edu.in/~wirocomm/resources/Social_Group_data.rar.

Table 3.7: Results for $K = 20$

Scenario	Load Requirements	Exhaustive Search	Proposed Algorithm
1.	$L_1 = 11$	$C_1^* = \{U_{15}\}, C_1^* = 1$	$C_1^{prop} = \{U_{15}\},$ $ C_1^{prop} = 1$
	$L_2 = 9$	$C_2^* = \{U_{15}, U_{16}\}, C_2^* = 2$	$C_1^{prop} = \{U_{15}, U_{16}\},$ $ C_1^{prop} = 2$
2.	$L_1 = 8$	$C_1^* = \{U_{16}, U_{19}\}, C_1^* = 2$	$C_1^{prop} = \{U_{15}, U_{16}\},$ $ C_1^{prop} = 2$
	$L_2 = 6$	$C_2^* = \{U_{16}, U_{19}\}, C_2^* = 2$	$C_2^{prop} = \{U_{15}, U_{16}, U_{19}\},$ $ C_2^{prop} = 3$
3.	$L_1 = 5$	$C_1^* = \{U_6, U_7, U_{19}\}, C_1^* = 3$	$C_1^{prop} = \{U_{15}, U_{16}, U_{19}\},$ $ C_1^{prop} = 3$
	$L_2 = 3$	$C_2^* = \{U_6, U_7, U_{19}, U_5\}, C_2^* = 4$	$C_1^{prop} = \{U_{15}, U_{16}, U_{19}, U_{17}\},$ $ C_1^{prop} = 4$

Further, we have demonstrated the search space reduction that can be achieved by the greedy algorithm. Consequently, in Section 3.3, we utilize greedy algorithm for the analysis of the cache selection framework.

3.2.2.2 Stage 2 - Optimization Due to Sudden Network Congestion

In a scenario where sudden network congestion is detected, the desired load, L_p on the network due to the social group will be lowered in the optimization problem at the p^{th} instant. Algorithm 4 presents the exhaustive search for Stage 2. Unlike Algorithm 2, Algorithm 4 initializes $|C_p| = |C_{p-1}^*|$ and $C_p = C_{p-1}^*$ such that the previously selected set of caches is not discarded. Then, the algorithm checks whether $|C_p| = |C_{p-1}^*|$ will meet the load constraint. If the load constraint is met, the algorithm will stop else $|C_p|$ will be incremented by 1. After this, similar to Algorithm 2, Algorithm 4 will carry out exhaustive search for the optimal solution. In Algorithm 4, $Select_1$ consists of all the candidate sets at a specific cardinality. $Select_2$ contains those sets, present in $Select_1$, that meet the load constraint. Further, the proposed algorithm for Stage 2 is illustrated in Algorithm 5. Initially, $|C_p| = |C_{p-1}^{prop}|$ and $C_p = C_{p-1}^{prop}$ such that the previously selected set of caches is not discarded. Then, the algorithm checks whether $|C_p| = |C_{p-1}^{prop}|$ will meet the load constraint. If the load constraint is met, the algorithm will stop else $|C_p|$ will be incremented by 1. After this, similar to Algorithm 3, Algorithm 5 will determine the solution.

Algorithm 4: Stage 2: Exhaustive Search

```
1 Step 1: Initialize  $|C_p| = |C_{p-1}^*|$  and  $C_p = C_{p-1}^*$ . Calculate user load for set  $C_p$ .
2 if Observable load  $\leq L_p$  then
3   | Optimal  $C_p^*$  is achieved.
4 else
5   | Go to Step 2.
6 end if
7 Step 2: Increment value of  $|C_p|$ .
8 if  $|C_p| > \lfloor K/2 \rfloor$  then
9   | No optimal solution.
10 else
11   | Calculate observable user load for all  $\binom{K}{|C_p|}$  sets of  $|C_p|$  caches.
12   | if Observable load  $\leq L_p$  for a set in Select1 then
13     |   | if  $C_{p-1}^*$  is a subset for a set in Select2 then
14       |     | Optimal solution  $C_p^*$ .
15     |   | else
16       |     | Go to Step 2
17     |   | end if
18   | else
19     | Go to Step 2
20   | end if
21 end if
```

Algorithm 5: Stage 2: Greedy Algorithm

- 1 Step 1: Initialize $|C_p| = |C_{p-1}^{prop}|$, $\mathcal{Q} = C_{p-1}^{prop}$ and $C_p = C_{p-1}^{prop}$.
 - 2 Calculate user load for set C_p .
 - 3 **if** *Observable load* $\leq L_p$ **then**
 - 4 | Solution C_p^{prop} is achieved.
 - 5 **else**
 - 6 | Go to Step 2.
 - 7 **end if**
 - 8 Step 2: Increment value of $|C_p|$.
 - 9 **if** $|C_p| > \lfloor K/2 \rfloor$ **then**
 - 10 | Exit search.
 - 11 **else**
 - 12 | Consider only those candidates in Step 3 which have \mathcal{Q} as a subset.
 - 13 **end if**
 - 14 Step 3: Calculate observable user load for candidate sets of $|C_p|$ caches.
 - 15 **if** *Observable Load* $\leq L_p$ **for at least one set** **then**
 - 16 | Solution obtained
 - 17 **else**
 - 18 | Update \mathcal{Q} with the candidate having minimum user load. Go to Step 2.
 - 19 **end if**
-

Now, let us say, for $K = 6$, due to sudden network congestion there will be a second instant of optimization with $L_2 = 1$ at the beginning of the 10^{th} slot of the specific day under consideration. As mentioned above, another optimization needs to be performed, and v_0 will now be the state at the 10^{th} slot. From the example above, $C_1^* = C_1^{prop} = \{U_5, U_6\}$. For the exhaustive search, Algorithm 4 is used. $|C_2|$ is initialized as $|C_1^*| = 2$. As the load constraint is not met at $|C_2| = 2$, $|C_2|$ is incremented to 3. The candidate sets at $|C_2| = 3$ are presented in Table 3.8 and are stored in $Select_1$. Out of the sets in $Select_2$, the set that has C_1^* as subset and has minimum load is obtained as $\{U_3, U_5, U_6\}$. Further, using Algorithm 5 for the greedy cache selection, $C_1^{prop} = \{U_3, U_5, U_6\}$.

Further, let there be a change in the load requirement from $L_1 = 8$ to $L_2 = 6$ for $K = 20$ as shown in Table 3.7. On using the exhaustive search the solution will be the same as before, i.e., $C_2^* = \{U_{16}, U_{19}\}$ and $|C_2^*| = 2$. However, proposed algorithm will result in $C_2^{prop} = \{U_{15}, U_{16}, U_{19}\}$ and $|C_2^{prop}| = 3$. Let there be a change in the load requirement to $L_2 = 6$. This will trigger a new optimization. On using the exhaustive search the solution will be the same as before, i.e., $C_2^* = \{U_{15}, U_{16}\}$ and $|C_2^*| = 2$. However, proposed algorithm will result in $C_2^{prop} = \{U_{15}, U_{16}, U_{19}\}$ and $|C_2^{prop}| = 3$. Therefore, the cardinality of selected set of caches differs by 1.

Table 3.8: Table for $|C_2| = 3$ and $K = 6$ in case of re-optimization at 10^{th} slot

Candidate Sets	User Load	Candidate Sets	User Load
$\{U_1, U_2, U_3\}$	0.887273	$\{U_2, U_3, U_4\}$	0.996364
$\{U_1, U_2, U_4\}$	0.898182	$\{U_2, U_3, U_5\}$	0.985455
$\{U_1, U_2, U_5\}$	1.058182	$\{U_2, U_3, U_6\}$	0.949091
$\{U_1, U_2, U_6\}$	1.134545	$\{U_2, U_4, U_5\}$	1.025455
$\{U_1, U_3, U_4\}$	1.014545	$\{U_2, U_4, U_6\}$	0.781818
$\{U_1, U_3, U_5\}$	0.978182	$\{U_2, U_5, U_6\}$	1
$\{U_1, U_3, U_6\}$	1.087273	$\{U_3, U_4, U_5\}$	1.087273
$\{U_1, U_4, U_5\}$	1.029091	$\{U_3, U_4, U_6\}$	0.890909
$\{U_1, U_4, U_6\}$	0.876364	$\{U_3, U_5, U_6\}$	0.785455
$\{U_1, U_5, U_6\}$	1.021818	$\{U_4, U_5, U_6\}$	0.82901

In general, the BSs present within the spatial spread of the social group of the K users can be informed about the selected caches. In this work, we have not directed the algorithm to the scenario where the transition matrix at a given slot boundary is also dynamic in nature. However, the work can easily be extended to such a scenario.

3.3 Result and Discussion

In this section, we demonstrate the results of the proposed cache selection framework for a specific working day at the campus using the greedy algorithm. Further, we have compared the performance of the mobility-unaware cache selection approach to the proposed cache selection framework. For a fair comparison, in mobility-unaware cache selection, the number of selected caches is kept equal to $|C_p^{\text{prop}}|$; however, any set can be randomly selected out of all the candidate sets [93]. We have considered social groups of size $K = 5, 10, 20$ to demonstrate the impact of the number of users in a social group on the selection of caches. The cardinality of the set of selected caches quantifies the caching load on the cellular network. The optimization has been performed in slot 1 of the specific day under consideration unless otherwise stated.

Table 3.9 exhibits the effect of γ_{cell} on the user load on the cellular network for $K = 5, 10, 20$ when $L_1 = 2, 4, 4$ respectively. It can be observed that with the increase in SNR threshold from 20 dB to 25 dB either user load on the core network increases or cardinality of the selected set of caches increases. This is because with increased threshold the size of feasibility sets reduces as explained in Section 3.1, and the load on the network will increase, or the network will require more caches to achieve the predefined user load constraint. As a result, for $K = 5$, U_5 is selected as the cache when threshold $\gamma_{\text{cell}} = 20$ dB whereas U_2 and U_5 are selected as caches at $\gamma_{\text{cell}} = 25$ dB.

Table 3.9: Table showing impact of γ_{cell}

K	L_1	γ_{cell}	Caches Selected	Load
5	2	20 dB	1 (U_5)	1.956
		25 dB	2 (U_2, U_5)	1.334
10	4	20 dB	2 (U_5, U_6)	2.781
		25 dB	2 (U_5, U_6)	3.204
20	4	20 dB	2 (U_{15}, U_{16}, U_{19})	3.765
		25 dB	2 (U_{15}, U_{16}, U_{19})	3.92

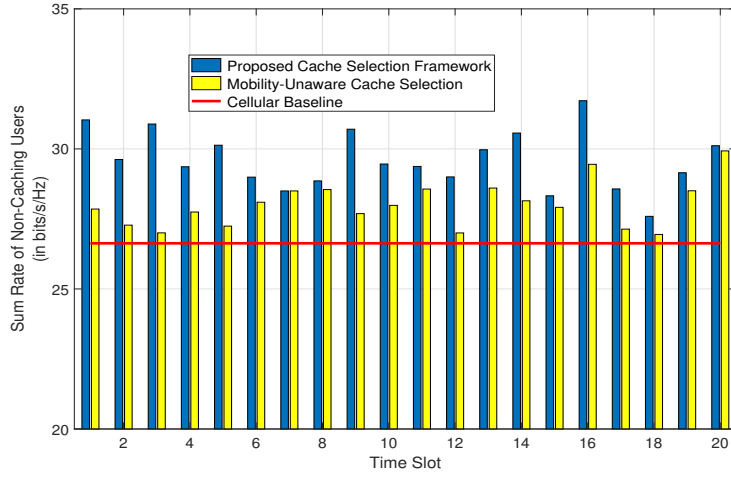


Figure 3.4: Sum-rate of non-caching users w.r.t. time slot when $\gamma_{cell} = 20$ dB.

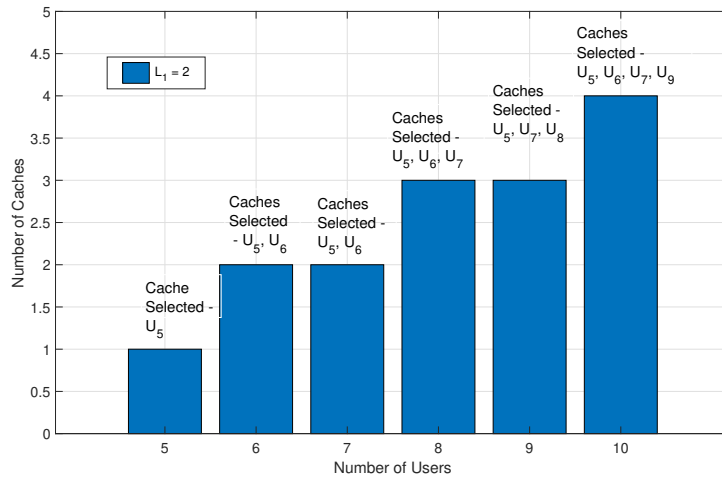


Figure 3.5: Number of Caches w.r.t. number of users to achieve $L_1 = 2$.

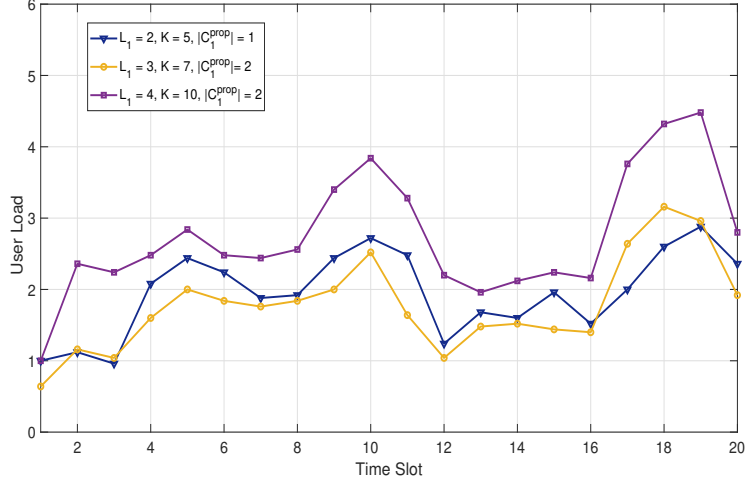


Figure 3.6: User load on cellular network w.r.t. time slot for $K = 5, 7, 10$.

Fig. 3.4 shows the achievable sum-rate of the non-caching users using the proposed spatio-temporal user behavior aware caching framework and mobility-unaware cache selection approach for $K = 5$ at $\gamma_{cell} = 20$ dB. The sum-rate expression for each slot is given as follows:

$$S_{rate} = \mathbb{E}_{v_l} \left[\sum_{i=1}^{(K-|C_1^{prop}|)} R_i \right], \quad (3.8)$$

where R_i is the rate achievable at the i^{th} user which may be served by one of the caches in C_1^{prop} via D2MD or cellular network. The sum rate is averaged over the observable states v_l in a given slot. From Table 3.9, it is known that U_5 should be selected as the cache at $K = 5$ and $\gamma_{cell} = 20$ dB, i.e., there are 4 non-caching users. In a conventional cellular communication with unicast links per user, the baseline cellular sum-rate performance will be $4 \times \log_2(1 + 100) = 26.632$ bits/s/Hz. It is obvious from Fig. 3.4 that the former performs worse than the proposed cache selection framework. This is because mobility-unaware cache selection fails to take into account the effect of spatio-temporal behavior and load constraint. Lesser sum-rate in mobility-unaware cache selection means more users are opting for cellular communication. Hence, the user load on the cellular network is higher in mobility-unaware cache selection. To achieve the desired user load on the cellular network, more caches have to be selected in the mobility-unaware cache selection approach. However, this will increase the caching load on the cellular network. Fig. 3.5 presents the set of caches required for $L_1 = 2$. It can be seen that as the number of users in the social group increases, the number of caches required will also increase. An interesting insight can be obtained through Fig. 3.5, i.e., even though the number of caches needed for $K = 8$ and 9 is equal, however, the caching users are not the same. This is because the total number of caches as well as the users that are to be selected as cache

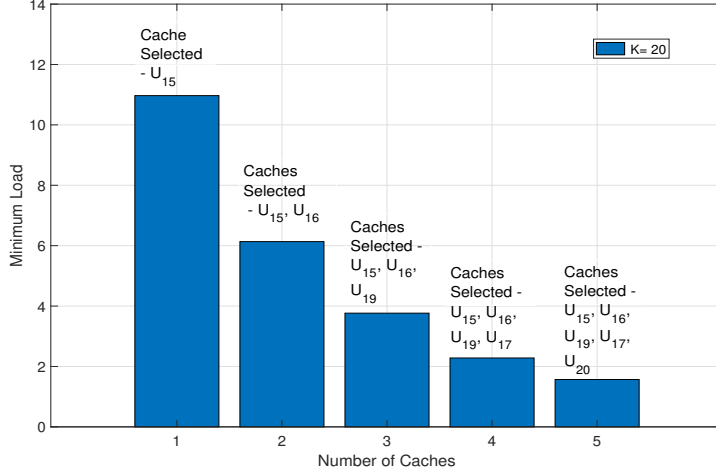


Figure 3.7: Minimum user load on cellular network w.r.t. number of caches for $K = 20$.

are the function of the joint mobility patterns and pairwise SINR values of users. Hence, for $K = 8$ and 9 sets $\{U_3, U_6, U_7\}$ and $\{U_5, U_7, U_8\}$ are selected as caches respectively.

Fig. 3.6 shows the per time slot user load on the cellular network at $K = 5, 7, 10$ and $L_1 = 2, 3, 4$ respectively when set C_1^{prop} of caches is utilized. The user load on the network varies with the time of the day as in each time slot different set of state vectors are observable depending on the transition matrices. This is also evident from the discussion in Section 3.1. Further, with a change in K and the desired user load L_1 , the user load on the network for each slot also shows a variation. This is because C_1^{prop} obtained using the greedy algorithm varies with the value of K and L_1 . Fig. 3.7 shows the impact of $K = 20$ and the number of caches selected on the minimum user load on the cellular network. It can be observed that as the number of caches increases, the minimum user load on cellular network decreases. However, an increase in the number of caches will strain the cellular network at the time of content caching.

Now, let us say, for $K = 6$, due to sudden network congestion, another optimization needs to be performed and v_0 is now the state of the network at the 10^{th} slot. $L_1 = 2$ is reduced to $L_2 = 1$ (example with $K = 6$ and $L_1 = 2$ illustrated in Section 3.2) at the beginning of the 10^{th} slot of the specific day other consideration. In case the constraint (3.7c) introduced in Section 3.2 is not applied to our proposed framework, using Table 3.8, set $\{U_2, U_4, U_6\}$ with minimum user load less than L_2 will be selected as caches. Thus, two new caches are added to the network. On the other hand, on incorporating the constraint (3.7c) and using $\mathcal{Q} = C_1^{prop} = \{U_5, U_6\}$, $\{U_3, U_5, U_6\}$ will be selected as the set of caches. Consequently, only one new cache, i.e., U_3 needs to be added. This implies that our proposed framework also helps in alleviating the caching load due to frequent optimizations.

3.4 Conclusion

A greedy algorithm for cache selection is proposed to solve the combinatorial optimization problem of selecting a set of the minimum number of caches to achieve the desired user load for a D2MD network. The set of caches is selected by utilizing real-world location information to obtain the spatio-temporal behavior of the users. The proposed work has been shown to alleviate the caching load on the cellular network. Moreover, the proposed framework does not discard the previously selected caches. Further, a discrete-time inhomogeneous Markov chain is presented to model the joint mobility pattern for the users in the D2MD network is developed. The selected caches are tagged to their social group and are responsible for doing D2D multicast to disseminate the popular multimedia files to the non-caching users. It is observed that the proposed algorithm is computationally less intensive with complexity $\mathcal{O}(K^2)$ as compared to complexity $\mathcal{O}(2^{K-1})$ of the exhaustive search. Further, the presented optimization has been shown to perform better than mobility-unaware cache selection.

Enabling Disaster Resilient Communication Using Multi-Hop D2D

In line with the objective of the thesis, in this chapter, we investigate multi-hop D2D to establish disaster-resilient communication networks. Multi-hop D2I/I2D routing can be utilized to connect to an active BS, which is present outside the dead-spot. For instance, in [48], the authors studied the benefits of multi-hop D2D in extending the coverage area of BS in public safety scenarios. It was shown that the average energy and spectral efficiencies due to multi-hop D2D are enhanced when the number of hops is increased. Like [47], they have utilized the SPR algorithm. Authors in [49] proposed a routing scheme utilizing the ant colony optimization to maximize the end-to-end throughput for all the data flows originating from the area without cellular network coverage. However, the aforementioned works of multi-hop D2I/I2D using SPR, for public safety and disaster scenarios, result in inefficient use of wireless radio resources in the presence of self-interference or contention among different users [18]. This is due to the fact that they have overlooked the half-duplex nature of the D2D relays, i.e., a D2D relay can either transmit or receive so that the transmitted and received data flows do not interfere while making the routing decisions. Further, it is important to restrict the data flows that can be transmitted or received simultaneously by a D2D relay to limit the contention among different flows/users [19]. Hence, it is imperative to jointly address the problem of routing and scheduling in the multi-hop D2D network.

Joint routing and scheduling algorithms have been widely discussed in the literature in the context of ad-hoc wireless networks [18, 50, 94–96]. For instance, in [94], authors have maximized the end-to-end data flow by solving the joint routing and scheduling problem. Authors in [18] propose an optimal joint design of cross-layer congestion control, routing and scheduling for ad-hoc networks. In [50], authors considered the nodes to be equipped with multi-radios. Hence, they solve the problem of routing and scheduling for multi-radio

ad-hoc networks. Similar to [50], [96] also considers multi-radio ad-hoc network. The work in [95] considers the wireless mesh network constituting the backbone of third-generation (3G) networks. In this work, the problem of joint routing and scheduling has been solved in the presence of directional antenna equipped nodes. In [97], authors have proposed routing and scheduling policies that optimize network throughput in multi-hop wireless networks where nodes are powered by renewable energy sources. However, the above works on routing and scheduling generally involve a single destination or fixed source-destination pairs. When the source-destination pairs are fixed prior to routing and scheduling, the users who are in close proximity may contend for the same wireless link to connect to the same active BS. Consequently, on applying the existing joint routing and scheduling algorithms to a time-bounded disaster-resilient communication network, fewer number of users will get covered.

In the proposed framework, a user is said to be covered by the network once it establishes connectivity with any of the active BSs within the given time frame. Hence, there is a possibility of an optimal selection of BS with which a user can be paired. By exploiting the above, we propose a joint optimization of source-destination pairing, routing, and scheduling to maximize the number of covered users within a given time frame in the dead spot. To the best of our knowledge, the existing literature has not dealt with joint source-destination pairing, routing, and scheduling for a multi-hop D2D network in a disaster scenario.

4.0.1 Contributions

In this work, we present a disaster-resilient multi-hop D2D network that employs a joint source-destination pairing, routing and scheduling framework to maximize the number of users that can be covered in the dead spot by an active BS within a given time frame. Unlike the prior works on multi-hop D2I/I2D routing in a disaster scenario, for a coverage maximization problem within T_s time slots, routing, scheduling and pairing are jointly optimized. The formulated optimization problem is an integer linear programming (ILP) problem and is computationally expensive. The existing routing and scheduling algorithms are not applicable as they do not decide the source-destination pairings. Consequently, we have proposed a low complexity graph-based scheduling constraint aware routing and pairing (SCARP) algorithm. SCARP is, as the name suggests, aware of the scheduling constraints and jointly determines the source-destination pairings and scheduled paths for the multi-hop D2D network. Further, SCARP results in up to 92% reduction in processing time as compared to the optimal solution.

The major contributions of the proposed work are summarized below:

- We propose a joint source-destination pairing, routing, and scheduling framework for a multi-hop D2D communication in the dead spot.

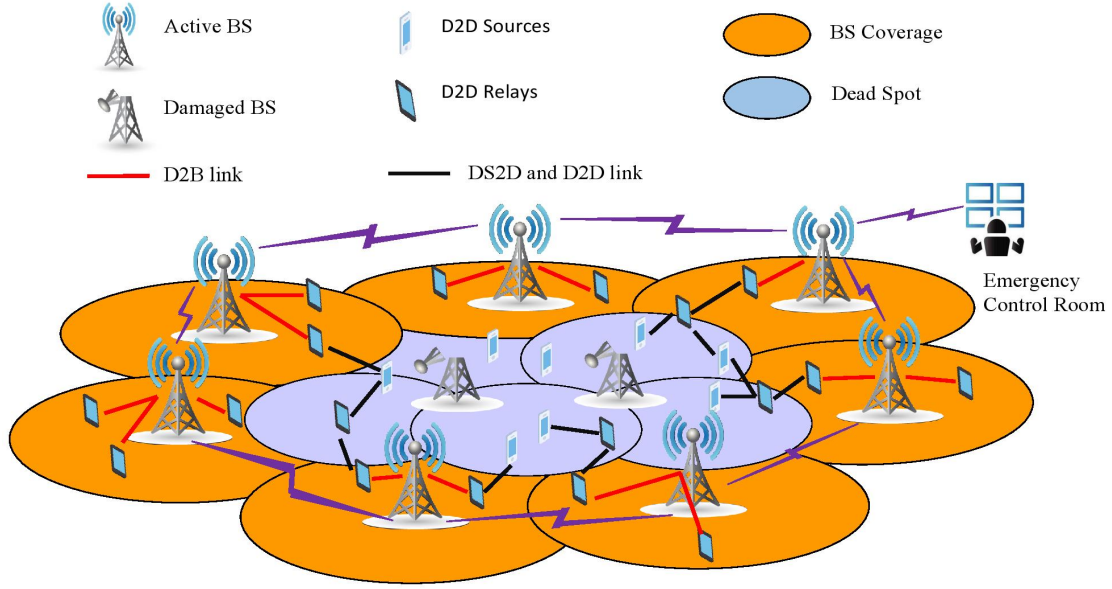


Figure 4.1: Disaster-resilient network model

- The maximization of covered users in the dead spot is formulated as an ILP. We show that the ILP is NP-hard; hence, a low complexity scheduling constraint aware routing and pairing (SCARP) algorithm is proposed to perform joint pairing, routing, and scheduling, unlike the existing algorithms.
- The reduction in the processing time required to obtain a solution to the ILP with SCARP as compared to the optimal solution is also demonstrated.
- In addition, we also compare SCARP with an SPR based scheduling. In the case of SPR based scheduling, we have considered fixed as well as unknown user-BS pairings¹. The improvement in the number of covered users is recorded in SCARP w.r.t. SPR based scheduling and pairing schemes.

4.1 Disaster-Resilient Network Model

We consider an uplink multi-cell network where a part of the network has been impacted due to disaster resulting in a dead spot containing non-functional BSs and users without any network connectivity, as shown in Fig. 4.1. Let us denote the set of users inside and outside the dead spot as D_1 and D_2 , respectively. In the aftermath of the disaster, at a given time instance, let there be a subset of the local population termed as “D2D sources” in the dead

¹Details of SPR based scheduling and pairing are provided in Section 4.4.

spot who are trying to communicate with close relatives, remote emergency control room, etc. Hence, the D2D sources will have to route their information to an active BS outside the dead spot. With the assistance of multi-hop D2D communication, a D2D source can expect to be connected to a BS in set B_s of active BSs, creating a disaster-resilient communication network. Let DS be the set of D2D sources where $DS \subset D_1$. Rest of the users in D_1 and the idle users in D_2 will constitute the set D of D2D relays. Fig. 4.1 shows the set D of D2D relays distributed across the multi-cell network. A connection between a D2D source and an active BS is assumed to be successful only when an active BS receives a data packet from the D2D source within time frame \mathcal{T} via multi-hop D2D communication. Let time frame \mathcal{T} consist of T_s time slots. It may be noted that in the present context, one time slot corresponds to one hop of communication. Further, in a disaster scenario, users will generally have limited mobility and are usually confined to specific locations. In a dead spot, it is reasonable to assume that the affected users will be given higher priority for establishing a connection as compared to cellular users outside the dead spot. Consequently, the D2D communication in the dead spot is allocated R RBs where one RB is assigned to each D2D source and $R > |DS|$ where $|\cdot|$ denotes the cardinality of a set. In a disaster scenario, there is a need to take centralized routing decisions as suggested in [49]. Hence, the joint pairing, routing, and scheduling decisions are assumed to be carried out at a central emergency control room. We assume D2D sources and D2D relays in the dead spot can be localized and this information is available at the control room. For instance, UAV networks can be utilized for such localization [98]. The details regarding the handling of the control-related traffic is out-of-scope of the present work. The disaster-resilient communication network in Fig. 4.1 can be modeled as a contact graph $G(V, E)$ where D2D sources, D2D relays and active BSs constitute the set of nodes V , and E is the set of edges. An edge will exist between any two nodes in the graph if the nodes are within each other's D2D transmission range. The weight of each edge is equal to w because all the edges satisfy the D2D transmission range criteria. Let L_{D2D} and L_{D2B} denote the set of links between D2D relays, and D2D relays and active BSs respectively. There is a set of links L_{DS2D}^m between a D2D source m and D2D relays. Let $L^m = L_{D2D} \cup L_{D2B} \cup L_{DS2D}^m$ be the set of all links through which D2D source m can possibly reach any active BS. As mentioned above, the motivation of the proposed work is to maximize the number of users covered in the dead spot, i.e., to facilitate more end-to-end connections from D2D sources in the dead spot to active BSs within T_s time slots. This optimization problem has been formulated as an ILP problem and is shown below to be NP-hard. Consequently, this chapter proposes a novel SCARP algorithm for multi-hop D2D in a disaster-resilient network.

4.2 Problem Formulation

In this section, we formulate an optimization problem where the objective function is to maximize the number of users that can be covered in the dead spot using multi-hop D2D within a deadline of T_s slots. A D2D source is said to be covered if there is an end-to-end connection from the D2D source routed through the D2D relays to an active BS within T_s slots. In the following, we introduce some key variables used in the problem formulation. Further, the details of the objective function and constraints of our problem are provided.

4.2.1 Variables

Let $(i, k) \in E$ be the link present between node i and k where $i, k \in V$. $F_{i,k}^m \in \{0, 1\}$ is the flow variable which is ‘1’ when there is a flow of a data packet through $(i, k)^{th}$ link originated at D2D source m and ‘0’ otherwise. Further, let $A_{i,k}^m(t) \in \{0, 1\}$ be the scheduling variable which is ‘1’ when the link (i, k) is scheduled to route the flow of D2D source m in time slot $n \in \mathcal{T}$ or ‘0’ otherwise.

4.2.2 Objective Function

The objective is to maximize the total number of covered users. If a D2D source m is getting covered $F_{i,k}^m$ will be ‘1’ for a link $(i, k) \in L_{D2B}$ else $F_{i,k}^m$ will be ‘0’ $\forall (i, k) \in L_{D2B}$. Hence, the total number of users covered is represented using $\sum_{m \in DS} \sum_{(i,k)} F_{i,k}^m$. In other words, $F_{i,k}^m = 1$ for $(i, k) \in L_{D2B}$, denotes that D2D source m will be paired with BS $k \in B$. The objective function is as given below:

$$\text{maximize } \sum_{m \in DS} \sum_{(i,k)} F_{i,k}^m, (i, k) \in L_{D2B}. \quad (4.1)$$

4.2.3 Constraints

In the following, we will discuss the constraints of our problem.

Constraint 1: In the considered graph-based network model, at each D2D relay, the flow constraints must be met, i.e., the amount of incoming flow of data must be equal to the outgoing flow of data and is formulated as:

$$\sum_{k:(i,k)} F_{i,k}^m - \sum_{l:(l,i)} F_{l,i}^m = 0, \forall i \in D, (i, k), (l, i) \in L_{D2D}, \forall m \in DS. \quad (4.2)$$

Constraint 2: There is a possibility that a D2D source is not covered by any BS. Hence, the flow constraint at a D2D source is as follows:

$$\sum_{k:(i,k)} F_{i,k}^m \leq 1, (i,k) \in L_{DS2D}^m, \forall m \in DS. \quad (4.3)$$

Constraint 3: It is not known beforehand which BS will receive D2D source m 's data. This flow constraint is as follows:

$$\sum_{i:(i,k)} F_{i,k}^m \leq 1, \forall k \in B, (i,k) \in L_{D2B}, \forall m \in DS. \quad (4.4)$$

The above constraint implies that at most one $D2B$ link will be active for the m^{th} D2D source.

Constraint 4: During the time frame \mathcal{T} link (i,k) can be active for a D2D source m at most once. This is because within \mathcal{T} a D2D source will be sending out only one packet of data. This scheduling constraint is formulated as:

$$\sum_{t \in \mathcal{T}} A_{i,k}^m(n) = F_{i,k}^m, \forall m \in DS, (i,k) \in L^m. \quad (4.5)$$

Constraint 5: There can be at most $|DS|$ data flows in the network originating from $|DS|$ D2D sources. In order to allow the reception of single data flow, originating from a single D2D source, at a D2D relay at a given time slot n , below reception constraint is added:

$$\sum_{m \in DS} \sum_{i:(i,k)} A_{i,k}^m(t) \leq 1, \forall t \in \mathcal{T}, \forall k \in D, (i,k) \in L_{DS2D}^m \cup L_{D2D}. \quad (4.6)$$

Constraint 6: In order to allow transmission of single data flow from the D2D sources and D2D relays at a given time slot n , a transmission constraint is introduced as follows:

$$\sum_{m \in DS} \sum_{k:(i,k)} A_{i,k}^m(t) \leq 1, \forall t \in \mathcal{T}, \forall i \in D \cup DS, (i,k) \in L_{DS2D}^m \cup L_{D2D}. \quad (4.7)$$

Constraint 7: The constraint to limit simultaneous transmission/reception in order to avoid self-interference at a node is given as follows:

$$\sum_{m \in DS} A_{i,k}^m(t) + \sum_{m \in DS} A_{k,l}^m(t) \leq 1, \forall t \in \mathcal{T}, \forall k \in D, (i,k), (k,l) \in L_{DS2D}^m \cup L_{D2D}. \quad (4.8)$$

This constraint is applicable for the links in L_{D2D} and L_{DS2D}^m .

Constraint 8: In the first time slot, it is obvious that no link can be established between the D2D relays, and a D2D relay and BS. This has been taken care of in the constraint given below:

$$A_{i,k}^m(1) = 0, \forall m \in DS, \forall (i,k) \in L_{D2D} \cup L_{D2B}. \quad (4.9)$$

Constraint 9: The flow of data originating at m arriving at $k \in D$ through any link (i,k) during \mathcal{T} , can be denoted by $\sum_{i:(i,k)} F_{i,k}^m$. Since, there is a possibility that a D2D source is not covered by a BS $\sum_{i:(i,k)} F_{i,k}^m$ can be '0' or '1'. Let $\sum_{i:(i,k)} F_{i,k}^m = 1$ and say data has not arrived at k till $n-1$ slots, i.e., $\sum_{\bar{n}=1}^{n-1} \sum_{i:(i,k)} A_{i,k}^m(\bar{n}) = 0$. Therefore, any link $(k,l) \in L_{D2D} \cup L_{D2B}$ must not be scheduled for m 's flow, i.e., $A_{k,l}^m(t)$ must be 0. To ensure the above causality constraint, i.e., a node k cannot transmit in the n^{th} slot unless it has received data in the previous $(n-1)$ slots the below condition needs to be satisfied:

$$\begin{aligned} \sum_{\bar{n}=1}^{n-1} \sum_{i:(i,k)} A_{i,k}^m(\bar{n}) + (1 - A_{k,l}^m(n)) &\geq \sum_{i:(i,k)} F_{i,k}^m, \forall n \in \mathcal{T}, \\ \forall m \in DS, (i,k) \in L_{DS2D}^m \cup L_{D2D}, \forall (k,l) \in L_{D2D} \cup L_{D2B}. \end{aligned} \quad (4.10)$$

Hence, the formulated optimization problem is given as follows:

$$\begin{aligned} &\text{maximize} && \sum_{m \in DS} \sum_{(i,k)} F_{i,k}^m, (i,k) \in L_{D2B} \\ &\text{subject to} && (4.2) - (4.10) \end{aligned} \quad (4.11)$$

4.2.4 Problem Complexity

The NP-hardness of our optimization problem can be proven by reducing from the preemptive scheduling problem in [99], which is proven to be NP-hard. In the preemptive scheduling problem, a set P of tasks is scheduled within a deadline. Each task $p \in P$ is subdivided into sub-tasks p_1, p_2, \dots, p_k . The duration of each task is the sum of the duration of the sub-tasks. The scheduling of sub-task p_i precedes the scheduling of p_{i-1} . Let the set of tasks P be the set of the routing paths of the D2D sources willing to connect to any active BS. Each task p will be a routing path and sub-tasks p_1 to p_k will be the links on the routing-path p . Consequently, preemptive scheduling maps to our problem. Hence, our problem is NP-hard with high time complexity.

4.3 Scheduling Constraint Aware Routing and Pairing (SCARP) Algorithm

In this section, we present the intuition behind the proposed SCARP algorithm. Then, we explain in detail the working of SCARP and illustrate it using an example. Further, we also evaluate the time complexity of SCARP.

4.3.1 Key Intuition

As shown above, the formulated ILP problem is NP-hard. Hence, an alternate approach is required to solve the formulated problem with lower complexity. For the considered time-bounded communication, SPR can find the routing paths with the minimum number of hops as the network links in the path are equally weighed. Given a single D2D source and corresponding destination, SPR seems to be an obvious approach to get the D2D source covered within time frame \mathcal{T} . However, when there are multiple D2D sources in the network, SPR will not ensure maximum coverage as it takes the routing decisions without considering the scheduling possibilities of the shortest paths within time frame \mathcal{T} [100]. Hence, we propose a SCARP algorithm that utilizes the shortest path algorithm in a recursive manner (explained in detail below) to select non-overlapping shortest paths to take care of the scheduling constraints. Further, SCARP also accounts for the fact that, in the formulated problem, the source-destination pairs are not known beforehand and decides the D2D source-BS pairing. SCARP is given $G(V, E)$ as input, and each D2D source's scheduled path is obtained as output, which gets stored in \mathcal{P} where \mathcal{P} is the list containing the scheduled paths of each D2D source. The scheduled paths for each D2D source are obtained recursively to keep re-selecting a scheduled path for a D2D source until it meets the scheduling constraints. Assuming a distance-dependent path loss, the proposed SCARP algorithm considers the link distances of the user pairs (D2D sources and D2D relays), and the user-BS pairs.

4.3.2 Detailed Working of SCARP

The proposed SCARP algorithm is presented in Algorithm 6. The working of the algorithm is as follows. $G(V, E)$ is provided as input to the algorithm. As mentioned above, the weight of each link is initialized with the same value w . The algorithm is executed only when $T_s \geq 2$, equivalent to a minimum of two hops that are required to establish a multi-hop D2D communication as direct communication between an active BS and a user in the dead spot is not feasible. For a D2D source m , function `Main()` is called which has the arguments: $G(V, E)$, \mathcal{P} and \mathcal{N} where \mathcal{N} is the list of D2D sources connected to each BS in B_s . Within `Main`, the

Algorithm 6: Proposed SCARP Algorithm

```

1 Initialize:  $\mathcal{P} = \emptyset, \mathcal{N} = \emptyset$ 
2 if  $T_s \geq 2$  then
3   for  $m = 1 : 1 : |DS|$  do
4     Function  $\text{Main}(G(V, E), \mathcal{P}, \mathcal{N}, m)$ :
5       Initialize:  $C_{relays} \leftarrow \emptyset, Nodes \leftarrow \emptyset, S_{DS} \leftarrow \emptyset, shortest\_paths \leftarrow \emptyset, B_{weights} \leftarrow \emptyset$ ;
6       Find shortest path from  $m$  to each BS; Update  $shortest\_paths$ ;  $B_{weights} \leftarrow \text{weight}$ 
          ( $shortest\_paths$ );
7       if  $\min(B_{weights}) \leq T$  then
8         Update  $C_m$ ;
9         if  $|C_m| > 1$  then
10          |  $\bar{j} \leftarrow \text{Select}(C_m)$ ;
11          end if
12          if  $\mathcal{N}_{\bar{j}} \neq \emptyset$  then
13            for  $k = 1 : 1 : |\mathcal{N}_{\bar{j}}|$  do
14              if  $|H_m - H_{\mathcal{N}_{\bar{j}}(k)}| < 2$  then
15                | Update  $Nodes$ ;
16                end if
17            end for
18            if  $|Nodes| > 0$  then
19              Add  $m$  to  $Nodes$ ;  $S_{DS} \leftarrow \min_{weights}(Nodes)$ ;
20              if  $|S_{DS}| < 2$  then
21                 $S_{re} = S_{DS}$ ;  $Nodes \leftarrow Nodes - \{S_{re}\}$ ; Update  $\mathcal{N}$  and  $\mathcal{P}$ ;
22                for  $r = 1 : 1 : |Nodes|$  do
23                  |  $C \leftarrow \text{Find}(S_{re}, Nodes(r))$ ; Add  $C$  to  $C_{relays}$ ;
24                  end for
25                  if  $C_{relays} \neq \emptyset$  then
26                    |  $G(V, E) \leftarrow \text{Update}(G(V, E), C_{relays})$ ;  $\text{Main}(G(V, E), \mathcal{P}, \mathcal{N}, S_{re})$ ;
27                    end if
28                else
29                   $Nodes \leftarrow Nodes - \{m\}$ ;
30                  for  $r = 1 : 1 : |Nodes|$  do
31                    |  $C \leftarrow \text{Find}(m, Nodes(r))$ ; Add  $C$  to  $C_{relays}$ ;
32                    end for
33                    if  $C_{relays} \neq \emptyset$  then
34                      |  $G(V, E) \leftarrow \text{Update}(G(V, E), C_{relays})$ ;  $\text{Main}(G(V, E), \mathcal{P}, \mathcal{N}, m)$ ;
35                      else
36                        | Update  $\mathcal{N}$  and  $\mathcal{P}$ ;
37                        end if
38                  end if
39                else
40                  | Update  $\mathcal{N}$  and  $\mathcal{P}$ ;
41                  end if
42                else
43                  | Update  $\mathcal{N}$  and  $\mathcal{P}$ ;
44                  end if
45              end if
46              return  $\mathcal{P}$ ;
47            End Function;
48          end for
49 end if

```

shortest path to each of the active BSs is calculated based on the sum of weights of the links between source m and each BS, and stored in $shortest_paths$. Finding the shortest paths with respect to each of the BSs takes into account the fact that the destination of each D2D source is not known beforehand. The function $weight()$ finds the sum of the link weights of each shortest path and is stored in $B_weights$. Function $min()$ provides the minimum of $B_weights$.

If the minimum of $B_weights$ is less than T_s , i.e., the shortest path can be scheduled within T_s slots, the information of the BSs having minimum $B_weights$ is stored in C_m . Otherwise, no optimal path is possible for the D2D source m . In case C_m contains more than one BS, the BS with the minimum sum of link distances from m is selected using $Select()$ function. Now, it has to be checked whether the selected BS is already selected as a destination for other D2D sources. This is because if any two D2D sources have a common destination (BS), higher is the possibility of having overlapping shortest paths. $\mathcal{N}_j \in \mathcal{N}$ is the set of D2D sources using BS j . If the m^{th} and other D2D sources have a common destination (BS) \bar{j} , the number of hops in their paths to the common BS are compared. Here, H_m and $H_{\mathcal{N}_{\bar{j}}(k)}$ are the number of hops in the path of D2D source m and k^{th} D2D source present in $\mathcal{N}_{\bar{j}}$. In Algorithm 6, $Nodes$ consist of the D2D sources with $|H_m - H_{\mathcal{N}_{\bar{j}}(k)}| < 2$. This step is required to meet the scheduling constraints, i.e., the constraints on simultaneous transmission and reception, and simultaneous multiple transmission/reception at a node. If m and any D2D source in $\mathcal{N}_{\bar{j}}$ fails to meet the scheduling constraints, i.e., $|Nodes| > 0$, then re-selection of the scheduled path of m or S_{re} is performed. Prior to making the decision which D2D sources' path needs to be re-selected, the D2D source m is added to the set $Nodes$. If there is only one D2D source in $Nodes$ with the minimum sum of link weights, i.e., $|S_{DS}| < 2$, S_{re} 's path is re-selected. However, if there is more than one D2D source with the minimum sum of weights, i.e., $|S_{DS}| > 1$, there is a re-selection of the path for m . The function $Find()$ has S_{re} or m and a D2D source in $Nodes - \{S_{re}\}$ or $Nodes - \{m\}$ as the input arguments (at steps 23 and 31) and outputs the common relays between their paths to the common destination \bar{j} , which is stored in C_{relays} . Using function $Update$ the weights of all the edges that connect to the relays in C_{relays} are updated to a higher weight $w' \gg w$. Updating the links with a higher weight will ensure that the previously chosen path for m is not selected again. This is due to the fact the previously chosen path for m or S_{re} will no longer be the shortest path to fulfill the criteria in line 7 of Algorithm 6. However, if C_{relays} is empty in case of m 's path re-selection, then \mathcal{P} is updated with the currently selected path of the D2D source m . In the above, the process of finding the shortest path is repeated recursively.

Correctness: SCARP algorithm schedules the D2D sources within the deadline of T_s slots. For a D2D source m , function $Main()$ is called for obtaining the scheduled path. If the condition in line 7 is not met for a D2D source, the function $Main()$ is exited. Then, SCARP moves on to the next D2D source. The algorithm terminates when the value of m equals to $|DS|$ or when $T_s < 2$.

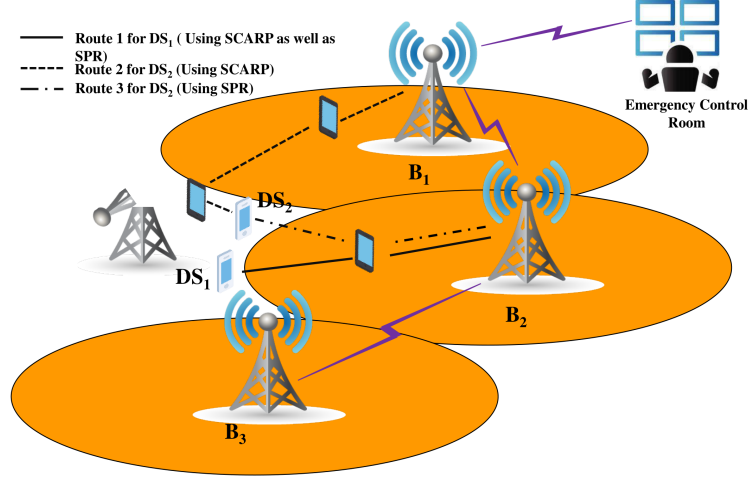


Figure 4.2: Network example for showing the route selection in SCARP and SPR.

Fig. 4.2 shows the routing decisions taken by the SCARP algorithm and SPR for the network. In the case of SPR, Route 1 and Route 3 are selected for D2D sources DS_1 and DS_2 , respectively, which have overlapping intermediate D2D relays. Consequently, for $T_s = 3$, at the time of scheduling DS_2 will not be covered by any active BS. On the other hand, for $T_s = 3$, in case of SCARP, DS_1 and DS_2 are paired with B_2 and B_1 respectively and their multi-hop paths can be scheduled within $T_s = 3$.

4.3.3 Time Complexity of SCARP

The time complexity of the for loop at step 3 of Algorithm I is $\mathcal{O}(|DS|)$, where $|DS|$ is the number of D2D sources in dead spot. The shortest path to each BS is calculated at step 6. Using Dijkstra's algorithm, the time taken is $\mathcal{O}(|B||E|\log|N|)$, where $|B|$, $|E|$ and $|N|$ are the number of BSs, edges and nodes respectively in the network. Further, for D2D source m , there are $(|DS| - 1)$ D2D sources with whom m can have overlapping paths. Therefore, in the worst case, for D2D source m function `Main()` will be called $|DS|$ times and hence, the time complexity of SCARP algorithm is $\mathcal{O}(|DS|^2|B||E|\log|N|)$.

4.4 Results and Discussion

In this section, we evaluate the performance of the proposed framework in a disaster-resilient communication network. The simulation set-up is described below:

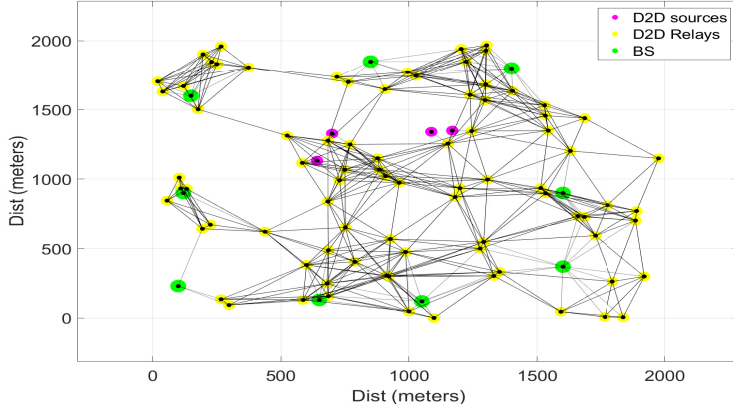


Figure 4.3: Network graph for a specific network realization.

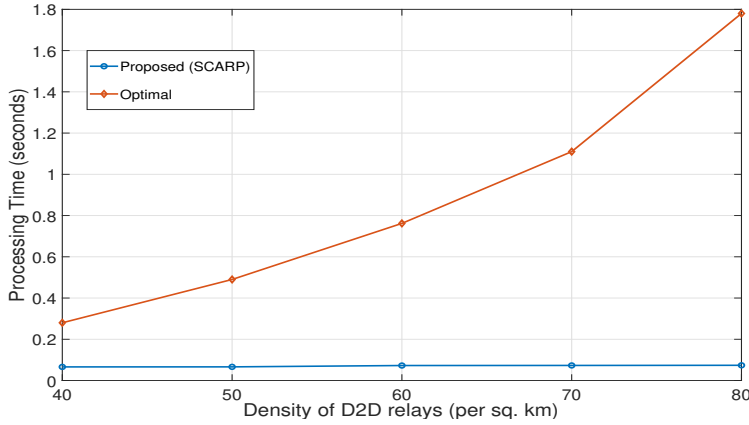


Figure 4.4: Processing time w.r.t. density of D2D relays

4.4.1 Simulation Set-Up

For each network realization, we have assumed a fixed number of active BSs having fixed locations that surround the dead spot. D2D sources in the dead spot are assumed to have random locations. D2D relays in the network are considered to follow a Poisson point process (PPP). The simulated network area is 2×2 sq. km. Unless otherwise stated, number of BSs, $|B| = 9$; number of D2D sources, $|DS| = 5$; density of D2D relays = $50/\text{km}^2$ and number of time slots, $T_s = 3$. For the graph-based network model, initial edge weight, $w = 1$ and updated edge weight (in Step 34, Algorithm 6), $w' = 100$. We have assumed a distance-dependent path loss model. Let the maximum transmission range of each BS and D2D device be 450 m and 150 m, respectively. The optimal solution to the formulated problem has been obtained using the IBM CPLEX solver.

The network simulations and SCARP algorithm execution have been done on MATLAB. All the above tasks have been run on an i5 CPU with 4 GB RAM and 64-bit operating system. The optimal solution and results corresponding to the SCARP algorithm have been demonstrated

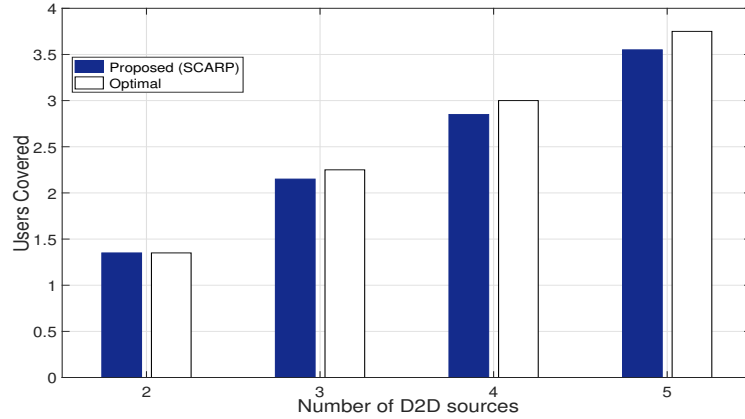


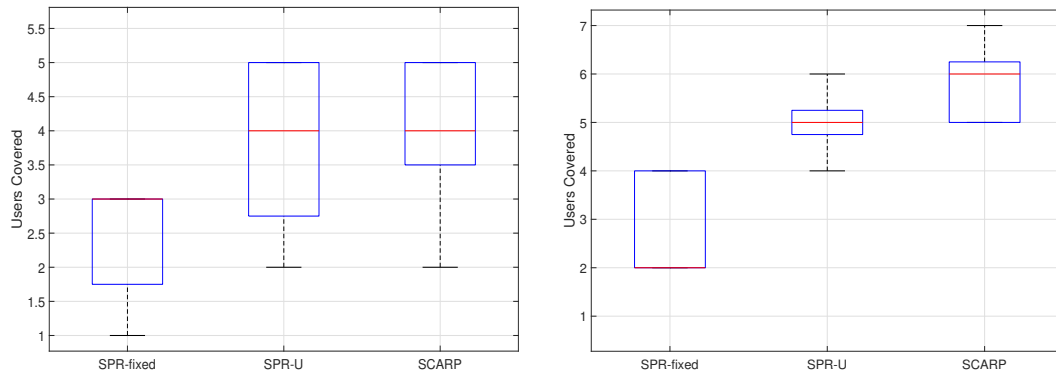
Figure 4.5: Users covered w.r.t. number of D2D sources

for a network with less number of D2D sources. However, the proposed SCARP algorithm can be implemented for large scale networks as well. The results shown in this section have been averaged over 100 network realizations. Network graph for a specific network realization with 9 BSs, 4 D2D sources, and D2D relays with density $20/\text{km}^2$ is shown in Fig. 4.3.

4.4.2 Performance Evaluation of SCARP

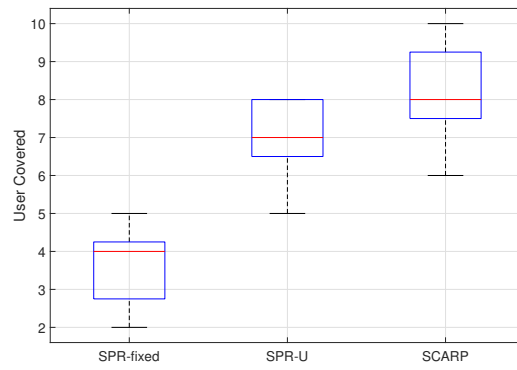
In the following, we have analyzed the processing time for the optimal solution using CPLEX solver and proposed SCARP algorithm. Fig. 4.4 compares the processing time for the optimal solution and SCARP. It can be observed that with the increase in the density of the D2D relays from $40/\text{km}^2$ to $80/\text{km}^2$ the processing time for the optimal solution is increasing rapidly. This is due to the fact that with the increase in the number of D2D relays in the network, the number of edges in the graph will increase. Consequently, the constraints in the formulated problem (4.11) will have to account for increasing number of nodes and edges. The optimal solution's processing time will increase exponentially. Unlike the exponential processing time for the optimal solution, the SCARP algorithm has a polynomial processing time.

Fig. 4.5 shows the plot for the number of users that can be provided access to the cellular network in the dead spot on using SCARP and the optimal solution. It can be observed that with an increase in the number of D2D sources from 2 to 5 in the dead spot, the number of users that can access the cellular network increases for both SCARP and optimal solution. However, there is a slight performance gap (up to 5.5%) between SCARP and optimal solution. This slight degradation in performance comes with a reduction of about 92% (as shown in Fig. 4.4) in the processing time on using SCARP. Therefore, SCARP provides a good low complexity alternative solution for the optimization problem.



(a)

(b)



(c)

Figure 4.6: Box plot with $|DS| = 10$, $T_y = 4$ and density of D2D relays (a) $40/\text{km}^2$ (b) $50/\text{km}^2$ and (c) $60/\text{km}^2$.

4.4.3 Comparison With Existing Work

As mentioned in Section 4.3, SPR used in the prior works on multi-hop D2D for a disaster scenario results in a minimum number of hops. For a time-bounded communication network considered in this work, SPR may appear to be efficient; however, it does not consider the necessary transmission/reception constraints (4.6)-(4.8). In SPR based scheduling, it is highly likely that D2D sources are assigned overlapping shortest paths to an active BS j . However, at the time of link scheduling, taking the transmission/reception constraints into account, only a few D2D sources with overlapping paths will access the cellular network within the desired deadline. Herein, we will compare the performance of SPR based scheduling with the proposed framework that utilizes SCARP. In our evaluation, we compare three approaches, namely: "SPR-fixed" (SPR with fixed pairings), "SPR-U" (SPR with unknown pairings) and "SCARP". In SPR-fixed, the disaster-affected area is divided into $|B|$ pie-shaped areas. These $|B|$ pie-shaped areas are assigned one BS each. Hence, a D2D source lying in a given area will be paired with the BS assigned to that area. In SPR-U, as the source-destination pairs are not known beforehand, a D2D source m is paired with the destination $j \in B$, where j results in a minimum hop path to m among all the BSs in B_s .

Fig. 4.6 presents the box plots to illustrate the impact of D2D relay density on the number of D2D sources covered in the dead spot for SPR-fixed, SPR-U, and SCARP at $T_s = 4$ and $|DS| = 10$. It can be seen that with the increase in the density of D2D relays, the gap between the range of D2D sources covered in SPR-fixed, SPR-U, and SCARP increases. SPR-fixed and SCARP perform the worst and best, respectively. This is because, in SCARP, due to the increased presence of D2D relays, the feasibility of non-overlapping paths from D2D sources to an active BS will increase.

Let us now analyze the performance of the above three approaches with respect to T_s . Fig. 4.6 (b) ($T_s = 4$) and Fig. 4.7 ($T = 5$) demonstrate the impact of T_s on the network set-up. With the increase in deadline T_s , the gap in the range of the number of users covered users decreases. However, the average number of users covered in SPR-fixed and SPR-U is still lower than SCARP.

The gain in the number of users covered using SCARP over SPR-U is small at higher T_s . However, there is another advantage of SCARP. It results in lower energy consumption at user devices on using SCARP at higher T_s as compare to SPR-U. To demonstrate this, we evaluate the energy consumption for SPR-U and SCARP. The expression for the average energy consumption per active D2D relay within T_s slots can be given as follows:

$$T_{energy} = \frac{\sum_l \omega_l \delta_E}{\delta_E |D_a|}, \forall l \in D_a \quad (4.12)$$

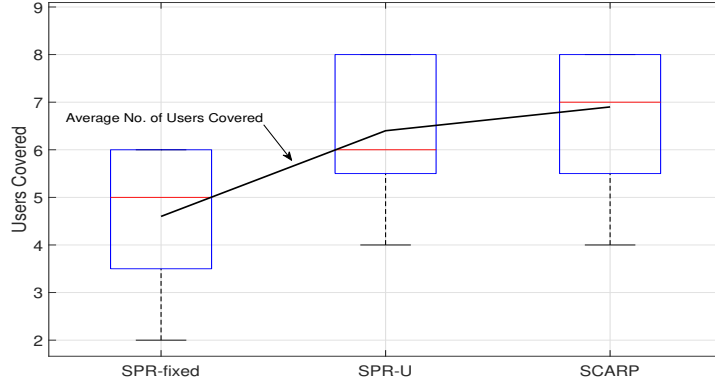


Figure 4.7: Box plot with $|DS| = 10$ at $T_s = 5$ and density of D2D relays = $50/\text{km}^2$.

where δ_E is the energy consumed per transmission and reception of a packet at a D2D relay. D_a is the set of active D2D relays within T_s slots. ω_l is the number of times relay, $l \in D_a$ is re-used within T_s . In Fig. 4.8, the plot for average energy consumption (normalized w.r.t. δ_E) per active D2D relay node within T_s slots has been presented w.r.t. T_s , where T_s varies from 3 to 6, for SCARP and SPR-U. With an increase in T_s , both SCARP and SPR-U will increase the number of users gaining access to the cellular network via multi-hop D2D. This is because the routing paths with more number of hops can be scheduled easily within higher T_s . Despite the above fact, it can be seen from Fig. 4.8 that the average energy consumption per active D2D relay is lower for the proposed SCARP algorithm as compared to SPR-U. This can be explained as follows. As mentioned above, SPR-U tends to assign overlapping shortest paths to the D2D sources. With the increase in T_s , the chances of overlapping paths being scheduled within T_s will increase. This will result in the re-use of relays in overlapping paths. In the case of SCARP, we minimize the selection of overlapping paths. Consequently, SPR-U has up to 21% $\left(= \frac{(1.63-1.35) \times 100\%}{1.35} \right)$ more energy consumption at higher T_s . Instead of re-using a set of D2D relays repeatedly, the proposed SCARP algorithm limits the overlap of the set of D2D relays corresponding to each data flow and hence results in spatially distributed energy consumption. This is beneficial for enhancing the longevity of the multi-hop D2D networks deployed for disaster-resilient communication.

Further, at lower values of T_s (such as $T_s = 3$), the chances of overlapping paths being scheduled within T_s are less, i.e., D2D sources covered are low, and the re-use of D2D relays may not be possible. Consequently, the average energy consumption per active D2D relay for SCARP is comparable at lower values of T_s .

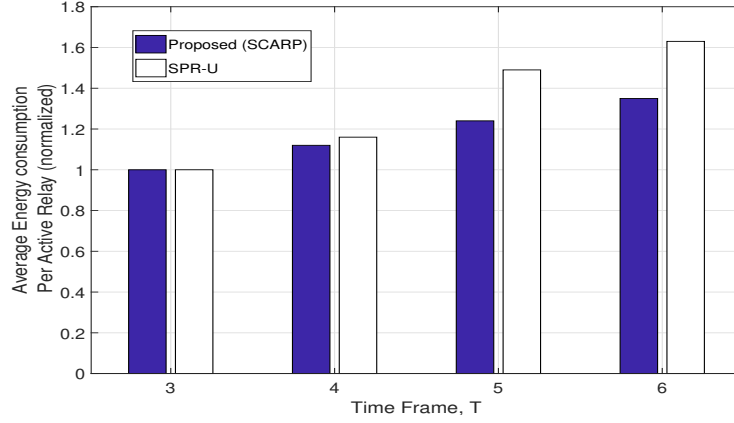


Figure 4.8: Average energy per active D2D relay node

4.5 Conclusion

We proposed a novel joint pairing, routing and scheduling framework for a disaster-resilient communication network based on multi-hop D2D. An optimization problem to maximize the number of users covered in the dead spot within a deadline of T time slots is formulated. The formulated problem is shown to be NP-hard. Hence, we proposed SCARP that shows a 92% reduction in processing time as compared to the optimal solution with an acceptable performance gap of 5.5% in terms of maximum users covered. It is also demonstrated that the proposed algorithm outperforms the widely used SPR based scheduling in terms of users covered at lower values of T (tighter deadline). This is because the D2D relays selected using SCARP are more spatially distributed which reduces wireless link contention at lower values of T . Further, the gain in users covered on using SCARP w.r.t. SPR based scheduling becomes more prominent with an increase in the density of D2D relays.

UAV-BS Placement in Aerial Communication Networks

Recent studies have demonstrated that the communication coverage in the disaster scenarios can also be enhanced by employing UAVs which helps in establishing disaster-resilient communication networks due to their flexibility and maneuverability [20]. Moreover, UAVs are capable of offering low-latency communication services. With the recent technological advances, it is feasible to deploy UAVs for the roles of flying/aerial BSs as well as aerial UEs [20]. Specifically, on using UAVs as flying BSs, the key challenge is the optimal 3-D placement of the UAVs for efficient network performance [54]. The 3-D placement of UAVs has drawn a lot of attention from the researchers over the last few years. For instance, authors in [55] have dealt with maximization of covered users via optimal 3-D placement of a single UAV. In [56], the optimal 3-D deployment of multiple UAVs is investigated to maximize the downlink coverage performance with minimum transmit power. In [59] and [60], 3-D placement UAV-BSs is studied for maximizing the sum logarithmic rate of the users and effectively prolonging the life-time of the network, respectively. Further, they analyzed the network performance for different user distributions such as Poisson point process and clustered user distribution.

However, the current literature on 3-D UAV placement has generally overlooked the mobility aspects of the ground users in a disaster scenario. Further, in order to facilitate UAV enabled communication network, a practical fly-hover-and-communicate protocol has been proposed in [101]. In this protocol, UAVs primarily have two modes of operation: flight and hover/communicate. In the hovering or communication mode, UAVs can provide coverage to the ground users, whereas in the flight mode, there will be no coverage. Hence, with respect to the UAV placement problem, if the next optimal UAV location is far away from the current UAV location the flight time of UAVs will increase. As stated in [102], the delay in UAV to ground communication networks is primarily due to the UAV flight time, which in turn

lowers the coverage time. Consequently, the flight time of UAVs is a crucial parameter in UAV placement problems. Hence, we propose a UAV based disaster-resilient communication network and optimize the 3-D placement of multiple UAVs in order to maximize the number of ground users covered in a disaster-affected area while taking into account the UAV flight time constraint. Unlike the prior works on UAV placement in a disaster scenario, the proposed strategy considers a disaster-specific ground user mobility model. Specifically, in this work, the emergency first responders (EFRs) are considered as the ground users and their mobility is modeled in the disaster-affected area. We demonstrate the results for single-UAV and multi-UAV scenarios in terms of average number of covered users and average coverage time in a disaster-resilient network. To the best of our knowledge, the proposed work is first of its kind for multi-UAV placement which utilizes a disaster-specific user mobility model and also considers a flight time constraint for UAVs. This work is discussed in part I of this chapter (Sections 5.2 to 5.4).

However, the part I of this chapter does not consider user coverage fairness. Further, the user mobility model is assumed to be synchronous, i.e., the users walk and pause at the same time. It may be noted that, in most of the prior works, the placement optimization of mobile UAV-BSs is generally carried out at specific time instants. Moreover, the user mobility behavior will play a key role in determining the time separation between two consecutive UAV-BS placement updates (or update interval). However, the prior works fail to characterize the relationship between user mobility and update interval. For example, [72] and [75] assume a fixed update interval whereas [73, 74] lack in quantifying such an update interval. To overcome these issues, we propose a joint optimization of the UAV-BS placement and the placement update interval. To the best of our knowledge, no existing work investigates the relationship between user mobility and UAV placement update interval. Specifically, the UAV-BS placement is optimized to maximize the number of users covered at an update instant while accounting for the user fairness as well as UAV-BS flight time. We propose the use of two metrics, i.e., total UAV-BS flight time and user coverage probability to optimize the update interval. We propose an iterative approach to solve the optimal UAV-BS placement and update interval problems jointly. Further, we derive the analytical expressions for user coverage probability in terms of user mobility parameter and update interval. This work is discussed in part II of the chapter (Sections 5.5 to 5.7).

5.0.1 Terminologies

Below are the key terms used throughout this chapter:

- **Operation Period:** It is the time period during which the UAV network is operational.

- Update Instant: An update instant denotes the point in time where the UAV network is optimized.
- Update Interval: The time interval between two consecutive update instants is termed as update interval.

Now, let us discuss the optimal UAV-BS placement for disaster-resilient communication network.

5.1 Network Model

We consider a disaster scenario with damaged cellular infrastructure as shown in Fig. 5.1. For the ground users (in this work, EFRs) to communicate with each other as well as with the disaster control room, a disaster-resilient network is established which consists of multiple UAVs. There UAVs in set \mathcal{U}_B in the 3-D space that act as flying BSs. A ground user in set \mathcal{U} , present in the 2-D space, is said to be covered by a UAV if it is present within the coverage radius of the UAV. Hence, as the ground users move over time, UAV locations need to be updated in order to provide coverage to the ground users. As mentioned before, UAVs are either in flight or hover mode. It is assumed that the ground users¹ periodically update their locations on the cloud [103]. Further, it is assumed that UAVs and users are allocated independent time-frequency resources for operation.

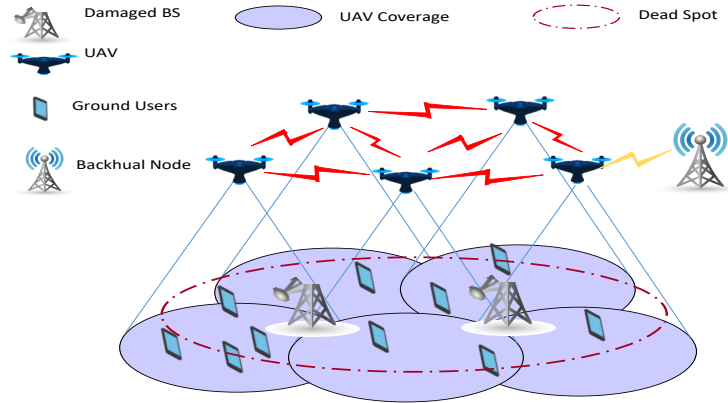


Figure 5.1: Network Model.

In case of disaster, the network topology information can be utilized by the centralized software defined network (SDN) to route the traffic of UAVs [104]. The centralized SDN controller also manages the control signals to the UAVs. Further, an active BS, closest to a

¹Specifically, EFRs can be equipped with satellite devices like satellite emergency notification device (SEND).

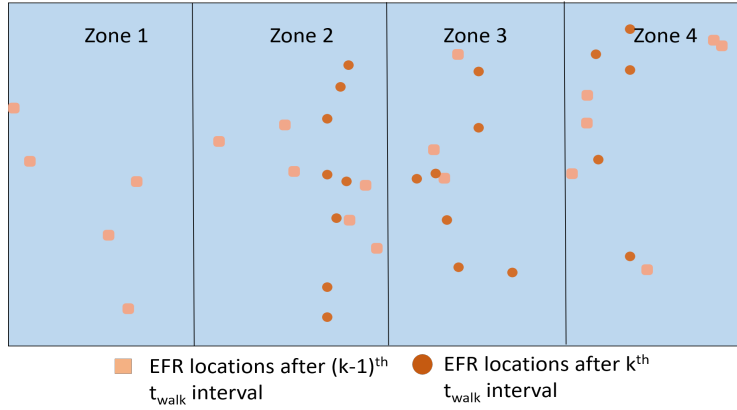


Figure 5.2: Zone-wise representation of Disaster Area

UAV deployed in the disaster area, can act as the backhual node [105]. UAVs can receive the control signals via inter-UAV² and UAV to backhual node links.

As mentioned before, the mobility of the ground users impacts the coverage performance of the network. Hence, the modelling of the ground users' mobility in a disaster scenario is required to do an efficient analysis of the UAV placement strategy.

5.1.1 Disaster Mobility Model

We adapt the mobility model for disaster area scenarios as given in [107] where the disaster area is divided into different zones based on the task to be performed in each zone. The disaster mobility model has been illustrated in Fig. 5.2. Zone 1 i.e., the incident site, consists of the disaster-affected people that need to be rescued. The affected people are taken to the zone 2 i.e., patient waiting area, where they are provided with the initial diagnosis and first aid treatment. Zone 3 i.e., the casualties clearing station, consists of the affected people who may need immediate care and hospitalization, and should be taken to the hospital. Zone 4 i.e., the technical operational command consists of the group of volunteers that strategize the search and rescue operations. Depending on the disaster rescue strategies, ground users will be assigned different zones. Similar to [107], we consider two types of ground users: stationary and transport ground users. Stationary ground users are restricted to a specific zone whereas transport ground users can move between two adjacent zones in a cyclic manner, i.e., zone z to zone $z + 1$ and back to zone z , $z \in \{1, 2, 3\}$. In our work, we assume that the ground users move in random directions and with random velocities within the assigned zones. Ground users walk from one location to another, take a pause at the new location to perform a task and then move to a different location. This means the walk time and pause time occur alternately. However, for the ease of analysis, the walk time, t_{walk} of a ground user is kept constant and

²There exists inter-UAV links as shown in Fig. 5.1 [106].

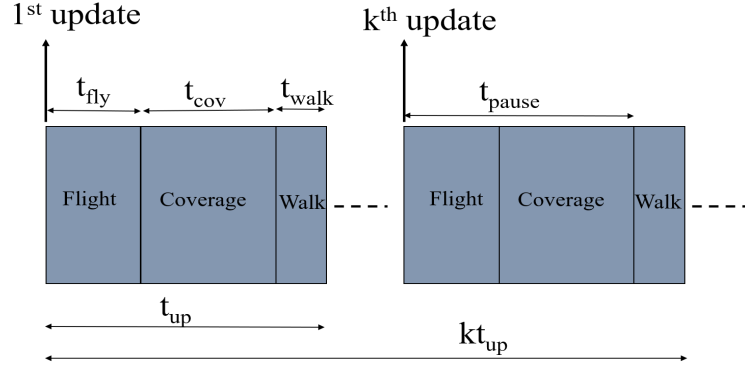


Figure 5.3: Illustration of UAV Placement timeline.

is assumed to be same for all ground users. Further, after each transition, the ground users are assumed to pause for a fixed amount of time. We term this as a synchronized mobility scenario, wherein the ground users walk and pause at the same time. As mentioned above, depending on the user mobility, UAV locations need to be updated at regular intervals [73]. Due to the transitions during t_{walk} interval user locations will be more dynamic; hence, let the UAV placement be updated at the beginning of each t_{pause} of the ground users. This will ensure that the UAV placement updates remain optimal for a longer time. Hence, the time interval between two updates, t_{up} consists of one t_{pause} and t_{walk} interval, i.e., $t_{up} = t_{pause} + t_{walk}$. Further, as mentioned above, UAV placements will not be optimal during t_{walk} interval. Hence, t_{pause} must consist of UAV flight time (t_{fly}) and coverage time (t_{cov}), where t_{cov} denotes the time within which maximum users are covered by UAVs. Therefore, in proposed analysis, two independent timelines denoting UAV operation and ground user mobility have been utilized. Specifically, UAVs' operation with parameters t_{fly} and t_{cov} , and ground users mobility with parameters t_{pause} and t_{walk} have been coupled together as a single timeline of UAV placement as illustrated in Fig. 5.3. At the k^{th} update instant, optimal UAV locations are determined for the upcoming k^{th} pause time in the network. It may be noted that unlike t_{pause} and t_{walk} , t_{fly} and t_{cov} are random variables and their values depend on the optimal UAV placements. Hence, in the optimization problem, the permissible limit of t_{fly} must be set less than t_{pause} to have a non-zero t_{cov} .

5.2 Problem Formulation

The aim is to find the optimal placement of the UAVs to maximize the covered users such that t_{fly} constraint is met. The coordinates of the user $j \in \mathcal{U}$ (represented by 'square' in Fig. 5.4) at the k^{th} update instant are $(w_{j,x}(k), w_{j,y}(k))$. Let the optimal coordinates of the UAV $i \in \mathcal{U}_{\mathcal{B}}$ after the k^{th} update instant be $(u_{i,x}(k), u_{i,y}(k), u_{i,z}(k))$ where $(u_{i,x}(k), u_{i,y}(k))$ denotes the

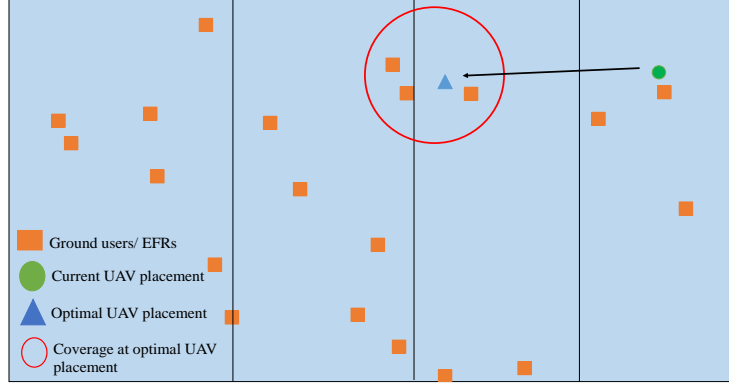


Figure 5.4: 2-D illustration of the ground users along with a UAV's current and new placement.

horizontal placement of UAV and $u_{i,z}(k)$ denotes the UAV altitude in the 3-D space. However, for a given quality of service (QoS) requirement, as shown in [57], to maximize the coverage evaluation of the optimal altitude, $u_{i,z}(k)$ can be decoupled from the horizontal placement, $(u_{i,x}(k), u_{i,y}(k))$ of UAV in the 2-D plane. The QoS requirement is the pathloss, PL_{th} which should not be exceeded³. Let all the UAVs be present at the optimal altitude, z_{opt} and have a maximum coverage radius R . Hence, we will focus on the 2-D placement of UAVs in the horizontal plane where $u_{i,x}(k)$ and $u_{i,y}(k)$ are continuous variables.

In Fig. 5.4, ‘circle’ represents the current 2-D placement of a UAV, $(u_{i,x}^*(k-1), u_{i,y}^*(k))$ which was optimal for the $(k-1)^{th}$ t_{int} interval. The ‘triangle’ denotes the newly selected placement of a UAV at the k^{th} update instant. Let I_{ij} be the distance between j^{th} user's location at the k^{th} update instant and i^{th} UAV's location selected at the k^{th} update instant i.e., the distance between the triangle and each of the squares in Fig. 5.4. Also, I_i denotes the distance between the i^{th} UAV's optimal location at the $(k-1)^{th}$ update instant and i^{th} UAV's location selected at the k^{th} update instant⁴.

$$I_{ij} = \sqrt{(w_{i,x}(k) - u_{i,x}(k))^2 + (w_{i,y}(k) - u_{i,y}(k))^2} \quad (5.1)$$

$$I_i = \sqrt{(u_{i,x}(k) - u_{i,x}^*(k-1))^2 + (u_{i,y}(k) - u_{i,y}^*(k-1))^2} \quad (5.2)$$

The formulated optimization problem at the k^{th} update instant is given in (5.3a)-(5.3e). (5.3a) is the number of users covered by the UAVs where $\mathbb{1}()$ is an indicator function that is ‘1’ when a user is within the coverage of a UAV. (5.3b) is the flight time constraint that ensures that $t_{fly} = I_i/V_{uav}$ is less than or equal to t_{fly}^{max} where t_{fly}^{max} is the permissible limit on t_{fly} . The

³For the air-to-ground (AtG) channel model and optimal altitude evaluation interested readers may refer to [57].

⁴It is assumed while travelling from one location to another UAVs follow a straight line trajectory.

constraint (5.3c) is applied such that a ground user is only served by a single UAV. The UAV's x and y coordinates should be within the range $[x_{min}, x_{max}]$ and $[y_{min}, y_{max}]$ as accounted for in (5.3d) and (5.3e), respectively.

$$\begin{aligned} & \text{maximize}_{\{u_{i,x}(k), u_{i,y}(k)\}} \sum_i \sum_j \mathbb{1}(I_{ij} \leq R), \quad i \in \mathcal{U}_{\mathcal{B}}, j \in \mathcal{U} \end{aligned} \quad (5.3a)$$

$$\text{subject to} \quad \frac{I_i}{v_{uav}} \leq t_{fly}^{max}, \quad \forall i \in \mathcal{U}_{\mathcal{B}}, \quad (5.3b)$$

$$\sum_i \mathbb{1}(I_{ij} \leq R) \leq 1, \quad \forall j \in \mathcal{U}, i \in \mathcal{U}_{\mathcal{B}}, \quad (5.3c)$$

$$x_{min} \leq u_{i,x}(k) \leq x_{max}, \quad \forall i \in \mathcal{U}_{\mathcal{B}}, \quad (5.3d)$$

$$y_{min} \leq u_{i,y}(k) \leq y_{max}, \quad \forall i \in \mathcal{U}_{\mathcal{B}} \quad (5.3e)$$

The objective function (5.3a) and constraint (5.3c) can be re-formulated by introducing a binary decision variable, Y_j^i to replace the indicator function as shown in (5.4a) and (5.4d), respectively. Consequently, a new constraint (5.4b) is introduced to take care that the condition $I_{ij} \leq R$ is met. Here, M is a large constant. If $I_{ij} \leq R$ then Y_j^i will have value 1. Otherwise, Y_j^i will be 0. Specifically, Y_j^i is an association variable which decides the users to be served by a UAV. The constraints (5.4e) and (5.4f) are same as constraints (5.3d) and (5.3e), respectively. Constraint (5.4g) is to ensure that Y_j^i has binary values. The optimization problem consists of continuous variables $u_{i,x}(k)$, $u_{i,y}(k)$ and binary/integer variables Y_j^i . Further, the objective function (5.4a) is linear and the constraints (5.4b) and (5.4c) are quadratic. Hence, this optimization problem is a mixed-integer quadratically constrained problem (MIQCP) [108]. The formulated problem (5.4a)-(5.4g) can be solved using IBM CPLEX solver because, in our case, the objective function is linear and the quadratic constraints are convex.

$$\begin{aligned} & \text{maximize}_{\{u_{i,x}(k), u_{i,y}(k), Y_j^i\}} \sum_i \sum_j Y_j^i, \quad i \in \mathcal{U}_{\mathcal{B}}, j \in \mathcal{U} \end{aligned} \quad (5.4a)$$

$$\text{subject to} \quad I_{ij} - (1 - Y_j^i)M \leq R, \quad \forall i \in \mathcal{U}_{\mathcal{B}}, \forall j \in \mathcal{U}, \quad (5.4b)$$

$$\frac{I_i}{v_{uav}} \leq t_{fly}^{max}, \quad \forall i \in \mathcal{U}, \quad (5.4c)$$

$$\sum_i Y_j^i \leq 1, \quad \forall j \in \mathcal{U}, i \in \mathcal{U}_{\mathcal{B}}, \quad (5.4d)$$

$$x_{min} \leq u_{i,x}(k) \leq x_{max}, \quad \forall i \in \mathcal{U}, \quad (5.4e)$$

$$y_{min} \leq u_{i,y}(k) \leq y_{max}, \quad \forall i \in \mathcal{U}, \quad (5.4f)$$

$$Y_j^i \in \{0, 1\}, \quad \forall i \in \mathcal{U}_{\mathcal{B}}, \forall j \in \mathcal{U} \quad (5.4g)$$

Table 5.1: User Placement in Disaster-Affected Area

	Users = 10		Users = 20	
	Stationary	Transport	Stationary	Transport
Zone 1	0	2	0	8
Zone 2	2	2	4	2
Zone 3	2	0	3	0
Zone 4	2	0	3	0

Table 5.2: Simulation Parameters

Parameter	Value
t_{pause}	60 seconds
t_{walk}	40 seconds
PL_{th}	83, 85 and 87 dB (Dense urban environment)
v_{uav}	8 to 18 m/s
v_u	[2, 3] m/s

5.3 Simulation Results

In this section, we analyze the performance of the proposed multi-UAV placement strategy for disaster-resilient communication network. We consider a disaster area of size 900×500 sq. m in a dense urban environment. Zone 1 has dimensions 300×500 sq. m and rest of the zones have dimensions 200×500 sq. m each. Table 5.1 states the zone wise assignment of the ground users considered in our analysis. A ground user is either a stationary or transport node as explained in Section 5.1. The other simulation parameters are provided in Table 5.2. We assume that $t_{pause} = 60$ seconds, $t_{walk} = 40$ seconds and the user speed, v_u is uniformly distributed in [2, 3] m/s. A ground user is covered by a UAV when the pathloss does not exceed PL_{th} . Further, UAVs are assumed to fly at a constant speed, v_{uav} . We consider $v_{uav} = 18$ m/s unless otherwise stated [23]. The simulation results are averaged over two hours of UAV operation in the disaster-affected area.

We know from Section 5.1 that t_{fly}^{max} must be less than t_{pause} to have a non-zero t_{cov} . Let us analyze the network for a range of t_{fly}^{max} by setting $t_{fly}^{max} = \rho t_{pause}$ where $\rho = \{0.1, 0.2, \dots, 0.9\}$.

Fig. 5.5 exhibits the performance of the proposed UAV placement strategy with a single UAV. It can be observed that with increase in ρ or t_{fly}^{max} , the average number of covered users increases whereas the average coverage time (\bar{t}_{cov}) decreases. This is because an increase in ρ provides a greater flexibility for UAV movement in order to increase the number of covered users; hence, a decrease in \bar{t}_{cov} . There exists a trade-off between the average number of covered users and \bar{t}_{cov} . However, the trade-off becomes insignificant at higher ρ as the rate of increase of the average number of covered users starts to decrease at higher ρ . For instance, with 10

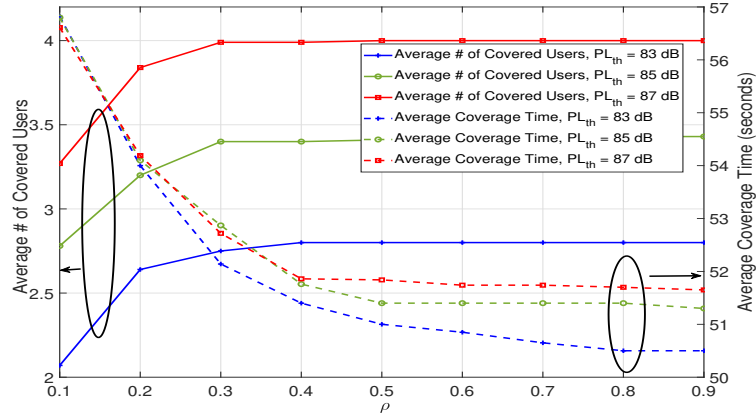


Figure 5.5: Performance of proposed UAV Placement Strategy with 10 users and single UAV, $t_{pause} = 60$ seconds.

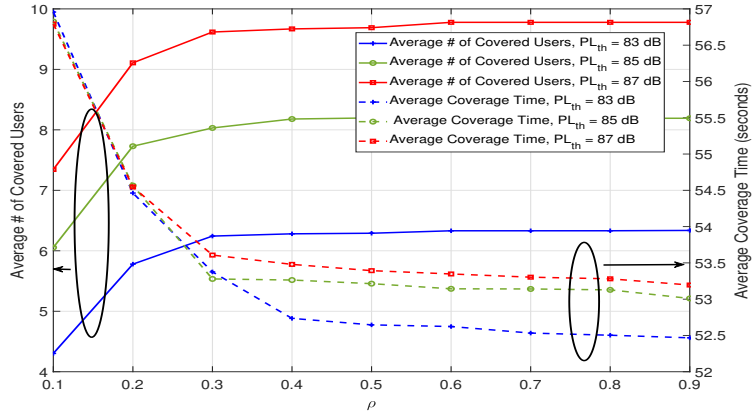


Figure 5.6: Performance of proposed UAV Placement Strategy with 20 users and single UAV, $t_{pause} = 60$ seconds.

users, at $PL_{th} = 83$ dB, the rate of increase of the average number of covered users starts to decrease after $\rho = 0.3$. However, the average coverage time keeps on decreasing at a faster rate. Lower \bar{t}_{cov} implies that the covered users are served by the UAVs for shorter time duration; hence, decreases the network resilience temporally. With efficient selection of the permissible limit of the UAV flight time reduction in \bar{t}_{cov} is feasible. Further, it can also be observed from Fig. 5.5 that \bar{t}_{cov} increases with the increase in PL_{th} . This is because at higher PL_{th} the maximum coverage radius of UAV increases. Consequently, on an average, UAV does not need to displace a lot from its current location and \bar{t}_{cov} is higher. It can also be observed from Fig.5.5 that with increase in PL_{th} the number of covered users increase. Fig. 5.6 exhibits the network performance with 20 users. With increase in the number of users in the network it is obvious that the average number of covered users will increase as seen in Fig. 5.6. Further, similar to the case with 10 users, the average number of covered users increases with ρ as well as PL_{th} . Figs. 5.7, 5.8 and 5.9 present results for two UAVs with 10 users and 20 users, and three UAVs

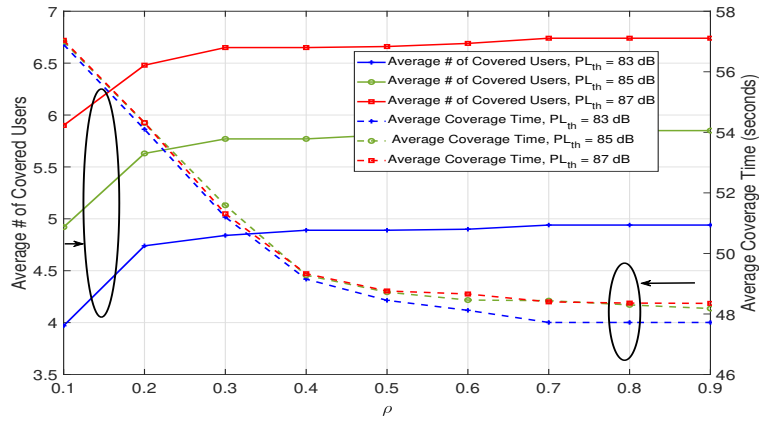


Figure 5.7: Performance of proposed multi-UAV placement strategy with 10 users, $t_{pause} = 60$ seconds, 2 UAVs.

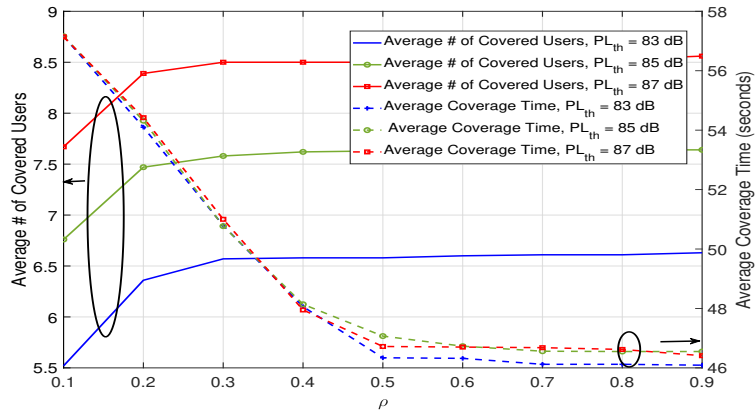


Figure 5.8: Performance of proposed multi-UAV placement strategy with 20 users, $t_{pause} = 60$ seconds, 2 UAVs.

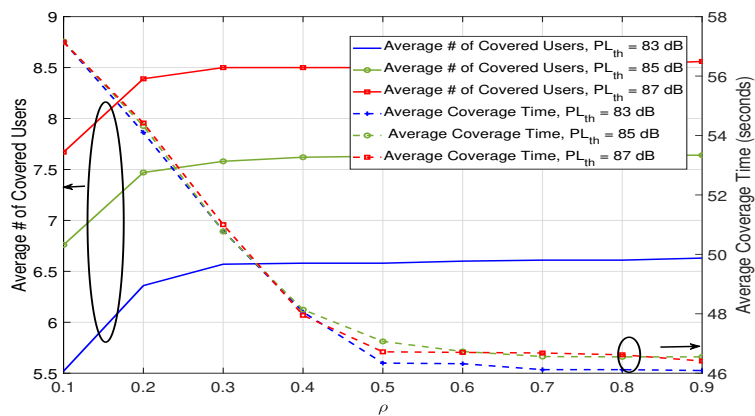


Figure 5.9: Performance of proposed multi-UAV placement strategy with 10 users, $t_{pause} = 60$ seconds, 3 UAVs.

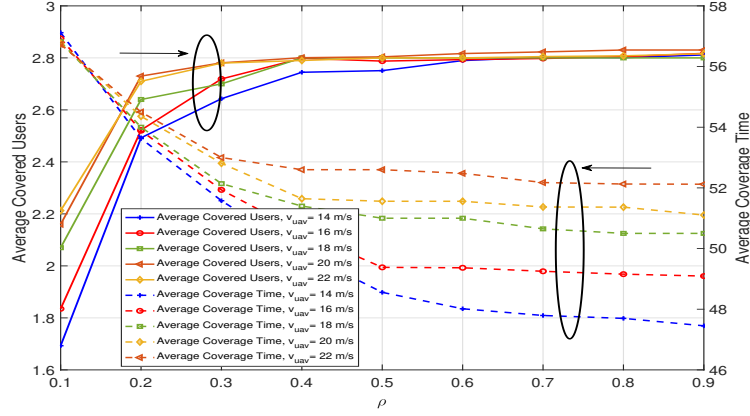


Figure 5.10: Performance of proposed multi-UAV Placement Strategy with 10 users, $t_{pause} = 60$ seconds, single UAV and v_{uav} ranges from 14 m/s to 22 m/s.

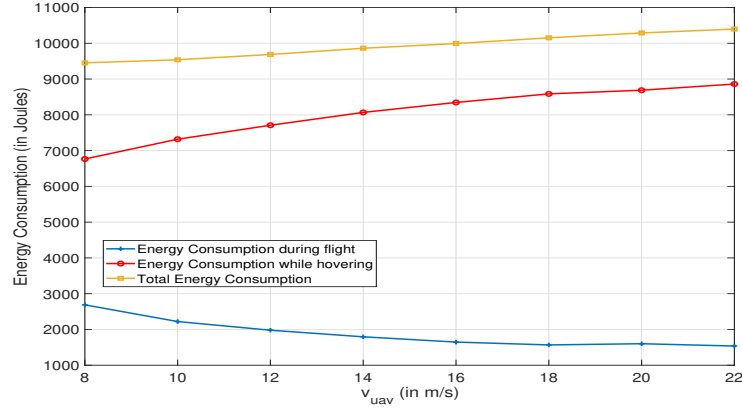


Figure 5.11: Performance of proposed multi-UAV Placement Strategy with 10 users, $t_{pause} = 60$ seconds, single UAV and v_{uav} ranges from 14 m/s to 22 m/s.

with 10 users, respectively. It can be observed that for the same ρ and PL_{th} average number of covered users increases with the number of UAVs. This is because, with increase in UAVs, the total area covered by the UAVs increases. Further, similar to the single UAV case, the rate of increase of the average number of covered users starts to decrease at higher ρ . In case of multiple UAVs, due to the (5.4d) constraint which requires no two UAVs to associate with the same user, UAVs need to fly a longer distance. Hence, \bar{t}_{cov} is lower as compared to single UAV. Further, we have also analyzed the proposed strategy for different values of V_{uav} . Fig. 5.10 demonstrates the plot for average covered user and \bar{t}_{cov} at $PL_{th} = 83$ dB with an increase in ρ . It can be observed that at higher values of ρ , irrespective of v_{uav} , average number of covered users saturates to the same value. However, higher the UAV velocity higher is \bar{t}_{cov} . Given this information, it may seem advantageous to have a higher UAV velocity. However, this may not hold true. Based on the power consumption model for rotary wing UAV, a promising contender for disaster scenarios, average energy consumption per t_{pause} interval at a specific value of V_{uav}

can be evaluated as follows⁵ [23]:

$$E_{total}(v_{uav}) = P_h \bar{t}_{cov} + P_f(v_{uav}) \bar{t}_{fly} \quad (5.5)$$

where P_h is the power consumption due to hovering and $P_f(V_{uav})$ is the power consumption due to flight of UAV. The energy consumption is plotted in Fig. 5.11. On increasing v_{uav} , there is a reduction in \bar{t}_{fly} and an increase in $P_f(v_{uav})$. However, the impact of \bar{t}_{fly} is dominant. Consequently, the energy consumption due to flight decreases with v_{uav} . Further, on increasing V_{uav} , \bar{t}_{cov} i.e., the hovering time during t_{pause} will increase and will result in increase in energy consumption due to hovering. Therefore, the overall energy consumption at the UAV increases.

Now, we will focus on the joint optimization of the UAV-BS placement and the placement update interval.

5.4 Network Model and Problem Formulation

We consider a network model with one UAV-BS and U multiple ground users, as shown in Fig. 5.12. The UAV-BS operates for T seconds to serve the ground users. The users are moving around following a random walk mobility model. The distance traveled by a user in each transition of random walk is assumed to be Rayleigh distributed with shape parameter σ . For analytical tractability, it is assumed that the velocity of the users is constant and same for all the users. The velocity of a user is denoted by v_u . Further, in line with the existing literature, it is assumed that the altitude of the UAV-BS is fixed, and it is denoted by H [70]. Hence, we focus only on the 2-D UAV-BS placement in the horizontal plane. The desired quality-of-service⁶ (QoS) determines the coverage radius R of the UAV-BS. Consequently, at any time instant, due to the user mobility and fixed R , the UAV-BS may not be able to cover all the users. So, UAV-placement needs to be updated after certain time intervals.

In our work, an update instant denotes the point in time where the UAV-BS placement is updated. Fig. 5.13 illustrates the UAV-BS placement update timeline for T seconds. At update instant k , $t_{up}(k)$ denotes the update interval. In other words, $t_{up}(k)$ is the time interval between the k^{th} and $(k+1)^{th}$ update instants. In the proposed work, at update instant k , we optimize the UAV-BS placement in order to maximize the number of users covered subject to the fairness of coverage. The 2-D UAV-BS placement at the k^{th} update instant is denoted as $[u_x(k), u_y(k)]$

⁵Energy consumption due to communication is negligible as compared to propulsion energy consumption; hence, not considered here [23]. Further, during t_{walk} UAV will always be hovering and will have a constant energy consumption irrespective of V_{uav} ; hence, not considered here.

⁶In our work, path-loss is utilized as a QoS metric. We compute the path-loss using the air-to-ground (AtG) path loss model in [70].

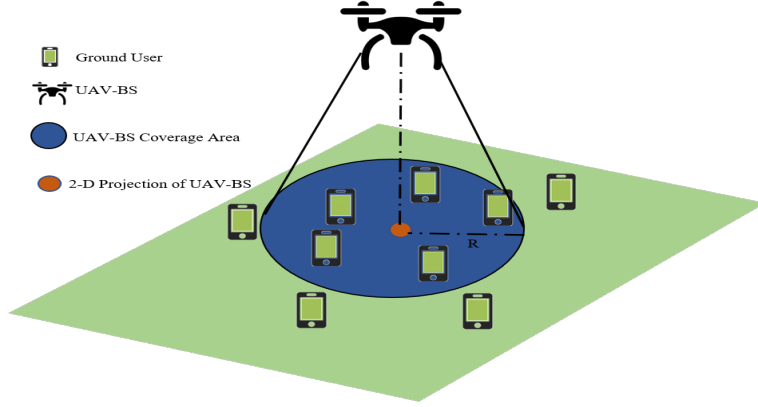


Figure 5.12: Network Model

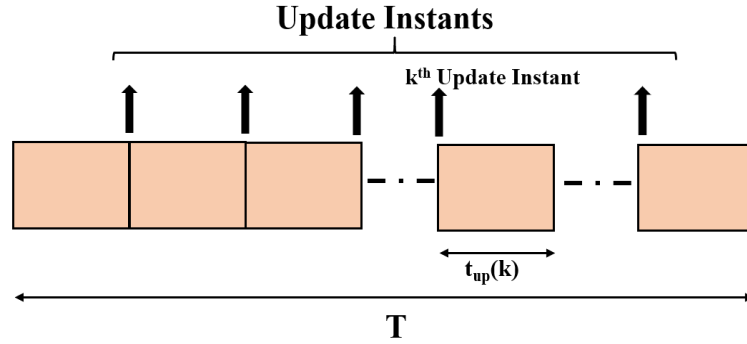


Figure 5.13: UAV-BS Placement Timeline

where $u_x(k)$ and $u_y(k)$ are the coordinates of the projection of UAV-BS in the 2-D plane. Let $A_j(k)$ be the indicator variable. If user j is covered at the update instant k , $A_j(k) = 1$ otherwise $A_j(k) = 0$. Since the users are moving, there is a possibility that some users get covered more often as compared to others. Hence, there is a need to maintain fairness of coverage. We define fairness at update instant k as follows [70]:

$$fair(k) = \frac{\left(\sum_{j=1}^U \sum_{l=0}^{l=k} A_j(l)\right)^2}{U \left(\sum_{j=1}^U \left(\sum_{l=0}^{l=k} A_j(l)\right)^2\right)} \quad (5.6)$$

Further, we consider the fly-hover communicate protocol for UAV-BS operation [23, 109]. This implies that UAV-BS will only serve users when it is hovering. Let $F(k)$ be the time taken by the UAV-BS to fly to its new location at k^{th} update and can be given as:

$$F(k) = \frac{1}{v_{uav}} \sqrt{(u_x(k) - u_x(k-1))^2 + (u_y(k) - u_y(k-1))^2} \quad (5.7)$$

where v_{uav} is the UAV-BS velocity.

Now, the optimal UAV-BS placement problem at k^{th} update instant can be formulated as follows:

$$(P1) \quad \max \sum_{j=1}^U A_j(k) \quad (5.8)$$

$$\text{s.t. } fair(k) > fair_{thres} \quad (5.9)$$

$$F(k) < t_{up}(k) \quad (5.10)$$

Here, (5.8) represents the total number of users covered at update instant k . Constraint in (5.9) is used to maintain a fairness above a threshold of $fair_{thres} \in [0, 1]$ in the network. To ensure non-zero service time during the upcoming $t_{up}(k)$ seconds, $F(k)$ must be less than $t_{up}(k)$. This has been applied using the constraint given in (5.10). For a given $t_{up}(k)$, (P1) is solved exhaustively. The exhaustive search details are provided in Section 5.5.

Now, the next problem is to find the optimal value of $t_{up}(k)$. The choice of $t_{up}(k)$ depends on two factors: 1) total UAV flight time and 2) temporal user coverage probability. As mentioned above, an increase in the UAV-BS flight time decreases its service time. Hence, minimization of total UAV-BS flight time during T seconds must be considered when optimizing $t_{up}(k)$. In the following, we will discuss in detail the temporal coverage probability metric.

5.4.1 Temporal Coverage Probability

The coverage probability of a user must capture the change in user locations over time. We define temporal coverage probability of user n , which is covered after the k^{th} update, as the probability of it being within UAV-BS coverage area for the upcoming $t_{up}(k)$ seconds. During these $t_{up}(k)$ seconds user may have multiple transitions. The number of transitions depends on σ and v_u . In other words, coverage probability can be defined as the probability that the displacement of the user from UAV-BS must be less than R for all the transitions within $t_{up}(k)$.

Let D_i denote the displacement of a user from UAV-BS after the i^{th} transition during its random walk. D_i is given as follows:

$$D_n = \sqrt{(X_{in} + X_1 + \dots + X_n - u_x(k))^2 + (Y_{in} + Y_1 + \dots + Y_n - u_y(k))^2}, \quad (5.11)$$

where X_{in} and Y_{in} are the initial coordinates of a user at k^{th} update instant. Further, $X_i, Y_i \sim \mathcal{N}(0, \sigma^2)$ are the change in the x and y coordinates of a user due to the i^{th} transition. The coverage probability will be as follows:

$$P_{j,t_{up}(k)} = P[D_1 < R]P[D_2 < R|D_1 < R] \cdots P[D_n < R|D_{n-1} < R] \cdots P[D_{q_{t_{up}(k)}} < R|D_{q_{t_{up}(k)}-1} < R], \quad (5.12)$$

where $q_{t_{up}(k)}$ denotes the average number of user transitions during $t_{up}(k)$.

Lemma 1. *When the distance traveled in each user transition is governed by a Rayleigh distribution, with shape parameter σ , $q_{t_{up}(k)}$ is given as:*

$$q_{t_{up}(k)} = \frac{t_{up}(k)v_u}{\sigma\sqrt{\pi/2}}. \quad (5.13)$$

Proof. During $t_{up}(k)$, the distance traveled by user moving with velocity v_u is $v_u t_{up}(k)$. Further, the average distance traveled by a user in each transition of random walk is $\sigma\sqrt{\pi/2}$. Hence, the average number of user transitions during $t_{up}(k)$ will be as follows:

$$q_{t_{up}(k)} = \frac{t_{up}(k)v_u}{\sigma\sqrt{\pi/2}}.$$

□

Let $W_1 = X_1 + X_{in} - u_x(k)$ and $W_2 = Y_1 + Y_{in} - u_y(k)$. Hence, $W_1 \in \mathcal{N}(\mu_1, \sigma^2)$ and $W_2 \in \mathcal{N}(\mu_2, \sigma^2)$ where $\mu_1 = X_{in} - u_x(k)$ and $\mu_2 = Y_{in} - u_y(k)$. Now, the first term on the R.H.S in (5.12) can be evaluated as follows:

$$\begin{aligned} P[D_1 < R] &= P\left[\sqrt{W_1^2 + W_2^2} < R\right] \\ &= \int_{-R}^R \int_{-C_1}^{C_1} f_{W_2 W_1}(w_2, w_1) dw_2 dw_1 \end{aligned} \quad (5.14)$$

where $f_{W_2 W_1}(w_2, w_1)$ is the joint pdf of W_1 and W_2 , and $C_1 = \sqrt{R^2 - w_1^2}$. Since X_1 and Y_1 are independent, W_1 and W_2 are independent. Now, (5.15) can be written as follows:

$$\begin{aligned} P[D_1 < R] &= \int_{-R}^R \int_{-C_1}^{C_1} f_{W_2}(w_2) f_{W_1}(w_1) dw_2 dw_1 \\ &= \frac{1}{2\pi\sigma^2} \int_{-R}^R \int_{-C_1}^{C_1} e^{-\frac{(w_1 - \mu_1)^2}{2\sigma^2}} e^{-\frac{(w_2 - \mu_2)^2}{2\sigma^2}} dw_2 dw_1 \end{aligned} \quad (5.15)$$

For the n^{th} term on the R.H.S. in (5.12), where $n > 1$, substitute $X = X_{in} - u_x(k) + X_1 + \dots + X_{n-1}$, $Y = Y_{in} - u_y(k) + Y_1 + \dots + Y_{n-1}$. The n^{th} term of the product in (5.12) will be:

$$P[D_n < R | D_{n-1} < R] = \frac{\int_{-R}^R \int_{-R-u}^{R-u} \int_{-C_2}^{C_2} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{v^2}{2\sigma^2}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} \Phi\left(\frac{C_3}{\sqrt{2}\sigma}\right) dw dv du}{\sigma \sqrt{2\pi} \int_{-R}^R \int_{-C_2}^{C_2} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} dw du}, \quad (5.16)$$

where $C_2 = \sqrt{R^2 - u^2}$ and $C_3 = \sqrt{R^2 - (u+v)^2} - w$. The details and proof for (5.16) is provided in Appendix D.

The proposed coverage probability metric helps in establishing a relationship between user mobility and $t_{up}(k)$. $t_{up}(k)$ must be selected in such a manner that it maximizes the coverage probability. Further, while optimizing $t_{up}(k)$, the coverage probability of all the users covered at k^{th} instant must be considered.

Since there are two factors impacting the choice of $t_{up}(k)$, we propose the minimization of a weighted single objective function as follows:

$$\min \left(\frac{\alpha \left(\sum_{l=1}^{k-1} F(l) + Q \mathbb{E} [F_{t_{up}(k)}] \right)}{T} \right) + \left(\frac{(1-\alpha)}{\sum_{n=1}^N A_j(k) P_{j,t_{up}(k)}} \right) \quad (5.17)$$

The first term corresponds to the fraction of T during which UAV-BS is in flight and cannot serve. Let $F_{t_{up}(k)}$ be the UAV-BS flight time during $t_{up}(k)$ with an average of $\mathbb{E} [F_{t_{up}(k)}]$. $Q = \left\lceil \frac{T - \sum_{j=0}^{k-1} t_{up}(j)}{t_{up}(k)} \right\rceil$ denotes the number of updates that may occur if update interval is fixed as $t_{up}(k)$ until completion of operation time T . The second term in (5.17) is the reciprocal of the sum of the coverage probability. A weight of α and $1 - \alpha$ has been assigned to the two terms, respectively. α can be tuned according to the network operator's requirement. For instance, if maximizing the coverage probability is the only requirement, the operator may set α as 0. However, if minimizing UAV-BS flight time is the only requirement, α may be set as 1.

In (5.17), the term $\sum_{l=1}^{k-1} F(l)$ will be same for all $t_{up}(k)$. Hence, the optimal $t_{up}(k)$ problem can be formulated as follows:

$$(P2) \quad \min \left(\frac{\alpha Q \mathbb{E} [F_{t_{up}(k)}]}{T} \right) + \left(\frac{(1-\alpha)}{\sum_{j=1}^U A_j(k) P_{j,t_{up}(k)}} \right) \quad (5.18)$$

$$\text{s.t. } t_{min} \leq t_{up}(k) \leq t_{max} \quad (5.19)$$

where t_{min} and t_{max} is the lower and upper limit decided for $t_{up}(k)$.

5.5 Problem Solution

In order to jointly optimize the UAV-BS placement and the update interval, we propose to solve (P1) and (P2) iteratively. First, at a given update instant, we solve (P1) exhaustively using Algorithm 7 to obtain the UAV-BS placement and $A_j(k) \forall j$. Then, $A_j(k)$ values are given as input to (P2). Using Algorithm 8, we obtain the update interval. The output of (P1) is then fed back to (P2). This goes on iteratively till the update interval value converges. In general, for the convergence of update interval, the condition $|t_{up}(k) - t_{out}(k)| < \tau$ must be met, where $t_{up}(k)$ and $t_{out}(k)$ are the input and output of Algorithm 7 and Algorithm 8, respectively. τ is the error tolerance value.

Algorithm 7: Solution to (P1) at k^{th} update instant

```

1 Input:  $t_{up}(k), u_x^*(k-1), u_y^*(k-1), fair_{thres}$ 
2  $\mathcal{X}$  - List of x coordinates of  $U$  users at  $k^{th}$  update instant
3  $\mathcal{Y}$  - List of y coordinates of  $U$  users at  $k^{th}$  update instant
4  $Hov_{loc}$  - List of 2-D UAV-BS Hovering locations
5  $Cov_{prev}$  - List of vectors containing coverage information of all users till  $(k-1)^{th}$  update instant
6 Output:  $\mathcal{A}$  - List of  $A_j(k) \forall j, u_x(k), u_y(k)$ 
7 Initialize:  $Cov_{mat} = \emptyset, Fly_{time} = \emptyset, Fairness = \emptyset$ 
8 for  $s = 1 : 1 : |Hov_{loc}|$  do
9    $Cov_{mat}(:,s) = Cov(\mathcal{X}, \mathcal{Y}, Hov_{loc})$ 
10   $Fly_{time}(s) = Fly((U_x^*(k-1), U_y^*(k-1), Hov_{loc}))$ 
11   $Fairness(s) = Fair(Cov_{mat}, Cov_{prev}),$ 
12 end for
13  $[u_x(k), u_y(k), Index1] = Find(Cov_{mat}, Fairness, Fly_{time}, fair_{thres}, t_{up}(k)),$ 
14  $Cov_{prev}(:,k) = Cov_{mat}(:, Index1),$ 
15  $\mathcal{A} = Cov_{mat}(:, Index1)$ 

```

In Algorithm 7, at $k = 1$, $t_{up}(1)$ is initialized as t_{min} . For $k > 1$, $t_{up}(k)$ is initialized as $t_{up}^*(k-1)$. Here, $t_{up}^*(k-1)$ is the optimal update interval at the $(k-1)^{th}$ update instant. $U_x^*(k-1)$ and $U_y^*(k-1)$ are the optimal coordinates at the $(k-1)^{th}$ update instant. Moreover, we consider that the 2-D hovering space of UAV-BS is discretized with a resolution of 20 m along the horizontal axis. Cov_{prev} is a list of $(k-1)$ vectors of length U . An element in a vector will be ‘1’ if user is covered otherwise ‘0’. For each of the hovering locations in Hov_{loc} , function $Cov()$ determines which of the U users may get covered. $Cov()$ assigns ‘1’ if a user is covered else ‘0’. The output of $Cov()$ is stored in Cov_{mat} . Then, the flight time of the UAV-BS is computed using function $Fly()$. Further, the fairness value corresponding to each hovering location is computed using $Fair()$ and stored in $Fairness$. Finally, function $Find()$ outputs $u_x(k), u_y(k)$ that maximize the number of users covered while satisfying the constraints (4) and (5). Also, we obtain the values of $A_n(k) \forall n$.

In Algorithm 8, $\mathbb{E}[F_{t_{up}(k)}]$ and $P_{j,t_{up}(k)}$ are utilized. The details on obtaining $\mathbb{E}[F_{t_{up}(k)}]$ are presented in Section 5.6. Moreover, $P_{j,t_{up}(k)}$ can be computed by substituting (5.15) and (5.16)

Algorithm 8: Solution to (P2) at k^{th} update instant

```
1 Input:  $\mathcal{A}$  - List of  $A_j(k) \forall j$ 
2  $t_{list}$  - List of update interval values,
3 Output:  $t_{up}(k)$ , Initialize:  $r = 1$ 
4 for  $t = t_{min} : 5 : t_{max}$  do
5    $Factor(r) = \frac{\alpha Q \mathbb{E}[F_i]}{T} + \frac{(1-\alpha)}{\sum_{j=1}^U A_j(k) P_{j,t}}$ 
6    $r=r+1$ 
7 end for
8  $Index2 = \text{In}(Factor)$ 
9  $t_{out}(k) = t_{list}(Index2)$ 
```

Table 5.3: Simulation Parameters

Parameter	Value	Parameter	Value
v_u	1.5 m/s	U	20
v_{uav}	25 m/s	$fair_{thres}$	0.7
σ	4 m	t_{min}	5 s
H	50 m	t_{max}	150 s
T	900 s	Simulated Area	400×400 sq. m

in (5.12). We discretize the update interval values with a resolution of 5 seconds. For each of the update interval values, the value of the function in (5.18) is computed and stored in $Factor$. $\text{In}()$ finds the index of the minimum value in $Factor$. Finally, $t_{out}(k)$ is obtained. In our work, we consider $\tau = 0$. Further, the update interval converges to $t_{up}^*(k)$ within 2 iterations⁷.

5.6 Results and Discussion

In the following, we present the results for the iterative solution to problems (P1) and (P2). For our study, we consider that all the users are initially inside the UAV-BS's coverage area. The initial distribution of the users and UAV-BS placement for a specific UAV-BS operation period of T seconds is given in Fig. 5.14. Further, we assume that the first UAV-BS placement update occurs after t_{min} seconds. On solving (P1) and (P2) iteratively, we obtain the optimal UAV-BS placement and $t_{up}(1)$. After this, the UAV-BS placement is updated at $(t_{min} + t_{up}(1))$ seconds. This process goes on till T seconds. In our study, we set the maximum allowable path-loss of 95 dB in a dense urban environment as a QoS metric. Based on the above, using the AtG path-loss model and $H = 50$ m, we obtain $R = 100$ m. The other simulation parameters are mentioned in Table 5.3. Further, the results are averaged over 100 UAV-BS operation periods.

⁷The number of iterations depend on the choice of τ as well as the resolution of update interval. With an increase in τ number of iterations will decrease. However, with an increase in resolution the number of iterations will also increase.

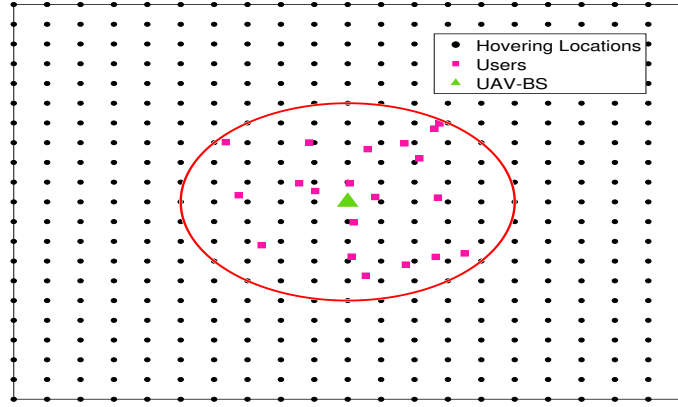


Figure 5.14: Initial Deployment

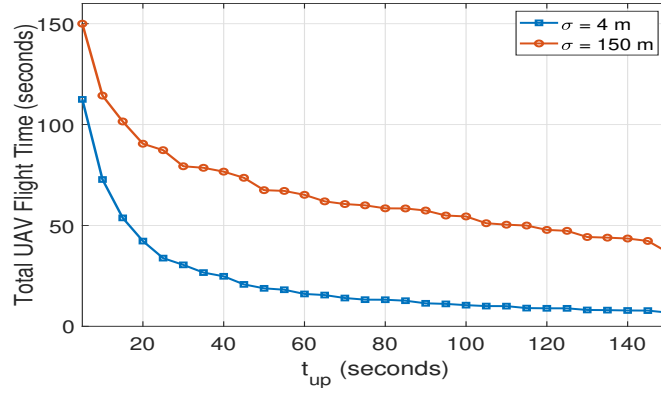


Figure 5.15: Total UAV-BS Flight time

As mentioned before, for any update interval t_{up} , the average UAV-BS flight time during t_{up} will be $\mathbb{E}[F_{t_{up}}]$ and can be written as follows:

$$\mathbb{E}[F_{t_{up}}] = \frac{\mathbb{E}\left[\sum_{i=1}^{T/t_{up}-1} F(i)\right]}{(T/t_{up}) - 1} \quad (5.20)$$

where $\mathbb{E}\left[\sum_{l=1}^{T/t_{up}-1} F(l)\right]$ is the average total UAV-BS flight time if updates are done after every t_{up} seconds. First, the term $\sum_{l=1}^{T/t_{up}-1} F(l)$ can be determined in an offline manner by solving (P1) for $\left(\frac{T}{t_{up}} - 1\right)$ update instants. Consequently, $\mathbb{E}\left[\sum_{l=1}^{T/t_{up}-1} F(l)\right]$ is the average of the above summation. A typical plot of average total UAV-BS flight time is shown in Fig. 5.15. It can be observed that as t_{up} increases the total UAV-BS flight time decreases. This is because lower update interval corresponds to frequent UAV-BS placement updates or vice versa. Consequently, the UAV-BS flies more frequently resulting in higher total UAV-BS flight time.

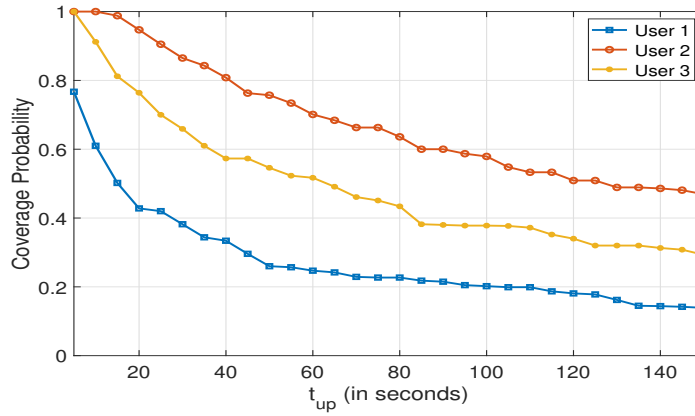


Figure 5.16: Coverage probability of users

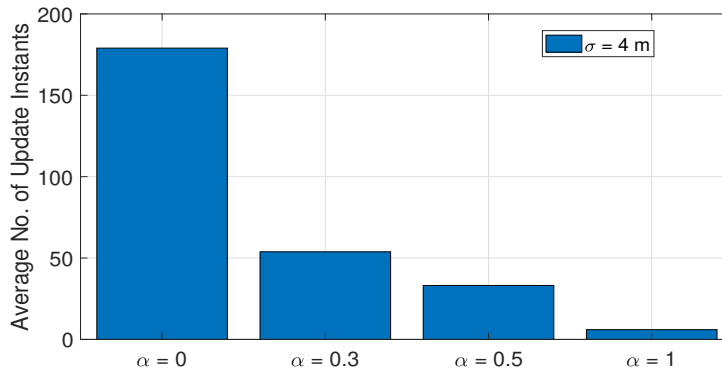


Figure 5.17: Average Number of update instants during $T = 900$ seconds

Fig. 5.16 shows the coverage probability of users 1, 2 and 3 at an initial position of $[56.8, -38.1]$ m, $[-4.1, -40.1]$ m and $[-0.3, 60.2]$ m respectively when the UAV-BS is located at $[0, 0]$ m. At lower update interval, UAV-BS updates its placement frequently that results in fewer occurrences of a user moving outside the UAV-BS coverage. However, at higher update interval, the probability of the user moving outside the UAV-BS coverage area increases. Consequently, for each user, coverage probability decreases with increase in t_{up} .

Figs. 5.17 and 5.18 show the average number of update instants and average update interval at $\sigma = 4$ m. With an increase in α the number of update instants reduces and update intervals increases. As evident from (13), it is because an increase in α results in a higher weight to reducing the total UAV-BS flight time. This means there would be less frequent UAV-BS placement updates. Further, Table 5.4 presents the standard deviation of update interval and number of update instants at $\alpha = 0, 0.3, 0.5$ and 1 .

Fig. 5.19 plots the average number of users covered in the network for $T = 900$ s. As α increases, the number of users covered is decreasing. This is due to the fact that less frequent

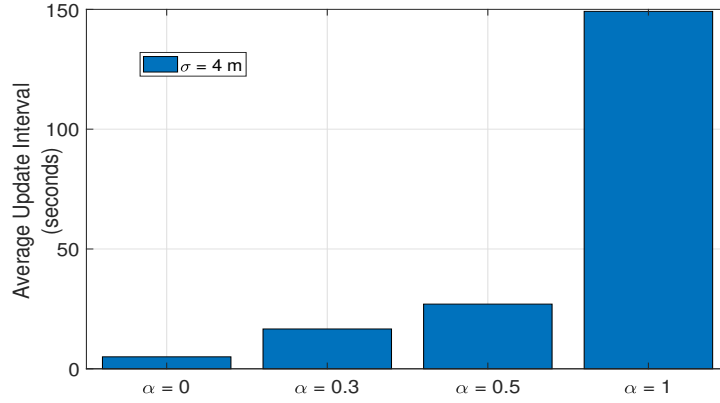


Figure 5.18: Average update interval at $T = 900$ seconds

Table 5.4: Standard Deviation

Parameter \ Weight, α	$\alpha = 0$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 1$
Std. Deviation of Update Instants	0 s	15.05 s	25.85 s	1.86 s
Std. Deviation of No. of update instants	0	9.35	5.76	0

UAV-BS placement updates result in lower coverage probability. Further, it can be observed that, on average, with one UAV-BS it is possible to provide coverage to around 14 users with the desired QoS. Further, the service time of UAV-BS can be written as:

$$\mathcal{S} = T - \mathbb{E}_p \left[\sum_{l=1}^{K_p} F(l) \right] \quad (5.21)$$

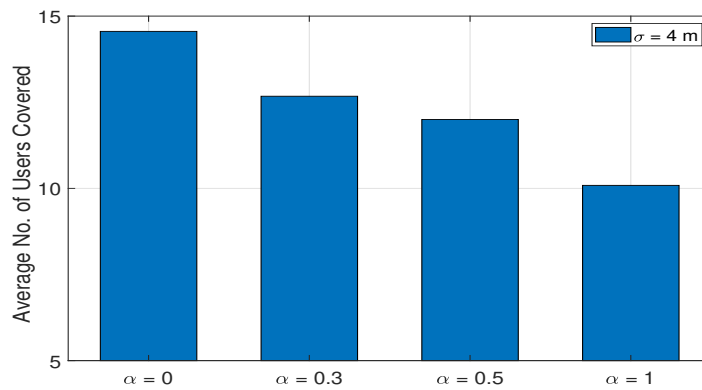


Figure 5.19: Average number of users covered at $T = 900$ seconds at $U = 20$

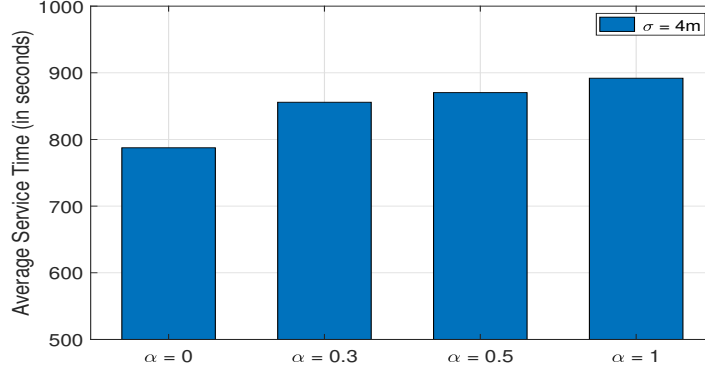


Figure 5.20: Average service time at $T = 900$ seconds

It can be observed from Fig. 5.20 that with an increase in α average service time increases. Hence, α can be tuned by the operator depending on whether more number of users should be covered or larger service time is desired.

The average number of users covered are determined by $A_j(k)$ as well as coverage probability and can be written as:

$$\bar{U} = \mathbb{E}_p \left[\frac{1}{K_p} \sum_{l=1}^{K_p} \sum_{j=1}^U A_j(k) P_{j,t_{up}(k)} \right] \quad (5.22)$$

where K_p denotes the number of update instants in p^{th} UAV-BS operation period.

5.7 Conclusion

We proposed a ground user mobility aware multi-UAV placement strategy for a disaster-resilient communication network. The disaster-affected area is divided into four zones and EFRs/ground users are assigned different zones. For developing a ground user mobility-aware strategy, we modeled the mobility of ground users within their assigned zones. Based on the above, an optimization problem is formulated to maximize the number of covered users while taking into account the UAV flight time constraint. We observed that there exists a trade-off between the average number of covered users and average coverage time. Further, UAV flight time constraint is more crucial in a multi-UAV scenario as compared to a single UAV scenario. This is due to the fact that multi-UAV scenario tends to have a lower average coverage time as compared to single UAV scenario with increase in UAV flight time permissible limit. It has also been observed that with an increase in UAV velocity, for a given average number of covered users, average coverage time increases. However, the overall energy consumption at a

UAV increases with an increase in UAV velocity because the increase in energy consumption due to hovering surpasses the decrease in energy consumption due to flight.

The above work in the multi-UAV network did not consider user coverage fairness. Further, it assumed a synchronous user mobility behavior and fixed update interval. Subsequently, we optimized the UAV-BS placement and update interval. Specifically, the UAV-BS placement is optimized to maximize the number of users covered at an update instant while accounting for the user fairness and UAV-BS flight time. Further, we introduced a weighted objective function depending on user coverage probability and total UAV-BS flight time for update interval optimization. Consequently, we proposed an iterative approach to solve the above two problems jointly. We evaluated the network service time and number of users covered for different weights in the optimal update interval problem formulation. Additionally, we derived the analytical expressions for the user coverage probability.

UAV-BS Placement Update with Resource Allocation

The previous chapter investigated the relationship between user mobility and UAV-BS placement update interval [110]. It considered a UAV-assisted communication network to serve the mobile ground users and optimize UAV-BS placement and update interval. However, the analysis was limited to a single UAV-BS network. Moreover, the UAV-BS placement was optimized to maximize the number of users covered at an update instant while accounting for the user fairness as well as UAV-BS flight time. Recently, research in UAV-BS placement optimization and resource allocation in UAV networks has also gained a lot of momentum [23–27]. This chapter broadly focuses on UAV-BS placement and resource allocation in the presence of mobile ground users.

Unlike [110], in this work, we consider a multi-UAV network where all the users are to be served during the UAV operation period. We aim to optimize the UAV-BS placement, resource allocation¹, user association, and update interval. We propose to divide the above optimization into two phases: phase 1 and phase 2. In phase 1, we jointly optimize UAV-BS placement, bandwidth allocation, and user association, whereas the update interval is optimized in phase 2. We propose the use of two metrics, i.e., total UAV-BS flight time and user coverage probability to optimize the update interval. Further, two different frameworks, namely, max sum rate and max min rate, are utilized. Specifically, in phase 1 of the max sum rate framework, the objective is to maximize the sum rate of the users. This framework is suitable for content caching applications where a single caching user may require a high bandwidth connection while a certain QoS is guaranteed at the rest of the users [111]. While in phase 1 of the max min rate framework, the minimum rate among the users is maximized. A sequential approach is proposed to solve phase 1 and phase 2. Further, the analytical expression for user coverage

¹In our work, resource denotes the bandwidth.

probability in terms of user mobility parameter and update interval are derived. The major contributions are summarized as follows:

Contributions

- We jointly optimize the UAV-BS placement, bandwidth allocation, user association, and update interval in a multi-UAV network. This optimization is divided into two phases, i.e., phase 1 and phase 2. In phase 1, the UAV-BS placement, bandwidth allocation, and user association are jointly optimized, whereas the update interval is optimized in phase 2. Two different frameworks, namely, max sum rate and max min rate, are utilized for the joint optimization.
- Specifically, in phase 1 of the max sum rate framework, the objective is to maximize the sum rate of the users, whereas, in phase 1 of the max min rate framework, the minimum rate among the users is maximized. The optimization problems in phase 1, for both the frameworks, are mixed-integer non-concave problems. Hence, the block coordinate descent algorithm is utilized. For phase 2, we propose a weighted objective function that minimizes the total UAV-BS flight time and maximizes the user coverage probability. A sequential approach is proposed to solve phase 1 and phase 2 jointly. We prove that convergence is guaranteed for the proposed sequential approach.
- The analytical expression for the coverage probability that accounts for the user mobility and update interval are also derived. It is shown that the analytical and simulated coverage probability are in agreement.
- We analyze the average update interval with respect to the number of UAV-BSs. Our work is also compared with a benchmark approach wherein the update interval is not optimized. It has been shown that, unlike the proposed work, the benchmark approach cannot adapt to the desired priority of service time and coverage probability.

It may be noted that simultaneous optimization of user association, bandwidth allocation, UAV-BS placement and update interval is not feasible. This is because the user association, bandwidth allocation and UAV-BS placement depends on the flight time of UAV-BS between two consecutive update instants, whereas the optimal update interval depends on the total UAV-BS flight time during the operation period. Consequently, in our work, the above optimization is divided into phase 1 and phase 2 and solved using the proposed sequential approach. This will be discussed in detail in Section 6.2.

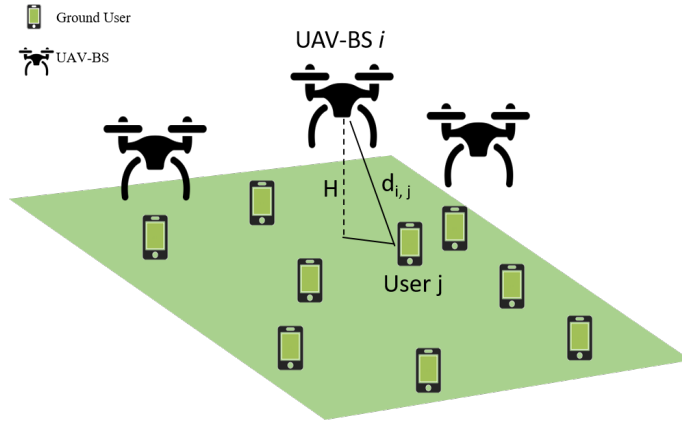


Figure 6.1: Network Model

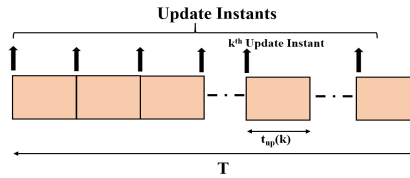


Figure 6.2: Network Optimization Timeline

6.1 Network Model

We consider a network model with U_B UAV-BSs and U mobile ground users, as shown in Fig. 6.1. The UAV-BS index is denoted as $i \in \{1, 2, \dots, U_B\}$ whereas the user index is denoted by $j \in \{1, 2, \dots, U\}$. The UAV-BS operates for T seconds to serve the ground users. The users are moving around following a random walk mobility model [112]. The distance traveled by a user in each transition of random walk is assumed to be Rayleigh distributed with shape parameter σ [113]. For analytical tractability, it is assumed that the velocity of the users is constant and each user has the same velocity, v_u . Further, in line with previous work, it is assumed that the altitude of each UAV-BS is fixed, and it is denoted by H [70]. Hence, we focus only on the 2-D UAV-BS placement in the horizontal plane. The total available bandwidth is assumed to be fixed (denoted as B_w), and each user is allocated a non-interfering orthogonal frequency band for transmission.

The mobility of the users necessitates recurring UAV network optimization, i.e., repeated optimization of user and UAV-BS association, user bandwidth allocation, and UAV-BS placement. Fig. 6.2 illustrates the network update timeline for T seconds. $t_{up}(k)$ is the time interval between the k^{th} and $(k+1)^{th}$ update instants. User mobility during the update interval will also impact the QoS at each user. Hence, it is important to optimize the update interval based on user mobility.

Let the 2-D placement of i^{th} UAV-BS at the k^{th} update instant be denoted as $U_i(k) = [u_{i,x}(k), u_{i,y}(k)]$ where $u_{i,x}(k)$ and $u_{i,y}(k)$ are the coordinates of the i^{th} UAV-BS in the horizontal plane. The 2-D location of user j at update instant k is denoted as $W_j(k) = [w_{j,x}(k), w_{j,y}(k)]$ where $w_{j,x}(k)$ and $w_{j,y}(k)$ are the coordinates of user j . The flight time of UAV-BS i between k^{th} and $(k+1)^{th}$ update instants is given as:

$$F_i(k) = \frac{\|U_i(k) - U_i(k-1)\|}{v_{uav}}, \quad (6.1)$$

where v_{uav} is the UAV-BS velocity. The path loss between UAV-BS i and user j at the k^{th} update instant assuming a dominant line-of-sight (LoS) component can be written as [67]:

$$L_{i,j}(k) = K_0 d_{i,j}^2(k) \delta_{LoS}, \quad (6.2)$$

where $K_0 = \left(\frac{4\pi f_c}{c}\right)^2$, c is the speed of light and f_c is the carrier frequency. δ_{LoS} is the attenuation in LoS path, $d_{i,j}(k)$ is the euclidean distance between UAV-BS i and user j at the k^{th} update instant and can be given as:

$$d_{i,j}(k) = \sqrt{H^2 + \|U_i(k) - W_j(k)\|^2}. \quad (6.3)$$

The SNR at user j when served by UAV-BS i will be

$$\gamma_{i,j}(k) = \frac{P_t}{\sigma_o^2 L_{i,j}(k)}, \quad (6.4)$$

where P_t is the transmit power of UAV-BS and σ_o^2 is the noise power at receiver.

6.2 Problem Formulation and Proposed Solution

As mentioned before, we plan to optimize the user association with UAV-BSs, bandwidth allocation, UAV-BS placement and update interval. The above optimization is divided into two phases: phase 1 and phase 2 which are discussed in detail below.

6.2.1 Phase 1

In phase 1, user association with UAV-BSs, user bandwidth allocation and UAV-BS placement are optimized at the k^{th} update instant. Let $Y_{i,j}(k)$ be the association indicator variable. $Y_{i,j}(k)$ is set as '1' when user j is associated to UAV-BS i or '0' otherwise at k^{th} update instant. $B_{i,j}(k)$

is the bandwidth allocated to the user j when associated to UAV-BS i at k^{th} update instant. The achievable rate between user j and UAV-BS i at k^{th} update instant can be written as:

$$R_{i,j}(k) = Y_{i,j}(k)B_{i,j}(k) \log_2(1 + \gamma_{i,j}(k)). \quad (6.5)$$

Let $\mathcal{Y} = \{Y_{i,j}(k), \forall i, j\}$, $\mathcal{B} = \{B_{i,j}(k), \forall i, j\}$ and $\mathcal{L} = \{U_i(k), \forall i\}$. Now, let us discuss the two optimization frameworks for phase 1:

6.2.1.1 Max Sum Rate Framework

In the max sum rate framework, the user association, bandwidth allocation and UAV-BS placement is optimized in order to maximize the sum rate across all the users.

$$(P3.1) \quad \max_{\mathcal{Y}, \mathcal{B}, \mathcal{L}} \sum_{i=1}^{U_B} \sum_{j=1}^U R_{i,j}(k) \quad (6.6)$$

$$\text{s.t.} \quad \sum_{i=1}^{U_B} \sum_{j=1}^U Y_{i,j}(k)B_{i,j}(k) \leq B_w, \quad (6.7)$$

$$\sum_{i=1}^{U_B} Y_{i,j}(k)B_{i,j}(k) \geq B_{thres}, \quad \forall j = 1, \dots, U, \quad (6.8)$$

$$\sum_{i=1}^{U_B} Y_{i,j}(k) = 1, \quad \forall j = 1, \dots, U, \quad (6.9)$$

$$F_i(k) < t_{up}(k), \quad \forall i = 1, \dots, U_B, \quad (6.10)$$

$$Y_{i,j}(k) \in \{0, 1\}, \quad \forall i, j. \quad (6.11)$$

Constraint (6.7) limits the total bandwidth availability as B_w . Constraint (6.8) ensures that the minimum bandwidth requirement of B_{thres} is met for all the users. Constraint (6.9) means that a user can only be associated with one UAV-BS at a time. Constraint (6.10) ensures that, at k^{th} update instant, $F_i(k)$ must be upper-bounded by the update interval. Further, constraint (6.11) represents that $Y_{i,j}(k)$ is a binary variable. Since the objective is a non-convex function of UAV-BS placement, the problem (P3.1) is a mixed-integer non-concave optimization problem. Hence, (P3.1) is a non-trivial problem to solve. Consequently, we propose to solve this problem using the block coordinate descent (BCD) algorithm. According to BCD, we split the problem into three sub-problems. These sub-problems are solved alternately in each iteration [67].

Sub-problem 1 Keeping UAV-BS placement and bandwidth allocation fixed, the user association $Y_{i,j}(k)$ is optimized. Hence, we solve the below optimization problem:

$$\begin{aligned} \max_{\mathcal{Y}} \quad & \sum_{i=1}^{U_B} \sum_{j=1}^U R_{i,j}(k) \\ \text{s.t.} \quad & (6.7) - (6.9), (6.11). \end{aligned} \quad (6.12)$$

The above problem is an integer linear programming (ILP) problem that can be solved using the IBM CPLEX solver.

Sub-problem 2 Keeping UAV-BS placement and user association fixed, the bandwidth allocation $B_{i,j}(k)$ is optimized. Hence, we solve the below optimization problem:

$$\begin{aligned} \max_{\mathcal{B}} \quad & \sum_{i=1}^{U_B} \sum_{j=1}^U R_{i,j}(k) \\ \text{s.t.} \quad & (6.7), (6.8). \end{aligned} \quad (6.13)$$

The above problem is a standard linear programming (LP) problem that can be solved using the IBM CPLEX solver.

Sub-problem 3 Keeping bandwidth allocation and user association fixed, UAV-BS placement is optimized. However, the objective function for sub-problem 3 is still a non-convex function of UAV-BS placement. To convexify the objective function, successive convex optimization needs to be carried out in each iteration of Algorithm 9. In successive convex optimization, a function is approximated by a more tractable function, in our case by using first-order Taylor approximation, at a given local point. The convexified problem is given as follows:

$$\begin{aligned} \max_{\mathcal{L}} \quad & \sum_{i=1}^{U_B} \sum_{j=1}^U \left[Y_{i,j}(k) B_{i,j}(k) \left(\log_2 \left(1 + \frac{A}{B + z_o} \right) - \frac{(\log_2 e) A (z - z_o)}{(B + z_o)(B + z_o + A)} \right) \right] \\ \text{s.t.} \quad & (6.10), \end{aligned} \quad (6.14)$$

where $A = \frac{P_i}{\delta_{Los} K_o \sigma_o^2}$, $B = H^2$, $z = \|U_i(k) - W_j(k)\|^2$ and $z_o = \|U_i(k, m-1) - W_j(k)\|^2$ at the m^{th} iteration² of Algorithm 9. It is a quadratic programming problem that can be solved using IBM CPLEX solver.

Algorithm 9 presents the BCD algorithm for solving sub-problems 1, 2, and 3. Here, m denotes the number of iterations. In each iteration, we first optimized the user association

²Refer to Appendix E for more details.

keeping bandwidth allocation and UAV-BS placement fixed. Accordingly, the user association is updated. Then, bandwidth allocation is optimized, keeping user association and UAV-BS placement fixed. This is followed by updating the bandwidth allocation. After this, UAV-BS is optimized, keeping user association and bandwidth allocation fixed. Finally, the UAV-BS placement is updated. The algorithm will converge when the fractional increment in the objective value, in this case, sum rate, is less than ε . The convergence of the BCD algorithm with successive convex optimization is guaranteed, as discussed in [66].

6.2.1.2 Max Min Rate Framework

In the max min rate framework, at update instant k , user association, bandwidth allocation and UAV-BS placement are optimized to maximize the minimum rate across the users.

$$\begin{aligned} \max_{\mathcal{Y}, \mathcal{B}, \mathcal{L}} \min_j \left(\sum_{i=1}^{U_B} R_{i,j}(k) \right) & \quad (6.15) \\ \text{s.t. (6.7), (6.9) - (6.11).} & \end{aligned}$$

The max min structure of the above problem can be reformulated as follows:

$$(P3.2) \quad \max_{\zeta, \mathcal{Y}, \mathcal{B}, \mathcal{L}} \zeta \quad (6.16)$$

$$\begin{aligned} \text{s.t. } \sum_{i=1}^{U_B} R_{i,j}(k) & \geq \zeta, \forall j & (6.17) \\ (6.7), (6.9) - (6.11). & \end{aligned}$$

where ζ is the lower bound on the rate of each of the users and ζ is to be maximized. In this framework, except for the minimum bandwidth constraint, all the constraints of max sum rate framework will be applicable. Since the max min rate framework is equivalent to ensuring a minimum rate at each of the user, the minimum bandwidth constraint becomes unnecessary. Due to the non-convex constraint (6.17), (P3.2) is a mixed-integer non-concave optimization problem. Similar to (P3.1), (P3.2) is non-trivial to solve. Hence, we solve this problem using the BCD algorithm. According to the BCD algorithm, we split the problem into three sub-problems.

Sub-problem 1 Keeping UAV-BS placements and bandwidth allocation fixed, the user association $Y_{i,j}(k)$ is optimized. Hence, we solve the below optimization problem:

$$\begin{aligned} \max_{\zeta, \mathcal{Y}} \zeta & \quad (6.18) \\ \text{s.t.} & \quad (6.7), (6.9), (6.11), (6.17). \end{aligned}$$

The above problem is an integer linear programming (ILP) problem that can be solved using the IBM CPLEX solver.

Sub-problem 2 Keeping UAV-BS placement and association fixed, the bandwidth allocation $Y_{i,j}(k)$ is optimized. Hence, we solve the below optimization problem:

$$\begin{aligned} \max_{\zeta, \mathcal{B}} \zeta & \quad (6.19) \\ \text{s.t.} & \quad (6.7), (6.17). \end{aligned}$$

The above problem is a standard linear programming (LP) problem that can be solved using the IBM CPLEX solver.

Sub-problem 3 Keeping bandwidth allocation and association fixed, UAV-BS placement is optimized. However, constraint (6.17) is still a non-convex function of UAV-BS placement. As mentioned above, successive convex optimization needs to be carried out in each iteration of the BCD algorithm. The convexified problem is given as follows:

$$\begin{aligned} \max_{\zeta, \mathcal{L}} \zeta & \quad (6.20) \\ \text{s.t.} & \quad (6.10), \end{aligned}$$

$$Y_{i,j}(k)B_{i,j}(k) \left(\log_2 \left(1 + \frac{A}{B+z_o} \right) - \frac{(\log_2 e)A(z-z_o)}{(B+z_o)(B+z_o+A)} \right) \geq \zeta, \forall j. \quad (6.21)$$

It is a quadratically constrained programming problem that can be solved using the IBM CPLEX solver.

Similar to max sum rate framework, Algorithm 9 will be used for solving the three sub-problems and its convergence is guaranteed. The objective value in this case will be the min rate.

Algorithm 9: Solution to phase 1 at k^{th} update instant

1 **Input:** $U_i(k-1) \forall i, W_j(k) \forall j, t_{up}(k), \varepsilon$
2 **Output:** $Y_{i,j}(k), B_{i,j}(k), U_i(k) \forall i, j$
3 **Initialize:** $B_{i,j}(k), U_i(k) \forall i, j$ and $m = 1$
4 **while** Fractional increment in objective value $> \varepsilon$ **do**
5 Solve Sub-problem 1 keeping $B_{i,j}(k), U_i(k)$ fixed $\forall i, j$
6 Update user association as $Y_{i,j}(k) = Y_{i,j}(k, m) \forall i, j$
7 Solve Sub-problem 2 keeping $Y_{i,j}(k), U_i(k)$ fixed $\forall i, j$
8 Update bandwidth allocation as $B_{i,j}(k) = B_{i,j}(k, m) \forall i, j$
9 Solve Sub-problem 3 keeping $Y_{i,j}(k), B_{i,j}(k)$ fixed $\forall i, j$
10 Update UAV-BS placement as $U_i(k) = U_i(k, m) \forall i$
11 $m=m+1$
12 **end while**

6.2.2 Phase 2

In phase 2, the update interval between k^{th} and $(k+1)^{th}$ update instants is optimized. According to the fly-hover-and-communicate protocol of UAV-BS operation, an increase in the flight time of UAV-BS decreases its service time [109]. Hence, minimization of total UAV-BS flight time must be considered. In our work, total UAV-BS flight time corresponds to the sum of the flight time during T , averaged over all the UAV-BSs. Further, the user mobility will impact the achievable rate at the user. Hence, we propose that the choice of update interval should depend on two factors: 1) total UAV-BS flight time and 2) user coverage probability. In the following, we will discuss in detail the coverage probability metric.

Coverage Probability: Coverage probability is defined as the probability that the rate achievable at a user is greater than the rate threshold R_{th} for all the user transitions within the update interval $t_{up}(k)$. It may be noted that during $t_{up}(k)$ user may have multiple transitions. The number of transitions depends on σ and v_u . The coverage probability expression for user j when associated to UAV-BS i for update interval $t_{up}(k)$ is given as follows:

$$P_{i,j,t_{up}(k)} = P[r_1 < \beta_j] P[r_2 < \beta_j | r_1 < \beta_j] \cdots P[r_n < \beta_j | r_{n-1} < \beta_j] \cdots P[r_{q_{t_{up}(k)}} < \beta_j | r_{q_{t_{up}(k)}-1} < \beta_j]. \quad (6.22)$$

In (6.22), r_n denotes the displacement of user j from UAV-BS i after its n^{th} transition and is given as:

$$r_n = \sqrt{(X_{in} + X_1 + \cdots + X_n - u_{i,x})^2 + (Y_{in} + Y_1 + \cdots + Y_n - u_{i,y})^2}. \quad (6.23)$$

X_{in} and Y_{in} are the coordinates of a user at k^{th} update instant. Further, $X_n, Y_n \sim \mathcal{N}(0, \sigma^2)$ are the change in the x and y coordinates of a user due to the n^{th} transition. β_j is the upper bound

on the displacement of user j and is given as:

$$\beta_j = \sqrt{\frac{P_t}{K_o(2^{R_{th}}/\Sigma_i B_{i,j}(k) - 1)N_o\delta_{LoS}} - H^2}. \quad (6.24)$$

Let $W_1 = X_1 + X_{in} - u_{i,x}(k)$ and $W_2 = Y_1 + Y_{in} - u_{i,y}(k)$. Hence, $W_1 \in \mathcal{N}(\mu_1, \sigma^2)$ and $W_2 \in \mathcal{N}(\mu_2, \sigma^2)$ where $\mu_1 = X_{in} - u_{i,x}(k)$ and $\mu_2 = Y_{in} - u_{i,y}(k)$. The first term on the R.H.S of (6.22) can be expressed as follows:

$$\begin{aligned} P[r_1 < \beta_j] &= P\left[\sqrt{W_1^2 + W_2^2} < \beta_j\right], \\ &= \int_{-\beta_j}^{\beta_j} \int_{-C_1}^{C_1} f_{W_2}(w_2) f_{W_1}(w_1) dw_2 dw_1, \\ &= \frac{1}{2\pi\sigma^2} \int_{-\beta_j}^{\beta_j} \int_{-C_1}^{C_1} e^{-\frac{(w_1-\mu_1)^2}{2\sigma^2}} e^{-\frac{(w_2-\mu_2)^2}{2\sigma^2}} dw_2 dw_1, \end{aligned} \quad (6.25)$$

where $C_1 = \sqrt{\beta_j^2 - w_1^2}$. The generalized expression for rest of the terms on the R.H.S of (6.22) will be:

$$P[r_n < \beta_j | r_{n-1} < \beta_j] = \frac{\int_{-\beta_j}^{\beta_j} \int_{-\beta_j-u}^{\beta_j-u} \int_{-C_2}^{C_2} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{v^2}{2\sigma^2}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} \Phi\left(\frac{C_3}{\sqrt{2}\sigma}\right) dw dv du}{\sigma\sqrt{2\pi} \int_{-\beta_j}^{\beta_j} \int_{-C_2}^{C_2} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} dw du}, \quad (6.26)$$

where $C_2 = \sqrt{\beta_j^2 - u^2}$, $C_3 = \sqrt{\beta_j^2 - (u+v)^2} - w$ and $\Phi(\cdot)$ is the error function [110], [114]. Now, we will discuss phase 2 optimization problem corresponding to max sum rate and max min rate frameworks.

6.2.2.1 Max Sum Rate Framework

Since there are two factors impacting the choice of $t_{up}(k)$, we propose the minimization of a weighted single objective function as given below [110]:

$$\begin{aligned}
\text{(P4.1)} \quad \min \quad & \underbrace{\alpha \left(\frac{Q \mathbb{E}_{\mathcal{A}} [F_{t_{up}(k)}] + \sum_{l=1}^{k-1} F(l)}{T} \right)}_{\Delta_1} \\
& + (1 - \alpha) \underbrace{\left(\frac{1}{\sum_{i=1}^{U_B} Y_{i,j}(k) P_{i,j,t_{up}(k)}}} \right)}_{\Delta_2}
\end{aligned} \tag{6.27}$$

$$\text{s.t. } t_{min} \leq t_{up}(k) \leq t_{max}, \tag{6.28}$$

where t_{min} and t_{max} is the lower and upper limit for $t_{up}(k)$. The numerator of Δ_1 is total UAV-BS flight time at $t_{up}(k)$. Hence, Δ_1 corresponds to the fraction of T during which UAV-BS is in flight and cannot serve. Let $F_{i,t_{up}(k)}$ denote the flight time during $t_{up}(k)$ for UAV-BS i . $F_{t_{up}(k)} = \frac{\sum_{i=1}^K F_{i,t_{up}(k)}}{K}$ is the flight time during $t_{up}(k)$ averaged over all U_B UAV-BSs. As mentioned before, user mobility impacts the UAV-BS placement which in turn decides the flight time. Hence, $F_{t_{up}(k)}$ is also averaged over random user locations $\mathcal{A} = \{W_j(k), \forall j, k\}$. Further, $Q = \left\lceil \frac{T - \sum_{l=1}^{k-1} t_{up}(l)}{t_{up}(k)} \right\rceil$ denotes the number of updates that may occur if update interval $t_{up}(k)$ is utilized for the remainder of operation period T . Here, $l \leq (k-1)$ denotes the previous update instants. $F(l)$ is the flight time between l^{th} and $(l+1)^{th}$ update instants, averaged over the UAV-BSs. It may be noted that $F(l)$ will be same for each $t_{up}(k)$ value. Δ_2 is the reciprocal of the coverage probability of the user with maximum rate. The reason for considering the maximum rate will be discussed in detail in Section 6.3. A weight of α and $1 - \alpha$ has been assigned to Δ_1 and Δ_2 , respectively. α can be tuned according to the network operator's requirement.

6.2.2.2 Max Min Rate Framework

$$\begin{aligned}
\text{(P4.2)} \quad \min \quad & \alpha \Delta_1 \\
& + (1 - \alpha) \underbrace{\left(\frac{U}{\sum_{i=1}^{U_B} \sum_{j=1}^U Y_{i,j,k} P_{i,j,t_{up}(k)}}} \right)}_{\Delta_3}
\end{aligned} \tag{6.29}$$

$$\text{s.t. (6.28)}.$$

The phase 2 formulation for max min rate framework is similar to the max sum rate framework except that the second term, i.e., Δ_3 is the reciprocal of the coverage probability averaged over all the users.

Algorithm 10 presents the solution to phase 2. In this, we exhaustively search for the optimal update interval. Algorithm 10 utilizes Δ_1 , Δ_2 and Δ_3 . The details of obtaining Δ_1 are presented in Section 6.3. Moreover, Δ_2 and Δ_3 can be computed using (6.22). We discretize the update interval values with a resolution of 5 seconds. For each of the update interval values, for (P4.1), Δ_1 and Δ_2 are computed and stored in T_1 and T_2 respectively. In case of (P4.2), Δ_1 and Δ_3 are computed and stored in T_1 and T_2 respectively. Further, the range of T_1 and T_2 is normalized. Functions $\text{Min}()$ and $\text{Max}()$ are used to find the minimum and maximum value in a list, respectively. Function $\text{In}()$ finds the index of the minimum value in a list.

Algorithm 10: Solution to phase 2 at k^{th} update instant

```

1 Input:  $\alpha, Y_{i,j}(k), B_{i,j}(k), U_i(k) \forall i, j$ 
2  $t_{list}$  - List of update interval values,
3 Output:  $t_{up}(k)$ , Initialize:  $s = 1$ 
4 for  $t = t_{min} : 5 : t_{max}$  do
5    $T_1(s) = \Delta_1$ 
6    $T_2(s) = \Delta_2$  for (P2.1)
7   or  $T_2(s) = \Delta_3$  for (P2.2)
8    $s = s + 1$ 
9 end for
10  $T_1 = \frac{T_1 - \text{Min}(T_1)}{\text{Max}(T_1) - \text{Min}(T_1)}$ 
11  $T_2 = \frac{T_2 - \text{Min}(T_2)}{\text{Max}(T_2) - \text{Min}(T_2)}$ 
12  $Index = \text{In}(\alpha T_1 + (1 - \alpha) T_2)$ 
13  $t_{up}(k) = t_{list}(Index)$ 

```

6.2.3 Proposed Sequential Approach for Jointly Solving Phase 1 and Phase 2

In this subsection, a sequential approach to jointly solve phase 1 and phase 2 is proposed [115]. First, at a given update instant, we solve the phase 1 problem using Algorithm 9 to obtain user association, bandwidth allocation, and UAV-BS placement. At the r^{th} iteration of the sequential approach, UAV-BS placement is initialized as follows:

$$\begin{aligned}
U_i(k) &= U_i(k, r - 1), \text{ for } r > 1, \\
&= U_i(k - 1), \text{ otherwise,}
\end{aligned} \tag{6.30}$$

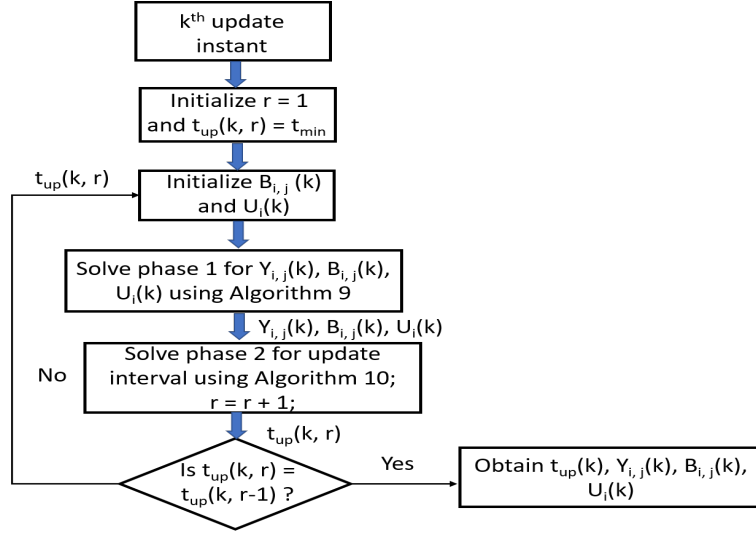


Figure 6.3: Flow diagram showing the Sequential Approach.

where $U_i(k, r-1)$ is the UAV-BS placement of UAV-BS i after $r-1$ iterations. Then, the output of phase 1 is given as input to phase 2. Using Algorithm 10, we obtain the update interval. The output of phase 2 is then fed back to phase 1. This goes on iteratively till the update interval value converges. In general, for the convergence of the sequential approach, the condition $|t_{up}(k, r) - t_{up}(k, r-1)| < \tau$ must be met. We have kept $\tau = 0$ to ensure that the flight time from phase 1 as well as total UAV-BS flight time and coverage probability in phase 2 are based on same update interval. Further, the update interval converges within 2 iterations³.

6.2.3.1 Convergence of Sequential Approach

In the sequential approach, at $r = 1$, we initialize update interval as t_{min} in phase 1. Further, Algorithm 10 will return $t_{up}(k, 1) \geq t_{min}$. Hence, in phase 1 and at $r = 2$, the flight time averaged over the UAV-BSs will be greater or equal to the flight time in iteration $r-1$, i.e.,

$$\frac{\sum_{i=1}^{U_B} F_i(k, r)}{U_B} \geq \frac{\sum_{i=1}^{U_B} F_i(k, r-1)}{U_B}. \quad (6.31)$$

This is because with increase in update interval, UAV-BSs may fly longer distances to maximize minimum rate in phase 1. Consequently,

³The number of iterations depend on the choice of τ as well as the resolution of update interval. With an increase in τ number of iterations will decrease. However, with an increase in resolution the number of iterations will also increase.

$$\begin{aligned}\Delta_2(k, r) &\leq \Delta_2(k, r-1), \text{ for (P2.1),} \\ \Delta_3(k, r) &\leq \Delta_3(k, r-1), \text{ for (P2.2),}\end{aligned}\tag{6.32}$$

where $\Delta_2(k, r)$ and $\Delta_3(k, r)$ is the reciprocal of coverage probability at update instant k after r^{th} iteration for max sum rate and max min rate frameworks, respectively. It may be noted that the first term in the objective function of phase 2 is independent of the instantaneous output of phase 1. Hence, the non-decreasing nature of coverage probability will ensure that update interval is also non-decreasing with each iteration, i.e.,

$$t_{up}(k, r+1) \geq t_{up}(k, r).\tag{6.33}$$

When the flight time and, in turn, coverage probability saturates with the change in update interval, convergence criteria will be met. Further, update interval is also upper bounded by t_{max} . Hence, the convergence of the sequential approach is guaranteed.

6.3 Results and Discussion

In the following, we present the results for the sequential approach to jointly solve phase 1 and phase 2. We consider that, initially, all the users and UAV-BSs are randomly distributed. Then, at $k = 1$, phase 1 and phase 2 are solved to optimize the initial placement of all the UAV-BSs, user association, bandwidth allocated to each user, and update interval, $t_{up}(k)$. After this, the network will be optimized at $t_{up}(k)$ seconds. This process goes on till T seconds. In our study, we consider $H = 100$ m, $f_c = 2$ GHz, $\sigma_o^2 = -100$ dBm, $v_u = 1.5$ m/s and $v_{uav} = 25$ m/s. Please note that $\alpha = 0$ and $\alpha = 1$ refers to update interval of t_{min} seconds and t_{max} seconds respectively. The other simulation parameters are mentioned in Table 6.2. Further, the results are averaged over 100 UAV-BS operation periods. First, we will present the results for standalone phase 1 followed by joint phase 1 and phase 2. We further compare our results with a benchmark approach. In the benchmark approach, the update interval is not optimized. Specifically, the network is optimized, or phase 1 occurs after a fixed update interval of $t_{max}/2$ seconds.

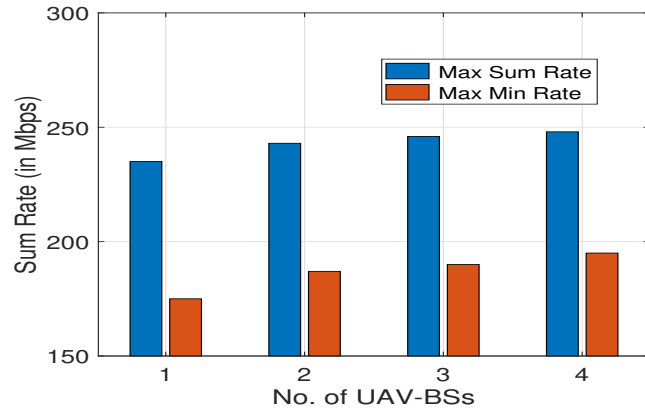


Figure 6.4: Sum Rate in Max Sum Rate Vs Max Min Rate Framework

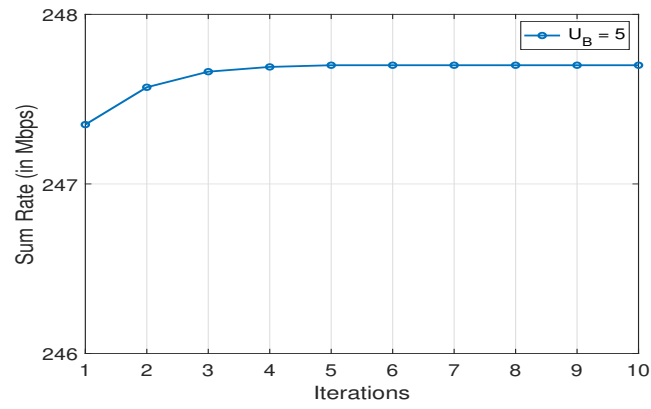


Figure 6.5: Convergence of sum rate Vs Algorithm 9 iterations

Table 6.1: Simulation Parameters

Parameter	Value
σ	8 m
ε	0.1
U	50
U_B	[1, 7]
T	900 s
P_t	0.15 W
B_w	20 MHz
B_{thres}	100 kHz
δ_{LoS}	3 dB [67]
Area	800 × 800 sq. m
t_{min}	5 s
t_{max}	100 s

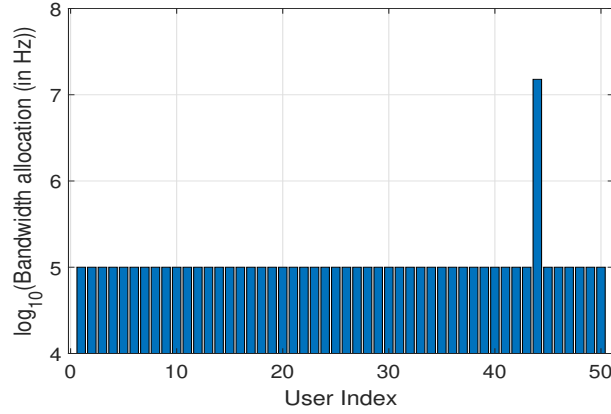


Figure 6.6: Bandwidth allocation per user in logarithmic scale for max sum rate framework

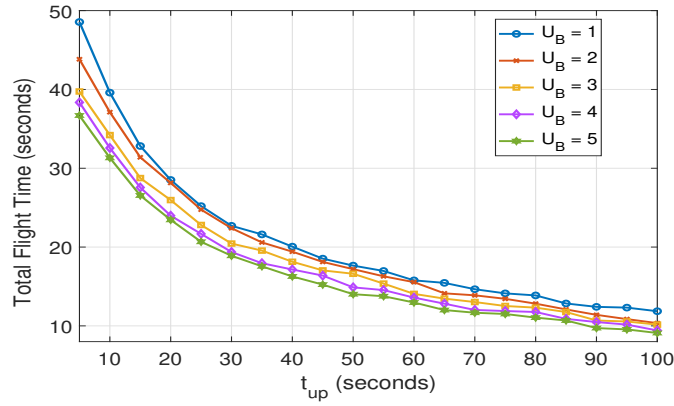


Figure 6.7: Total UAV-BS Flight Time for Max Sum Rate framework

6.3.1 Phase 1

6.3.1.1 Max Sum Rate Framework

Fig. 6.4 presents the sum rate of the users with respect to the number of UAV-BSs in the network for an update interval $t_{up} = 100$ s. It can be observed that as U_B increases sum rate also increases. We also observe in Fig. 6.4 that the sum rate in max min rate framework is lower than that of the max sum rate framework. Further, Fig. 6.5 depicts the convergence of sum rate with the increase in iterations of Algorithm 9. It can be seen that the convergence is obtained in three iterations. The optimal bandwidth allocation at one of the update instants is demonstrated in Fig. 6.6. The bandwidth threshold of $B_{thres} = 100$ kHz (in log scale, $\log_{10}(10^5) = 5$) is met at all the users. However, there is one user who is allocated the maximum bandwidth. Please note that this user may vary during the operation period.

As mentioned before, for update interval t_{up} , the flight time of UAV-BS averaged over user locations will be $\mathbb{E}_{\mathcal{A}} [F_{t_{up}}]$ and can be given as:

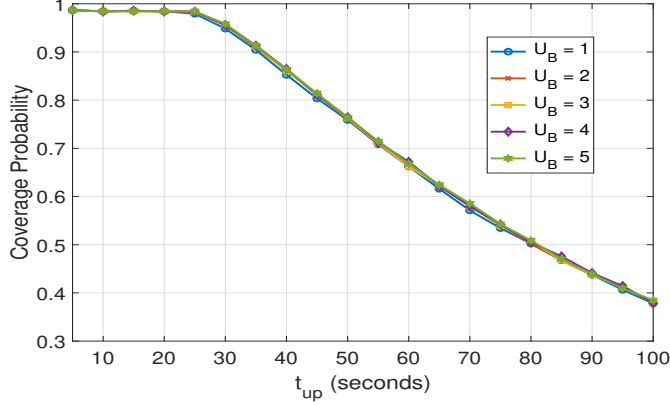


Figure 6.8: Coverage Probability for Max Sum Rate framework at $R_{th} = 190$ Mbps

$$\mathbb{E}_{\mathcal{A}} [F_{t_{up}}] = \mathbb{E}_{\mathcal{A}} \left[\frac{\sum_{i=1}^{U_B} \sum_{k=1}^{\lceil T/t_{up} \rceil} F_i(k)}{U_B (\lceil T/t_{up} \rceil)} \right], \quad (6.34)$$

assuming the updates are done after every t_{up} seconds. The term $\sum_{k=1}^{\lceil T/t_{up} \rceil} F_i(k)$ can be determined in an offline manner by solving (P1.1) for $\lceil \frac{T}{t_{up}} \rceil$ update instants. A typical plot of total UAV-BS flight time is shown in Fig. 6.7. It can be observed that as $t_{up} \in [t_{min}, t_{max}]$ increases the total UAV-BS flight time decreases. This is because lower update interval corresponds to frequent UAV-BS placement updates or vice versa. Consequently, the UAV-BSs fly more frequently resulting in higher total UAV-BS flight time. Further, with increase in U_B total UAV-BS flight time decreases. This is because SNR of the user with maximum bandwidth governs the maximum sum rate. Hence, with increase in U_B , UAV-BSs will have to fly less to cater to that maximum bandwidth user.

Coverage probability is defined as the probability that the rate achieved at the maximum bandwidth user is above the rate threshold for the complete update interval, t_{up} . Fig. 6.8 shows the plot for the coverage probability at each $t_{up} \in [t_{min}, t_{max}]$. Due to the fact that the network will always try to improve the SNR of the maximum bandwidth user, the achievable rate at each update instant will be the same for each U_B . Hence, it is interesting to observe that the coverage probability of the user with maximum bandwidth is unchanged due to increase in U_B . However, the coverage probability degrades with an increase in t_{up} .

6.3.1.2 Max Min Rate Framework

As shown in Fig. 6.9, it is quite intuitive that with an increase in U_B , the minimum user rate increases. Further, Fig. 6.10 exhibits that the minimum rate converges in two iterations of Algorithm 9. Fig. 6.11 shows that, unlike the max sum rate framework, the optimal bandwidth

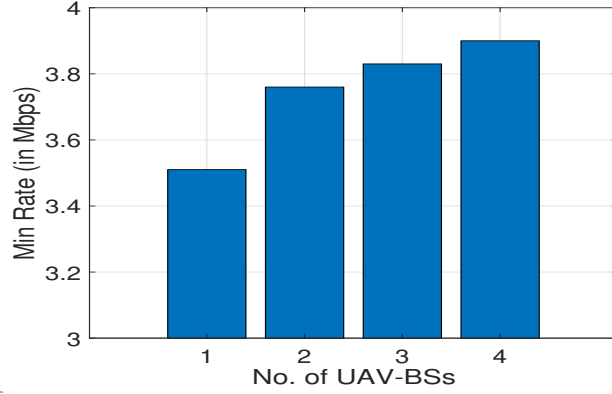


Figure 6.9: Minimum user rate in Max Min Rate framework

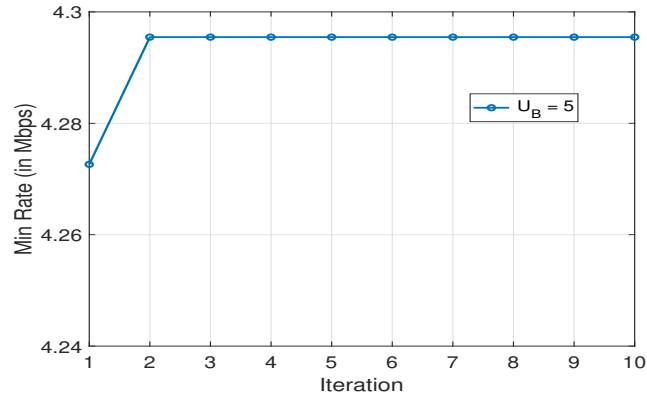


Figure 6.10: Minimum user rate Vs Algorithm 9 Iterations

allocation for the max min rate framework is well distributed across the users. This is because, unlike the max sum rate framework, in max min rate framework, there is a lower bound on the user rate.

Similar to the max sum rate framework, total UAV-BS flight time can be evaluated for the max min rate framework. Fig. 6.12 presents the total UAV-BS flight time with respect to update interval. Unlike the max sum rate framework, the total UAV-BS flight time initially increases with increase in U_B and then decreases (as observed for $U_B = 7$). This is because, in order to maximize the minimum rate, the UAV-BSs will tend to fly more; hence, increasing the total UAV-BS flight time. However, for a given area, when U_B increases sufficiently (in the present case $U_B = 7$), the need to fly around decreases.

Unlike the max sum rate framework, the coverage probability for max min rate framework is averaged over the coverage probability of all the users in the network. For a fair comparison of coverage probability with respect to U_B , we consider only those realizations where the min rate at $k = 1$ is greater than or equal to R_{th} . Fig. 6.13 presents the coverage probability with

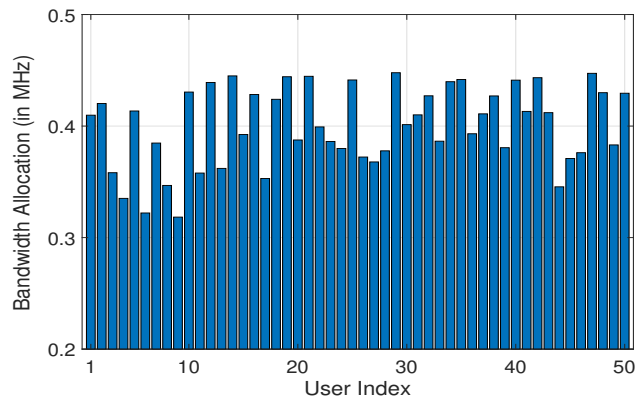


Figure 6.11: Bandwidth allocation in Max Min Rate framework

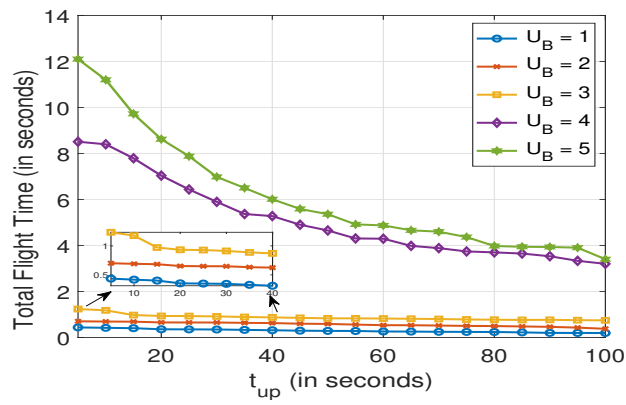


Figure 6.12: Total UAV-BS Flight time for Max Min Rate framework

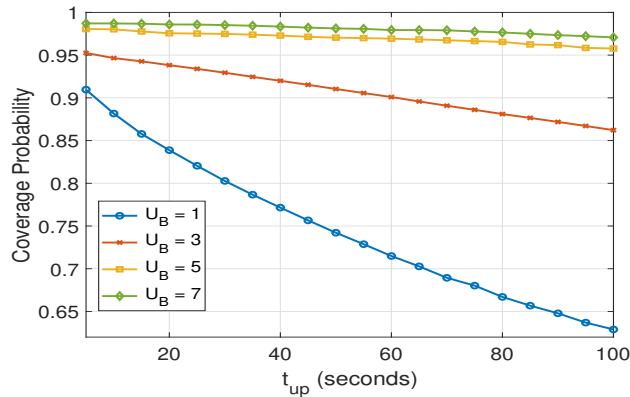


Figure 6.13: Coverage Probability for Max Min Rate framework at $R_{th} = 3.6$ Mbps

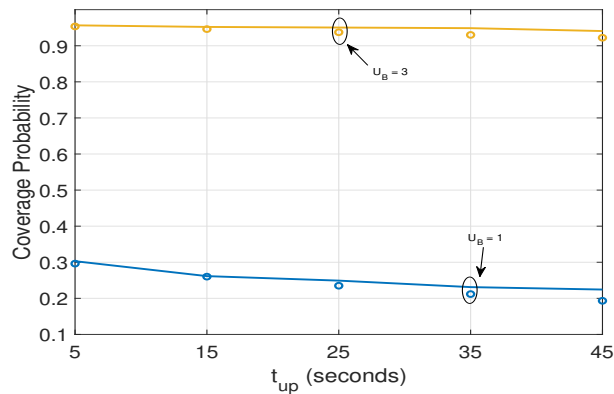


Figure 6.14: Comparison of Analytical and Simulated coverage probability at $R_{th} = 3.5$ Mbps

respect to update interval. As evident, the coverage probability decreases with increase in update interval. It is because, on an average, with increase in update interval, the displacement of a user from its associated UAV-BS increases. Consequently, the user rate decreases more at a higher update interval, and this decreases the coverage probability. Further, the coverage probability improves with increase in U_B . This is because the displacement of a user from its associated UAV-BS decreases with increase in U_B . We also compared our analytical coverage probability results with the simulated coverage probability. In Fig. 6.14, it can be observed that the analytical plot upper bounds the simulated plot. This is because for analytical tractability we considered only the average number of user transitions per update interval, as shown in (22).

6.3.2 Joint Solution of Phase 1 and Phase 2

To jointly solve phase 1 and phase 2, we employ the sequential approach proposed in Section 6.2. We have compared the results of the sequential approach to the benchmark approach.

Further, in scenarios where the coverage probability is zero for all users at each update interval, t_{min} is selected as the optimal update interval.

6.3.2.1 Max Sum Rate Framework

In the max sum rate framework, as shown in Fig. 6.15, the average update interval increases with increase in α . This is because, with increase in α , more priority is given to minimizing the total UAV-BS flight time. To minimize the total UAV-BS flight time, the average update interval must be longer so that UAV-BSs take less number of flights during T . Further, the average update interval is the same for all U_B . The first factor responsible for this behavior is that, as seen in Fig. 6.7, the rate of decrease of total UAV-BS flight time with increase in update interval is the same for all U_B . The other factor is that the coverage probability is the same for all U_B as shown in Fig. 6.8. Figs. 6.16 and 6.17 show that the benchmark approach prioritizes service time⁴ over coverage probability. Unlike the proposed optimization of update interval, the benchmark approach cannot adapt to the desired priority of service time or coverage probability.

6.3.2.2 Max Min Rate Framework

In Fig. 6.18, we observe an increase in average update interval with an increase in U_B . From Fig. 6.13, it is obvious that the gradient of coverage probability, with an increase in update interval, lowers at higher U_B . Hence, lesser number of updates will be required at higher U_B , i.e., longer update interval at higher U_B . Further, similar to max sum rate framework, the average update interval increases with increase in α . This is because, with increase in α , more priority is given to minimizing the total UAV-BS flight time. In Figs. 6.19 and 6.20, the sequential approach is compared with the benchmark approach. It can be observed that the benchmark approach prioritizes coverage probability over service time. Once again, the benchmark approach cannot adapt to the desired priority of service time or coverage probability. Please note that, in Fig. 6.19, the variation in service time will become more significant with increase in area.

6.4 Conclusion

We considered a multiple UAV-BS network where all the mobile ground users are served during the operation period. The joint optimization of user association, UAV-BS placement, bandwidth allocation, and update interval has been proposed. The optimization is divided into two phases: phase 1 and phase 2. Phase 1 optimized the user association, UAV-BS

⁴Service time is obtained by subtracting total UAV-BS flight time from T .

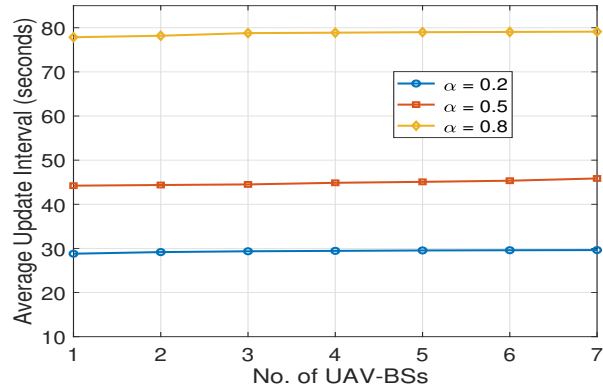


Figure 6.15: Average Update Interval Vs number of UAV-BSs

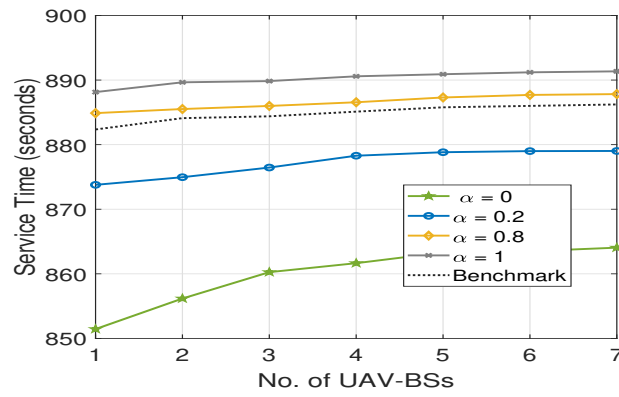


Figure 6.16: Comparison of service time with benchmark approach in Max sum rate framework

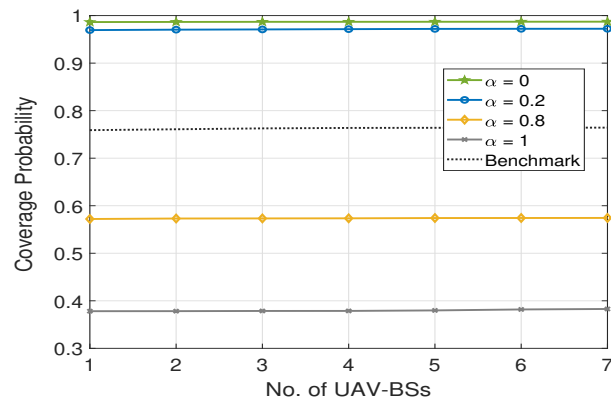


Figure 6.17: Comparison of coverage probability with benchmark approach in Max sum rate framework

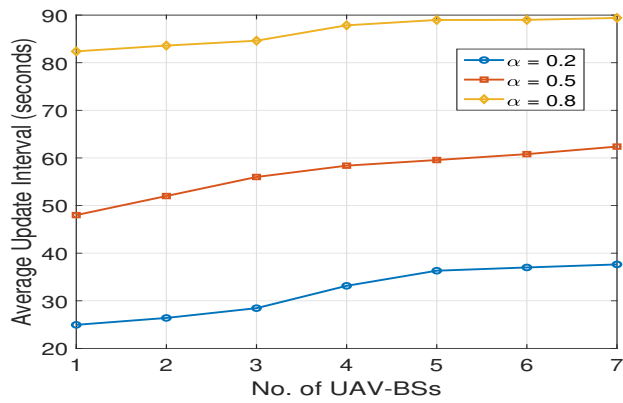


Figure 6.18: Average Update Interval Vs Number of UAV-BSs for $\alpha = \{0.2, 0.5, 0.8\}$

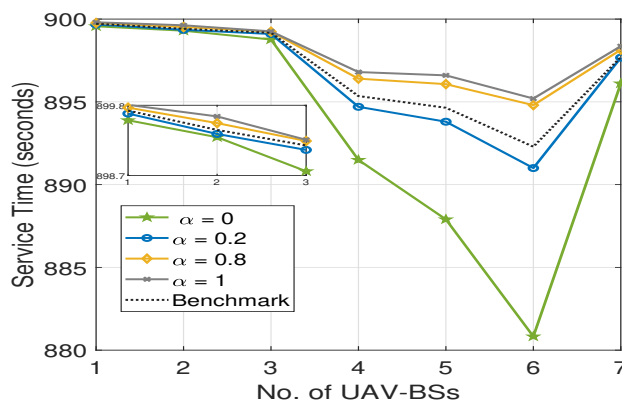


Figure 6.19: Comparison of service time with benchmark approach in max min rate framework

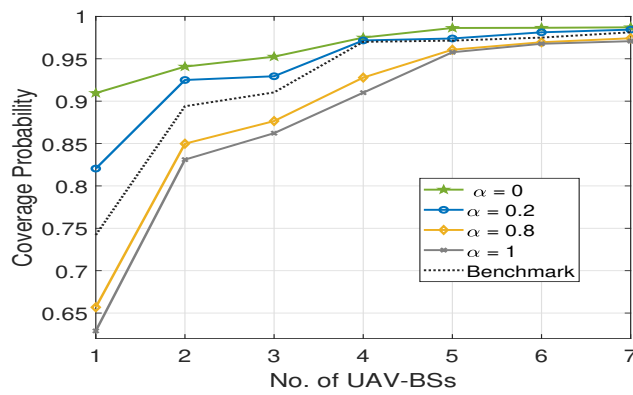


Figure 6.20: Comparison of coverage probability with benchmark approach in max min rate framework

placement, and bandwidth allocation, whereas update interval is optimized in phase 2. These phases are solved using a sequential approach. It has been shown that the convergence of the sequential approach is guaranteed. Further, max sum rate and max min rate frameworks have been utilized for the joint optimization. Phase 1 of the max sum rate framework focused on maximizing the sum rate of the users, whereas phase 1 of max min rate framework maximized the worst-off user rate. We showed that the total UAV-BS flight time decreases with an increase in UAV-BSs in the max sum rate framework. However, in the max min rate framework, the total UAV-BS flight time first increases and then decreases. In the max sum rate framework, the coverage probability does not change with the number of UAV-BSs. However, in the max min rate framework, coverage probability improves with an increase in UAV-BSs. The analytical expression for the coverage probability while accounting for the user mobility has also been derived. It has been shown that the analytical and simulated coverage probability are in agreement. In the max sum rate framework, the update interval does not change with increase in UAV-BSs whereas in the max min rate framework, the update interval increases with increase in UAV-BSs. Our work has been compared with a benchmark approach wherein the update interval is not optimized. It has been observed that, in the max sum rate framework, the benchmark approach prioritizes service time, whereas, in the max min rate framework, the benchmark approach prioritizes coverage probability. However, in both frameworks, the proposed work can adapt to the desired priority of service time and coverage probability.

Conclusion and Future Work

In this thesis, we focused on characterizing user mobility to determine and optimize the performance of D2D and aerial communication networks. We investigated D2D and aerial networks for different applications like content caching and emergency communication.

We began with an analysis of the D2D+cellular framework by making use of real-world user mobility traces [79]. The real-world data helped to exploit the spatio-temporal correlations present in user mobility. Subsequently, a realistic performance evaluation of two D2MD networks, namely D2MD-U and D2MD-M utilizing the joint spatio-temporal behavior of the users, is presented [116]. To obtain the joint spatio-temporal behavior of the users, two novel methods, OME and EME, are proposed. In OME, the joint spatio-temporal behavior is extracted from the past location information of the requesting users, whereas in EME, the location information of the requesting users is unknown, and the expected occupancies were estimated using a training dataset of users. The results demonstrated the impact of spatiotemporally correlated user mobility on the data offloading capability of D2MD networks.

The above observations motivated us to investigate real-world user mobility-aware optimization of D2D networks. Further, content caching coupled with D2D reduces the content delivery time and helps in offloading traffic from the cellular network. Hence, we explored user mobility aware cache selection in D2MD networks [111]. We proposed a greedy algorithm for cache selection to solve the combinatorial optimization problem of selecting a set of the minimum number of caches to achieve the desired user load for a D2MD network. The set of caches is selected by utilizing real-world location information to obtain the spatio-temporal behavior of the users. The proposed work has been shown to alleviate the caching load on the cellular network. Further, a discrete-time inhomogeneous Markov chain is presented to model the joint mobility pattern for the users in the D2MD network. The selected caches are tagged to their social group and are responsible for doing D2D multicast to disseminate the popular multimedia files to the non-caching users. Further, the presented optimization has been shown

to perform better than mobility-unaware cache selection. Moreover, the proposed framework is shown to further reduce the caching load by not discarding the previously selected caches.

Next, in order to establish emergency communication networks, we proposed a novel joint routing, scheduling, and pairing framework for a disaster-resilient communication network based on multi-hop D2D [117, 118]. An optimization problem to maximize the number of users covered in the dead spot within a certain deadline is formulated. The formulated problem is shown to be NP-hard. Hence, we proposed SCARP. It is demonstrated that the proposed algorithm outperforms the widely used SPR based scheduling in terms of users covered at stringent deadlines and energy consumption per active D2D relay at the less stringent deadline. This is because the D2D relays selected using SCARP are more spatially distributed, reducing wireless link contention at stringent deadlines and lowering energy consumption at less stringent deadlines. Further, the gain in users covered on using SCARP w.r.t. SPR based scheduling becomes more prominent with an increase in the density of D2D relays.

In addition to the above, we also proposed a ground user mobility aware multi-UAV placement strategy for a disaster-resilient communication network [109]. The disaster-affected area is divided into four zones and EFRs/ground users are assigned different zones. For developing a ground user mobility-aware strategy, we modeled the mobility of ground users within their assigned zones. Based on the above, an optimization problem is formulated to maximize the number of covered users while taking into account the UAV flight time constraint. We observed that there exists a trade-off between the average number of covered users and average coverage time. Further, UAV flight time constraint is more crucial in a multi-UAV scenario as compared to a single UAV scenario. This is due to the fact that multi-UAV scenario tend to have a lower average coverage time as compared to single UAV scenario with increase in UAV flight time permissible limit. It has also been observed that with an increase in UAV velocity, for a given average number of covered users, average coverage time increases. However, the overall energy consumption at a UAV increases with an increase in UAV velocity because the increase in energy consumption due to hovering surpasses the decrease in energy consumption due to flight.

The above work in the multi-UAV network did not consider user coverage fairness. Further, it assumed a synchronous user mobility behavior and fixed update interval. Subsequently, we optimized the UAV-BS placement and update interval [110]. Specifically, the UAV-BS placement is optimized to maximize the number of users covered at an update instant while accounting for the user fairness and UAV-BS flight time. Further, we introduced a weighted objective function depending on user coverage probability and total UAV-BS flight time for update interval optimization. Consequently, we proposed an iterative approach to solve the above two problems jointly. We evaluated the network service time and number of users

covered for different weights in the optimal update interval problem formulation. Additionally, we derived the analytical expressions for the user coverage probability.

The above work is further extended to a multi-UAV network while accounting for resource allocation. The joint optimization of user association, UAV-BS placement, bandwidth allocation, and update interval is proposed. The optimization is divided into two phases: phase 1 and phase 2. Phase 1 optimized the user association, UAV-BS placement, and bandwidth allocation, whereas update interval is optimized in phase 2. These phases are solved using a sequential approach. It has been shown that the convergence of the sequential approach is guaranteed. Further, max sum rate and max min rate frameworks have been utilized for joint optimization. Phase 1 of the max sum rate framework focused on maximizing the sum rate of the users, whereas phase 1 of max min rate framework maximized the worst-off user rate. We showed that the total UAV-BS flight time decreases with an increase in UAV-BSs in the max sum rate framework. However, in the max min rate framework, the total UAV-BS flight time first increases and then decreases. In the max sum rate framework, the coverage probability does not change with the number of UAV-BSs. However, in the max min rate framework, coverage probability improves with an increase in UAV-BSs. The analytical expression for the coverage probability while accounting for the user mobility has also been derived. It has been shown that the analytical and simulated coverage probability are in agreement. In the max sum rate framework, the update interval does not change with increase in UAV-BSs whereas in the max min rate framework, the update interval increases with increase in UAV-BSs. Our work has been compared with a benchmark approach wherein the update interval is not optimized. It has been observed that, in the max sum rate framework, the benchmark approach prioritizes service time, whereas, in the max min rate framework, the benchmark approach prioritizes coverage probability. However, in both frameworks, the proposed work can adapt to the desired priority of service time and coverage probability.

7.1 Future Work

This thesis will pave the way for some exciting future research work. The key research directions that can be explored are:

- In our work on joint optimization of UAV-BS placement, resource allocation, and update interval, we considered that the UAV-BSs are operating on orthogonal resource blocks. Further, we only optimized the 2-D UAV-BS placement. It will be interesting to investigate the 3-D UAV-BS placement, resource allocation, and update interval while considering inter UAV-BS interference in the presence of mobile ground users [119].

- In a UAV-assisted communication network, user handovers will happen across the UAV-BSs due to user mobility. However, the higher the number of handovers higher will be the communication overhead [120]. Minimizing the number of user handovers while guaranteeing desired user rate in the network is another future research direction. Further, we also need to study the impact of minimizing user handovers on the update interval.
- Recently, research on the intelligent reflecting surface (IRS) assisted wireless communication has gained a lot of traction. IRS consists of many reflecting units which are jointly adjusted to reconfigure the wireless signal transmitting environment. The existing works on IRS assisted communication networks mainly consider terrestrial IRS [121]. In the future, aerial IRS (mounted on UAV) assisted UAV networks may be explored. The joint optimization of aerial IRS placement and UAV-BS placement in the presence of mobile ground users will be a challenging problem to solve.

Publications

Journals

1. **M. Peer**, V. A. Bohara, A. Srivastava, G. Ghatak, "User Mobility-Aware UAV-BS Placement Update with Optimal Resource Allocation," submitted to *IEEE Transactions on Communications*.
2. D. N. Anwar, **M. Peer**, K. Lata, A. Srivastava, V. A. Bohara, "3-D Deployment of VLC Enabled UAV Networks with Energy and User Mobility Awareness," in *IEEE Transactions on Green Communications and Networking*, 2022.
3. **M. Peer**, V. A. Bohara and A. Srivastava, "Enabling Disaster-Resilient Communication Using Multi-Hop Device-to-Device Framework," in *Wireless Networks (Springer)*, 2020.
4. **M. Peer**, V. A. Bohara and A. Srivastava, "Cache Selection in Dynamic D2D Multicast Networks Using Inhomogeneous Markov Model," in *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3235-3245, 1 Oct.-Dec. 2020.
5. **M. Peer**, V. A. Bohara and A. Srivastava, "Real-World Spatio-Temporal Behavior Aware D2D Multicast Networks," in *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1675-1686, 1 July-Sept. 2020.

Conferences

1. S. Dohadwalla, **M. Peer**, V. A. Bohara, "Fair Multiple Subchannel Assignment and 3-D UAV-BS Placement in UAV-Enabled Networks," accepted in *IEEE VTC Spring*, 2022.
2. Y. Gupta, **M. Peer**, V. A. Bohara, "Performance Analysis of RF/VLC Enabled UAV Base Station in Heterogeneous Network," accepted in *IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021.

3. **M. Peer**, V. A. Bohara, A. Srivastava and G. Ghatak, "User Mobility-Aware Time Stamp for UAV-BS Placement," 2021 *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2021, pp. 1-6, doi: 10.1109/WCNCW49093.2021.9420032.
4. **M. Peer**, V. A. Bohara and A. Srivastava, "Multi-UAV Placement Strategy for Disaster-Resilient Communication Network," 2020 *IEEE 92nd Vehicular Technology Conference (VTC-Fall)*, 2020, pp. 1-7, doi: 10.1109/VTC2020-Fall49728.2020.9348687.
5. **M. Peer**, V. A. Bohara and A. Srivastava, "Multi-Hop D2D Framework for Disaster-Resilient Communication Network," 2020 *IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 584-589, doi: 10.1109/5GWF49715.2020.9221256.
6. **M. Peer**, V. A. Bohara and A. Srivastava, "On the performance of network-assisted indoor device-to-device communication using location awareness and realistic path loss models," 2017 *IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1-7, doi: 10.1109/PIMRC.2017.8292583.

PDF of Minimum of Channel Gains

The instantaneous rate expressions for D2MD as well as cellular multicast involve a minimum of channel gains which are exponentially distributed. Let the random variable Y denote the minimum of channel gains in (2.25) of the main document. The pdf of Y can be given as follows [122]:

$$f_y = \left(\sum_B x_B \Lambda^B \right) e^{-\sum_B x_B \Lambda^B y}, \quad (\text{A.1})$$

where Λ^B is the parameter of channel gain distribution in building B . The channels from users in building B to BS are identically distributed where as the channel gains corresponding to users present in different buildings are non-identically distributed. The value of U as well as the building occupancy alters the pdf of Y .

The pdf of the minimum of channel gains Z in γ_B expression in (2.20) of the main document can be given as:

$$f_z = (\Lambda_1^B + \dots + \Lambda_{W'}^B) e^{-(\Lambda_1^B + \dots + \Lambda_{W'}^B)z}, \quad (\text{A.2})$$

where Λ_i^B is the parameter of channel gain distribution of i^{th} user's link to C in the D2MD group. With increase in W' , the mean of Z gets reduced.

Computational Complexity

Exhaustive Search: When exhaustive algorithm is employed, the decision can be carried out only after evaluating the cellular load at each candidate set of caches with cardinality $|C_p| = 1$ to $|C_p| = \lfloor K/2 \rfloor$. Hence, the number of candidate sets will be

$$\begin{aligned} \mathcal{S} &= \binom{K}{1} + \binom{K}{2} + \dots + \binom{K}{\lfloor K/2 \rfloor}, \\ &\approx \mathcal{O}(2^{(K-1)}). \end{aligned}$$

This implies the complexity of exhaustive search will be $\approx \mathcal{O}(2^{(K-1)})$.

Greedy Algorithm: For the proposed algorithm the number of candidate sets of caches will be given as:

$$\begin{aligned} \mathcal{S} &= K + (K-1) + \dots + (K - \lfloor K/2 \rfloor), \\ &= \frac{K(K+1)}{2} - \frac{(K - \lfloor K/2 \rfloor)(K - \lfloor K/2 \rfloor + 1)}{2} \end{aligned} \tag{B.1}$$

$$\approx \mathcal{O}(K^2). \tag{B.2}$$

Hence, the complexity of the proposed caching algorithm is $\approx \mathcal{O}(K^2)$, which is significantly less than the exhaustive search.

Appendix C

Load Values for $K = 10$

Table C.1: Table for $|C_1|=1$ and $K = 10$

Candidate Sets	Load	Candidate Sets	Load
$\{U_1\}$	5.975	$\{U_6\}$	5.696875
$\{U_2\}$	6.121875	$\{U_7\}$	6.778125
$\{U_3\}$	6.415625	$\{U_8\}$	5.5
$\{U_4\}$	6.553125	$\{U_9\}$	5.43125
$\{U_5\}$	5.3875	$\{U_{10}\}$	5.434375

Table C.2: Table for $|C_1|=2$ and $K = 10$

Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2\}$	3.75625	$\{U_2, U_6\}$	3.315625	$\{U_4, U_5\}$	3.1875	$\{U_6, U_8\}$	3.171875
$\{U_1, U_3\}$	3.79375	$\{U_2, U_7\}$	3.721875	$\{U_4, U_6\}$	3.465625	$\{U_6, U_9\}$	2.8625
$\{U_1, U_4\}$	3.15	$\{U_2, U_8\}$	3.4625	$\{U_4, U_7\}$	4.471875	$\{U_6, U_{10}\}$	3.453125
$\{U_1, U_5\}$	3.25	$\{U_2, U_9\}$	3.4	$\{U_4, U_8\}$	3.06875	$\{U_7, U_8\}$	3.33125
$\{U_1, U_6\}$	3.028125	$\{U_2, U_{10}\}$	3.546875	$\{U_4, U_9\}$	3.553125	$\{U_7, U_9\}$	3.453125
$\{U_1, U_7\}$	3.775	$\{U_3, U_4\}$	3.628125	$\{U_4, U_{10}\}$	3.234375	$\{U_7, U_{10}\}$	3.43125
$\{U_1, U_8\}$	3.8	$\{U_3, U_5\}$	3.24375	$\{U_5, U_6\}$	2.78125	$\{U_8, U_9\}$	2.975
$\{U_1, U_9\}$	3.51875	$\{U_3, U_6\}$	3.18125	$\{U_5, U_7\}$	3.371875	$\{U_8, U_{10}\}$	3.03125
$\{U_1, U_{10}\}$	3.61875	$\{U_3, U_7\}$	3.871875	$\{U_5, U_8\}$	2.9875	$\{U_9, U_{10}\}$	3.421875
$\{U_2, U_3\}$	3.559375	$\{U_3, U_8\}$	3.584375	$\{U_5, U_9\}$	3.096875		
$\{U_2, U_4\}$	3.45625	$\{U_3, U_9\}$	3.225	$\{U_5, U_{10}\}$	3.134375		
$\{U_2, U_5\}$	3.240625	$\{U_3, U_{10}\}$	2.85625	$\{U_6, U_7\}$	3.565625		

Table C.3: Table for $|C_1|=3$ and $K=10$

Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2, U_3\}$	2.359375	$\{U_1, U_7, U_8\}$	2.325	$\{U_2, U_7, U_10\}$	2.165625	$\{U_4, U_6, U_7\}$	2.684375
$\{U_1, U_2, U_4\}$	1.94375	$\{U_1, U_7, U_9\}$	2.296875	$\{U_2, U_8, U_9\}$	2.134375	$\{U_4, U_6, U_8\}$	2.05
$\{U_1, U_2, U_5\}$	2.3875	$\{U_1, U_7, U_10\}$	2.3625	$\{U_2, U_8, U_10\}$	2.253125	$\{U_4, U_6, U_9\}$	2.06875
$\{U_1, U_2, U_6\}$	2.15625	$\{U_1, U_8, U_9\}$	2.3	$\{U_2, U_9, U_10\}$	2.5375	$\{U_4, U_6, U_10\}$	2.25625
$\{U_1, U_2, U_7\}$	2.35	$\{U_1, U_8, U_10\}$	2.384375	$\{U_3, U_4, U_5\}$	2.253125	$\{U_4, U_7, U_8\}$	2.259375
$\{U_1, U_2, U_8\}$	2.61875	$\{U_1, U_9, U_10\}$	2.528125	$\{U_3, U_4, U_6\}$	2.175	$\{U_4, U_7, U_9\}$	2.634375
$\{U_1, U_2, U_9\}$	2.31875	$\{U_2, U_3, U_4\}$	2.2625	$\{U_3, U_4, U_7\}$	2.7875	$\{U_4, U_7, U_10\}$	2.384375
$\{U_1, U_2, U_10\}$	2.659375	$\{U_2, U_3, U_5\}$	2.321875	$\{U_3, U_4, U_8\}$	2.146875	$\{U_4, U_8, U_9\}$	2.00625
$\{U_1, U_3, U_4\}$	2.209375	$\{U_2, U_3, U_6\}$	2.18125	$\{U_3, U_4, U_9\}$	2.38125	$\{U_4, U_8, U_10\}$	1.86875
$\{U_1, U_3, U_5\}$	2.340625	$\{U_2, U_3, U_7\}$	2.28125	$\{U_3, U_4, U_10\}$	1.925	$\{U_4, U_9, U_10\}$	2.30625
$\{U_1, U_3, U_6\}$	2.125	$\{U_2, U_3, U_8\}$	2.43125	$\{U_3, U_5, U_6\}$	1.884375	$\{U_5, U_6, U_7\}$	1.815625
$\{U_1, U_3, U_7\}$	2.5625	$\{U_2, U_3, U_9\}$	2.134375	$\{U_3, U_5, U_7\}$	2.371875	$\{U_5, U_6, U_8\}$	1.94375
$\{U_1, U_3, U_8\}$	2.634375	$\{U_2, U_3, U_10\}$	2.03125	$\{U_3, U_5, U_8\}$	2.196875	$\{U_5, U_6, U_9\}$	1.81875
$\{U_1, U_3, U_9\}$	2.334375	$\{U_2, U_4, U_5\}$	2.153125	$\{U_3, U_5, U_9\}$	2.23125	$\{U_5, U_6, U_10\}$	2.065625
$\{U_1, U_3, U_10\}$	2.203125	$\{U_2, U_4, U_6\}$	2.303125	$\{U_3, U_5, U_10\}$	1.940625	$\{U_5, U_7, U_8\}$	2.04375
$\{U_1, U_4, U_5\}$	2.0375	$\{U_2, U_4, U_7\}$	2.565625	$\{U_3, U_6, U_7\}$	2.18125	$\{U_5, U_7, U_9\}$	2.134375
$\{U_1, U_4, U_6\}$	1.771875	$\{U_2, U_4, U_8\}$	2.03125	$\{U_3, U_6, U_8\}$	2.234375	$\{U_5, U_7, U_10\}$	2.08125
$\{U_1, U_4, U_7\}$	2.41875	$\{U_2, U_4, U_9\}$	2.25	$\{U_3, U_6, U_9\}$	1.778125	$\{U_5, U_8, U_9\}$	1.984375
$\{U_1, U_4, U_8\}$	1.928125	$\{U_2, U_4, U_10\}$	2.171875	$\{U_3, U_6, U_10\}$	2.053125	$\{U_5, U_8, U_10\}$	2.021875
$\{U_1, U_4, U_9\}$	2.221875	$\{U_2, U_5, U_6\}$	2.109375	$\{U_3, U_7, U_8\}$	2.365625	$\{U_5, U_9, U_10\}$	2.20625
$\{U_1, U_4, U_10\}$	2.075	$\{U_2, U_5, U_7\}$	2.146875	$\{U_3, U_7, U_9\}$	2.203125	$\{U_6, U_7, U_8\}$	2.109375
$\{U_1, U_5, U_6\}$	1.84375	$\{U_2, U_5, U_8\}$	2.221875	$\{U_3, U_7, U_10\}$	1.8625	$\{U_6, U_7, U_9\}$	1.78125
$\{U_1, U_5, U_7\}$	2.290625	$\{U_2, U_5, U_9\}$	2.2125	$\{U_3, U_8, U_9\}$	2.015625	$\{U_6, U_7, U_10\}$	2.175
$\{U_1, U_5, U_8\}$	2.225	$\{U_2, U_5, U_10\}$	2.378125	$\{U_3, U_8, U_10\}$	1.934375	$\{U_6, U_8, U_9\}$	1.815625
$\{U_1, U_5, U_9\}$	2.296875	$\{U_2, U_6, U_7\}$	2.125	$\{U_3, U_9, U_10\}$	1.99375	$\{U_6, U_8, U_10\}$	2.175
$\{U_1, U_5, U_10\}$	2.290625	$\{U_2, U_6, U_8\}$	2.19375	$\{U_4, U_5, U_6\}$	2.03125	$\{U_6, U_9, U_10\}$	2.2875
$\{U_1, U_6, U_7\}$	1.959375	$\{U_2, U_6, U_9\}$	2.165625	$\{U_4, U_5, U_7\}$	2.4375	$\{U_7, U_8, U_9\}$	1.859375
$\{U_1, U_6, U_8\}$	2.225	$\{U_2, U_6, U_10\}$	2.484375	$\{U_4, U_5, U_8\}$	2.003125	$\{U_7, U_8, U_10\}$	1.915625
$\{U_1, U_6, U_9\}$	1.915625	$\{U_2, U_7, U_8\}$	2.159375	$\{U_4, U_5, U_9\}$	2.190625	$\{U_7, U_9, U_10\}$	2.10625
$\{U_1, U_6, U_10\}$	2.29375	$\{U_2, U_7, U_9\}$	2.0125	$\{U_4, U_5, U_10\}$	2.13125	$\{U_8, U_9, U_10\}$	2.028125

Table C.4: Table for $|C_1|=4$ and $K=10$

Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load	Candidate Sets	Load
$\{U_1, U_2, U_3, U_4\}$	1.153125	$\{U_1, U_4, U_5, U_{10}\}$	1.371875	$\{U_2, U_4, U_5, U_7\}$	1.33125	$\{U_3, U_5, U_7, U_8\}$	1.353125
$\{U_1, U_2, U_3, U_5\}$	1.5	$\{U_1, U_4, U_6, U_7\}$	1.175	$\{U_2, U_4, U_5, U_8\}$	1.296875	$\{U_3, U_5, U_7, U_9\}$	1.3875
$\{U_1, U_2, U_3, U_6\}$	1.346875	$\{U_1, U_4, U_6, U_8\}$	1.246875	$\{U_2, U_4, U_5, U_9\}$	1.365625	$\{U_3, U_5, U_7, U_{10}\}$	1.1375
$\{U_1, U_2, U_3, U_7\}$	1.30625	$\{U_1, U_4, U_6, U_9\}$	1.203125	$\{U_2, U_4, U_5, U_{10}\}$	1.4	$\{U_3, U_5, U_8, U_9\}$	1.334375
$\{U_1, U_2, U_3, U_8\}$	1.6125	$\{U_1, U_4, U_6, U_{10}\}$	1.31875	$\{U_2, U_4, U_6, U_7\}$	1.553125	$\{U_3, U_5, U_8, U_{10}\}$	1.28125
$\{U_1, U_2, U_3, U_9\}$	1.3375	$\{U_1, U_4, U_7, U_8\}$	1.275	$\{U_2, U_4, U_6, U_8\}$	1.328125	$\{U_3, U_5, U_9, U_{10}\}$	1.35625
$\{U_1, U_2, U_3, U_{10}\}$	1.365625	$\{U_1, U_4, U_7, U_9\}$	1.5	$\{U_2, U_4, U_6, U_9\}$	1.33125	$\{U_3, U_6, U_7, U_8\}$	1.353125
$\{U_1, U_2, U_4, U_5\}$	1.25625	$\{U_1, U_4, U_7, U_{10}\}$	1.325	$\{U_2, U_4, U_6, U_{10}\}$	1.459375	$\{U_3, U_6, U_7, U_9\}$	0.965625
$\{U_1, U_2, U_4, U_6\}$	1.203125	$\{U_1, U_4, U_8, U_9\}$	1.321875	$\{U_2, U_4, U_7, U_8\}$	1.275	$\{U_3, U_6, U_7, U_{10}\}$	1.125
$\{U_1, U_2, U_4, U_7\}$	1.225	$\{U_1, U_4, U_8, U_{10}\}$	1.228125	$\{U_2, U_4, U_7, U_9\}$	1.340625	$\{U_3, U_6, U_8, U_9\}$	1.09375
$\{U_1, U_2, U_4, U_8\}$	1.2	$\{U_1, U_4, U_9, U_{10}\}$	1.4625	$\{U_2, U_4, U_7, U_{10}\}$	1.290625	$\{U_3, U_6, U_8, U_{10}\}$	1.328125
$\{U_1, U_2, U_4, U_9\}$	1.196875	$\{U_1, U_5, U_6, U_7\}$	1.053125	$\{U_2, U_4, U_8, U_9\}$	1.23125	$\{U_3, U_6, U_9, U_{10}\}$	1.209375
$\{U_1, U_2, U_4, U_{10}\}$	1.25	$\{U_1, U_5, U_6, U_8\}$	1.35625	$\{U_2, U_4, U_8, U_{10}\}$	1.159375	$\{U_3, U_7, U_8, U_9\}$	1.1375
$\{U_1, U_2, U_5, U_6\}$	1.359375	$\{U_1, U_5, U_6, U_9\}$	1.28125	$\{U_2, U_4, U_9, U_{10}\}$	1.375	$\{U_3, U_7, U_8, U_{10}\}$	1.04375
$\{U_1, U_2, U_5, U_7\}$	1.353125	$\{U_1, U_5, U_6, U_{10}\}$	1.3875	$\{U_2, U_5, U_6, U_7\}$	1.14375	$\{U_3, U_7, U_9, U_{10}\}$	1.05625
$\{U_1, U_2, U_5, U_8\}$	1.55625	$\{U_1, U_5, U_7, U_8\}$	1.321875	$\{U_2, U_5, U_6, U_8\}$	1.3625	$\{U_3, U_8, U_9, U_{10}\}$	1.128125
$\{U_1, U_2, U_5, U_9\}$	1.49375	$\{U_1, U_5, U_7, U_9\}$	1.415625	$\{U_2, U_5, U_6, U_9\}$	1.31875	$\{U_4, U_5, U_6, U_7\}$	1.346875
$\{U_1, U_2, U_5, U_{10}\}$	1.659375	$\{U_1, U_5, U_7, U_{10}\}$	1.353125	$\{U_2, U_5, U_6, U_{10}\}$	1.565625	$\{U_4, U_5, U_6, U_8\}$	1.35
$\{U_1, U_2, U_6, U_7\}$	1.15	$\{U_1, U_5, U_8, U_9\}$	1.471875	$\{U_2, U_5, U_7, U_8\}$	1.18125	$\{U_4, U_5, U_6, U_9\}$	1.25625
$\{U_1, U_2, U_6, U_8\}$	1.5	$\{U_1, U_5, U_8, U_{10}\}$	1.534375	$\{U_2, U_5, U_7, U_9\}$	1.190625	$\{U_4, U_5, U_6, U_{10}\}$	1.453125
$\{U_1, U_2, U_6, U_9\}$	1.3125	$\{U_1, U_5, U_9, U_{10}\}$	1.61875	$\{U_2, U_5, U_7, U_{10}\}$	1.271875	$\{U_4, U_5, U_7, U_8\}$	1.35
$\{U_1, U_2, U_6, U_{10}\}$	1.575	$\{U_1, U_6, U_7, U_8\}$	1.303125	$\{U_2, U_5, U_8, U_9\}$	1.346875	$\{U_4, U_5, U_7, U_9\}$	1.39375
$\{U_1, U_2, U_7, U_8\}$	1.36875	$\{U_1, U_6, U_7, U_9\}$	1.0375	$\{U_2, U_5, U_8, U_{10}\}$	1.465625	$\{U_4, U_5, U_7, U_{10}\}$	1.36875
$\{U_1, U_2, U_7, U_9\}$	1.240625	$\{U_1, U_6, U_7, U_{10}\}$	1.29375	$\{U_2, U_5, U_9, U_{10}\}$	1.559375	$\{U_4, U_5, U_8, U_9\}$	1.253125
$\{U_1, U_2, U_7, U_{10}\}$	1.384375	$\{U_1, U_6, U_8, U_9\}$	1.340625	$\{U_2, U_6, U_7, U_8\}$	1.1875	$\{U_4, U_5, U_8, U_{10}\}$	1.284375
$\{U_1, U_2, U_8, U_9\}$	1.50625	$\{U_1, U_6, U_8, U_{10}\}$	1.5625	$\{U_2, U_6, U_7, U_9\}$	1.0625	$\{U_4, U_5, U_9, U_{10}\}$	1.421875
$\{U_1, U_2, U_8, U_{10}\}$	1.615625	$\{U_1, U_6, U_9, U_{10}\}$	1.53125	$\{U_2, U_6, U_7, U_{10}\}$	1.26875	$\{U_4, U_6, U_7, U_8\}$	1.375
$\{U_1, U_3, U_4, U_5\}$	1.684375	$\{U_1, U_7, U_8, U_9\}$	1.284375	$\{U_2, U_6, U_8, U_9\}$	1.265625	$\{U_4, U_6, U_7, U_9\}$	1.353125
$\{U_1, U_3, U_4, U_6\}$	1.421875	$\{U_1, U_7, U_8, U_{10}\}$	1.309375	$\{U_2, U_6, U_8, U_{10}\}$	1.44375	$\{U_4, U_6, U_7, U_{10}\}$	1.465625
$\{U_1, U_3, U_4, U_7\}$	1.21875	$\{U_1, U_7, U_9, U_{10}\}$	1.409375	$\{U_2, U_6, U_9, U_{10}\}$	1.6125	$\{U_4, U_6, U_8, U_9\}$	1.2125
$\{U_1, U_3, U_4, U_8\}$	1.54375	$\{U_1, U_8, U_9, U_{10}\}$	1.51875	$\{U_2, U_7, U_8, U_9\}$	1.08125	$\{U_4, U_6, U_8, U_{10}\}$	1.340625
$\{U_1, U_3, U_4, U_9\}$	1.3375	$\{U_2, U_3, U_4, U_5\}$	1.415625	$\{U_2, U_7, U_8, U_{10}\}$	1.084375	$\{U_4, U_6, U_9, U_{10}\}$	1.415625
$\{U_1, U_3, U_4, U_{10}\}$	1.45625	$\{U_2, U_3, U_4, U_6\}$	1.409375	$\{U_2, U_7, U_9, U_{10}\}$	1.209375	$\{U_4, U_7, U_8, U_9\}$	1.275
$\{U_1, U_3, U_5, U_6\}$	1.25	$\{U_2, U_3, U_4, U_7\}$	1.45	$\{U_2, U_8, U_9, U_{10}\}$	1.41875	$\{U_4, U_7, U_8, U_{10}\}$	1.15625
$\{U_1, U_3, U_5, U_7\}$	1.275	$\{U_2, U_3, U_4, U_8\}$	1.346875	$\{U_3, U_4, U_5, U_6\}$	1.36875	$\{U_4, U_7, U_9, U_{10}\}$	1.415625
$\{U_1, U_3, U_5, U_8\}$	1.51875	$\{U_2, U_3, U_4, U_9\}$	1.340625	$\{U_3, U_4, U_5, U_7\}$	1.540625	$\{U_4, U_8, U_9, U_{10}\}$	1.209375
$\{U_1, U_3, U_5, U_9\}$	1.5875	$\{U_2, U_3, U_4, U_{10}\}$	1.16875	$\{U_3, U_4, U_5, U_8\}$	1.4	$\{U_5, U_6, U_7, U_8\}$	1.1125
$\{U_1, U_3, U_5, U_{10}\}$	1.546875	$\{U_2, U_3, U_5, U_6\}$	1.35	$\{U_3, U_4, U_5, U_9\}$	1.459375	$\{U_5, U_6, U_7, U_9\}$	0.984375
$\{U_1, U_3, U_6, U_7\}$	1.46875	$\{U_2, U_3, U_5, U_7\}$	1.334375	$\{U_3, U_4, U_5, U_{10}\}$	1.321875	$\{U_5, U_6, U_7, U_{10}\}$	1.153125
$\{U_1, U_3, U_6, U_8\}$	1.228125	$\{U_2, U_3, U_5, U_8\}$	1.48125	$\{U_3, U_4, U_6, U_7\}$	1.5	$\{U_5, U_6, U_8, U_9\}$	1.209375
$\{U_1, U_3, U_6, U_9\}$	1.54375	$\{U_2, U_3, U_5, U_9\}$	1.39375	$\{U_3, U_4, U_6, U_8\}$	1.365625	$\{U_5, U_6, U_8, U_{10}\}$	1.3625
$\{U_1, U_3, U_6, U_{10}\}$	1.171875	$\{U_2, U_3, U_5, U_{10}\}$	1.378125	$\{U_3, U_4, U_6, U_9\}$	1.221875	$\{U_5, U_6, U_9, U_{10}\}$	1.39375
$\{U_1, U_3, U_7, U_8\}$	1.421875	$\{U_2, U_3, U_6, U_7\}$	1.225	$\{U_3, U_4, U_6, U_{10}\}$	1.290625	$\{U_5, U_7, U_8, U_9\}$	1.128125
$\{U_1, U_3, U_7, U_9\}$	1.503125	$\{U_2, U_3, U_6, U_8\}$	1.421875	$\{U_3, U_4, U_7, U_8\}$	1.44375	$\{U_5, U_7, U_8, U_{10}\}$	1.140625
$\{U_1, U_3, U_7, U_{10}\}$	1.446875	$\{U_2, U_3, U_6, U_9\}$	1.171875	$\{U_3, U_4, U_7, U_9\}$	1.5625	$\{U_5, U_7, U_9, U_{10}\}$	1.21875
$\{U_1, U_3, U_8, U_9\}$	1.271875	$\{U_2, U_3, U_6, U_{10}\}$	1.36875	$\{U_3, U_4, U_7, U_{10}\}$	1.21875	$\{U_5, U_8, U_9, U_{10}\}$	1.340625
$\{U_1, U_3, U_8, U_{10}\}$	1.465625	$\{U_2, U_3, U_7, U_8\}$	1.353125	$\{U_3, U_4, U_8, U_9\}$	1.28125	$\{U_6, U_7, U_8, U_9\}$	0.96875
$\{U_1, U_3, U_8, U_{10}\}$	1.44375	$\{U_2, U_3, U_7, U_9\}$	1.1375	$\{U_3, U_4, U_8, U_{10}\}$	1.1	$\{U_6, U_7, U_8, U_{10}\}$	1.2
$\{U_1, U_3, U_9, U_{10}\}$	1.415625	$\{U_2, U_3, U_7, U_{10}\}$	1.040625	$\{U_3, U_4, U_9, U_{10}\}$	1.271875	$\{U_6, U_7, U_9, U_{10}\}$	1.140625
$\{U_1, U_4, U_5, U_6\}$	1.2	$\{U_2, U_3, U_8, U_9\}$	1.2375	$\{U_3, U_5, U_6, U_7\}$	1.10625	$\{U_6, U_8, U_9, U_{10}\}$	1.3375
$\{U_1, U_4, U_5, U_7\}$	1.378125	$\{U_2, U_3, U_8, U_{10}\}$	1.271875	$\{U_3, U_5, U_6, U_8\}$	1.346875	$\{U_7, U_8, U_9, U_{10}\}$	1.021875
$\{U_1, U_4, U_5, U_8\}$	1.275	$\{U_2, U_3, U_9, U_{10}\}$	1.265625	$\{U_3, U_5, U_6, U_9\}$	1.178125		
$\{U_1, U_4, U_5, U_9\}$	1.403125	$\{U_2, U_4, U_5, U_6\}$	1.409375	$\{U_3, U_5, U_6, U_{10}\}$	1.25		

Proof for Equation (5.16)

As mentioned before, $X = X_{in} - U_x(k) + X_1 + \cdots + X_{n-1}$ and $Y = Y_{in} - U_y(k) + Y_1 + \cdots + Y_{n-1}$. Hence, $X \in \mathcal{N}(\mu_1, \sigma^2(n-1))$ and $Y \in \mathcal{N}(\mu_2, \sigma^2(n-1))$. The n^{th} term of the product on R.H.S of (5.12) can be written as:

$$\begin{aligned}
 & P[D_n < R | D_{n-1} < R] \\
 &= P \left[\sqrt{(X + X_n)^2 + (Y + Y_n)^2} < R \mid \sqrt{X^2 + Y^2} < R \right] \\
 &= \frac{\int_{-R}^R \int_{-R-u}^{R-u} \int_{-C_2}^{C_2} \int_{-C_3}^{C_3} f_{XX_n} f_{YY_n}(u, v, w, y) dy dw dv du}{\int_{-R}^R \int_{-C_2}^{C_2} f_{XY}(u, w) dw du} \tag{D.1}
 \end{aligned}$$

where $C_2 = \sqrt{R^2 - u^2}$ and $C_3 = \sqrt{R^2 - (u+v)^2} - w$. Since X_n 's and Y_n 's are independent, X and Y are also independent. Now, (D.1) can be written as follows:

$$\begin{aligned}
 & P[D_n < R | D_{n-1} < R] = \\
 & \frac{\int_{-R}^R \int_{-R-u}^{R-u} \int_{-C_2}^{C_2} \int_{-C_3}^{C_3} f_X(u) f_{X_n}(v) f_Y(w) f_{Y_n}(y) dy dw dv du}{\int_{-R}^R \int_{-C_2}^{C_2} f_X(u) f_Y(w) dw du} \\
 &= \frac{\int_{-R}^R \int_{-R-u}^{R-u} \int_{-C_2}^{C_2} \int_{-C_3}^{C_3} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{v^2}{2\sigma^2}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} e^{-\frac{y^2}{2\sigma^2}} dy dw dv du}{(\sigma\sqrt{2\pi})^2 \int_{-R}^R \int_{-C_2}^{C_2} e^{-\frac{(u-\mu_1)^2}{2\sigma^2(n-1)}} e^{-\frac{(w-\mu_2)^2}{2\sigma^2(n-1)}} dw du} \tag{D.2}
 \end{aligned}$$

Finally, using (D.2), symmetric pdf property of normal distribution and [114], we obtain (5.16).

First-Order Taylor Expansion of Rate

SNR at user j due to UAV-BS i can be written as:

$$\begin{aligned} & \log_2 \left(1 + \frac{P_i}{\delta_{LOS} K_o N_o d_{i,j}^2(k)} \right) \\ &= \log_2 \left(1 + \frac{\frac{P_i}{\delta_{LOS} K_o N_o}}{H^2 + \|U_i(k) - W_j(k)\|^2} \right). \end{aligned} \quad (\text{E.1})$$

Using first-order Taylor approximation of $f(z) = \log_2 \left(1 + \frac{A}{B+z} \right)$ at point $z = z_o$ will be [66]:

$$\begin{aligned} \log_2 \left(1 + \frac{A}{B+z} \right) &\geq \log_2 \left(1 + \frac{A}{B+z_o} \right) \\ &\quad - \frac{(\log_2 e)A(z-z_o)}{(B+z_o)(B+z_o+A)}. \end{aligned} \quad (\text{E.2})$$

Hence, at the m^{th} iteration of Algorithm 9, the first-order Taylor approximation of (E.1) can be obtained by substituting $A = \frac{P_i}{\delta_{LOS} K_o N_o}$, $B = H^2$, $z = \|U_i(k) - W_j(k)\|^2$ and $z_o = \|U_i(k, m-1) - W_j(k)\|^2$ in (E.2).

References

- [1] Future mobile data usage and traffic growth. [Online]. Available: <https://www.ericsson.com/en/mobility-report/future-mobile-data-usage-and-traffic-growth>.
- [2] K. Gomez, L. Goratti, T. Rasheed, and L. Reynaud, "Enabling disaster-resilient 4g mobile communication networks," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 66–73, Dec. 2014.
- [3] G. Baldini, S. Karanasios, D. Allen, and F. Vergari, "Survey of wireless communication technologies for public safety," *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 619–641, Second 2014.
- [4] K. Flynn, "The mobile broadband standard." [Online]. Available: https://www.3gpp.org/news-events/1875-mc_services
- [5] F. S. Shaikh and R. Wismüller, "Routing in multi-hop cellular device-to-device (d2d) networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 2622–2657, Fourthquarter 2018.
- [6] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Rihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015.
- [7] A. Al-Rimawi and D. Dardari, "Analytical characterization of device-to-device and cellular networks coexistence," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5537–5548, Aug. 2017.
- [8] M. Ji, G. Caire, and A. F. Molisch, "Optimal throughput-outage trade-off in wireless one-hop caching networks," in *Proc. IEEE International Symposium on Information Theory*, 2013, pp. 1461–1465.
- [9] T. X. Vu, S. Chatzinotas, and B. Ottersten, "Edge-caching wireless networks: Performance analysis and optimization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2827–2839, Apr. 2018.
- [10] S. T. ul Hassan, M. Bennis, P. H. J. Nardelli, and M. Latva-Aho, "Modeling and analysis of content caching in wireless small cell networks," in *Proc. International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2015, pp. 765–769.

- [11] W. Wen, Y. Cui, F. Zheng, and S. Jin, "Random caching based cooperative transmission in heterogeneous wireless networks," in *Proc. IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.
- [12] R. Wang, J. Zhang, S. H. Song, and K. B. Letaief, "Mobility-aware caching in d2d networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5001–5015, Aug. 2017.
- [13] R. Amer, M. Majid Butt, M. Bennis, and N. Marchetti, "Inter-cluster Cooperation for Wireless D2D Caching Networks," *ArXiv e-prints*, 2018.
- [14] R. Wang, J. Zhang, S. H. Song, and K. B. Letaief, "Exploiting mobility in cache-assisted d2d networks: Performance analysis and optimization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5592–5605, Aug. 2018.
- [15] I. Pappalardo, G. Quer, B. D. Rao, and M. Zorzi, "Caching strategies in heterogeneous networks with d2d, small bs and macro bs communications," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.
- [16] "Study on further enhancements to LTE device to device (D2D), user equipment (UE) to network relays for internet of things (IoT) and wearables version 15.1.1," Third Gener. Partnership Project, Sophia Antipolis, France, 3GPP Rep. TR 36.746, 2018, Tech. Rep.
- [17] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Stochastic geometry study on device-to-device communication as a disaster relief solution," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3005–3017, May 2016.
- [18] L. Chen, S. H. Low, M. Chiang, and J. C. Doyle, "Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks," in *Proc. 25TH IEEE International Conference on Computer Communications*, Apr. 2006, pp. 1–13.
- [19] D. Ebrahimi, H. Elbiaze, and W. Ajib, "Device-to-device data transfer through multihop relay links underlying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9669–9680, Oct. 2018.
- [20] Z. Kaleem, M. Yousaf, A. Qamar, A. Ahmad, T. Q. Duong, W. Choi, and A. Jamalipour, "Uav-empowered disaster-resilient edge architecture for delay-sensitive communication," *IEEE Network*, vol. 33, no. 6, pp. 124–132, 2019.
- [21] A. S. Abdalla and V. Marojevic, "Communications standards for unmanned aircraft systems: The 3GPP perspective and research drivers," 2021.
- [22] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on uav cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019.
- [23] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing uav," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329–2345, 2019.

- [24] L. Ruan, J. Wang, J. Chen, Y. Xu, Y. Yang, H. Jiang, Y. Zhang, and Y. Xu, "Energy-efficient multi-UAV coverage deployment in UAV networks: A game-theoretic framework," *China Communications*, vol. 15, no. 10, pp. 194–209, 2018.
- [25] B. Wang, R. Zhang, C. Chen, X. Cheng, L. Yang, H. Li, and Y. Jin, "Graph-based file dispatching protocol with D2D-enhanced UAV-NOMA communications in large-scale networks," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8615–8630, 2020.
- [26] Y. Lin, T. Wang, and S. Wang, "UAV-assisted emergency communications: An extended multi-armed bandit perspective," *IEEE Communications Letters*, vol. 23, no. 5, pp. 938–941, 2019.
- [27] Q. Wang, W. Zhang, Y. Liu, and Y. Liu, "Multi-UAV dynamic wireless networking with deep reinforcement learning," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2243–2246, 2019.
- [28] K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "Outage probability analysis for device-to-device system," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–5.
- [29] J. Guo, S. Durrani, X. Zhou, and H. Yanikomeroglu, "Device-to-device communication underlying a finite cellular network region," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 332–347, Jan. 2017.
- [30] G. George, R. K. Mungara, and A. Lozano, "An analytical framework for device-to-device communication in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6297–6310, Nov. 2015.
- [31] Winner ii channel models. [Online]. Available: https://www.researchgate.net/publication/234055761_WINNER_II_channel_models
- [32] X. Zhao, S. Li, X. Liang, Q. Wang, L. Hentilä, and J. Meinilä, "Measurements and modelling for d2d indoor wideband mimo radio channels at 5 ghz," *IET Communications*, vol. 10, no. 14, pp. 1839–1845, 2016.
- [33] S. L. Cotton, "A statistical characterization of shadowed device-to-device communications in an indoor environment," in *Proc. 8th European Conference on Antennas and Propagation (EuCAP 2014)*, Apr. 2014, pp. 3362–3365.
- [34] P. Janis, C. Yu, K. Doppler, C. Ribeiro, C. Wijting, K. Hugl, O. Tirkkonen, and V. Koivunen, "Device-to-device communication underlying cellular communications systems," *International Journal of Communications, Network and System Sciences*, 2009.
- [35] B. Peng, T. Peng, Z. Liu, Y. Yang, and C. Hu, "Cluster-based multicast transmission for device-to-device (d2d) communication," in *Proc. IEEE 78th Vehicular Technology Conference (VTC Fall)*, Sep. 2013, pp. 1–5.
- [36] M. Condoluci, L. Militano, G. Araniti, A. Molinaro, and A. Iera, "Multicasting in lte-a networks enhanced by device-to-device communications," in *Proc. IEEE Globecom Workshops*, Dec. 2013.

- [37] L. Jiang, H. Tian, Z. Xing, K. Wang, K. Zhang, S. Maharjan, S. Gjessing, and Y. Zhang, "Social-aware energy harvesting device-to-device communications in 5g networks," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 20–27, Aug. 2016.
- [38] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware d2d communications: qualitative insights and quantitative analysis," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 150–158, Jun. 2014.
- [39] L. Wang, G. Araniti, C. Cao, W. Wang, and Y. Liu, "Device-to-device users clustering based on physical and social characteristics," *International Journal of Distributed Sensor Networks*, vol. 11, no. 8, p. 165608, 2015. [Online]. Available: <https://doi.org/10.1155/2015/165608>
- [40] G. Zhang, K. Yang, and H. H. Chen, "Socially aware cluster formation and radio resource allocation in d2d networks," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 68–73, Aug. 2016.
- [41] L. Feng, P. Zhao, F. Zhou, M. Yin, P. Yu, W. Li, and X. Qiu, "Resource allocation for 5g d2d multicast content sharing in social-aware cellular networks," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 112–118, Mar. 2018.
- [42] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4346–4359, Aug. 2014.
- [43] B. Chen and C. Yang, "Caching policy optimization for d2d communications by learning user preference," in *Proc. IEEE 85th Vehicular Technology Conference (VTC Spring)*, Jun. 2017, pp. 1–6.
- [44] J. Li, M. Liu, J. Lu, F. Shu, Y. Zhang, S. Bayat, and D. N. K. Jayakody, "On social-aware content caching for d2d-enabled cellular networks with matching theory," *IEEE Internet of Things Journal*, pp. 1–1, 2018.
- [45] P. Ren, Q. Du, and L. Sun, "Interference-aware routing for hop-count minimization in wireless D2D networks," in *Proc. IEEE/CIC International Conference on Communications in China - Workshops (CIC/ICCC)*, Aug. 2013, pp. 65–70.
- [46] H. Yuan, W. Guo, and S. Wang, "Emergency route selection for D2D cellular communications during an urban terrorist attack," in *Proc. IEEE International Conference on Communications Workshops (ICC)*, Jun. 2014, pp. 237–242.
- [47] S. Wang, W. Guo, Z. Zhou, Y. Wu, and X. Chu, "Outage probability for multi-hop D2D communications with shortest path routing," *IEEE Communications Letters*, vol. 19, no. 11, pp. 1997–2000, Nov. 2015.
- [48] K. Ali, H. X. Nguyen, P. Shah, Q. Vien, and E. Ever, "D2D multi-hop relaying services towards disaster communication system," in *Proc. 24th International Conference on Telecommunications (ICT)*, May 2017, pp. 1–5.
- [49] M. Tanha, D. Sajjadi, F. Tong, and J. Pan, "Disaster management and response for modern cellular networks using flow-based multi-hop device-to-device communications," in *Proc. IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Sep. 2016, pp. 1–7.

- [50] X. Lin and S. B. Rasool, “Distributed and provably efficient algorithms for joint channel-assignment, scheduling, and routing in multichannel ad hoc wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 17, no. 6, pp. 1874–1887, Dec. 2009.
- [51] C. Buratti and R. Verdone, “Joint routing and scheduling for centralised wireless sensor networks,” in *Proc. IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*, Sep. 2016, pp. 1–6.
- [52] C. J. Katila and C. Buratti, “A novel routing and scheduling algorithm for multi-hop heterogeneous wireless networks,” in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Jun. 2018, pp. 1–6.
- [53] A. Chowdary, Y. Ramamoorthi, A. Kumar, and L. R. Cenkeramaddi, “Joint resource allocation and uav scheduling with ground radio station sleeping,” *IEEE Access*, vol. 9, pp. 124 505–124 518, 2021.
- [54] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, “A tutorial on uavs for wireless networks: Applications, challenges, and open problems,” 2018.
- [55] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, “Efficient 3-d placement of an aerial base station in next generation cellular networks,” in *Proc. IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–5.
- [56] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage,” *IEEE Communications Letters*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.
- [57] M. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, “3-d placement of an unmanned aerial vehicle base station (uav-bs) for energy-efficient maximal coverage,” *IEEE Wireless Communications Letters*, vol. 6, no. 4, pp. 434–437, Aug. 2017.
- [58] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, “Placement optimization of uav-mounted mobile base stations,” *IEEE Communications Letters*, vol. 21, no. 3, pp. 604–607, 2017.
- [59] E. Kalantari, I. Bor-Yaliniz, A. Yongacoglu, and H. Yanikomeroglu, “User association and bandwidth allocation for terrestrial and aerial base stations with backhaul considerations,” in *Proc. IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.
- [60] J. Lu, S. Wan, X. Chen, and P. Fan, “Energy-efficient 3d uav-bs placement versus mobile users’ density and circuit power,” in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–6.
- [61] F. Lagum, I. Bor-Yaliniz, and H. Yanikomeroglu, “Strategic densification with uav-bss in cellular networks,” *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 384–387, 2018.
- [62] B. Galkin, J. Kibilda, and L. A. DaSilva, “Deployment of uav-mounted access points according to spatial user locations in two-tier cellular networks,” in *Proc. Wireless Days (WD)*, 2016, pp. 1–6.

- [63] X. Sun and N. Ansari, "Jointly optimizing drone-mounted base station placement and user association in heterogeneous networks," in *Proc. IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.
- [64] C. Zhan, Y. Zeng, and R. Zhang, "Trajectory design for distributed estimation in uav-enabled wireless sensor network," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 10 155–10 159, 2018.
- [65] J. Cui, Z. Ding, Y. Deng, A. Nallanathan, and L. Hanzo, "Adaptive UAV-trajectory optimization under quality of service constraints: A model-free solution," *IEEE Access*, vol. 8, pp. 112 253–112 265, 2020.
- [66] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.
- [67] F. Zeng, Z. Hu, Z. Xiao, H. Jiang, S. Zhou, W. Liu, and D. Liu, "Resource allocation and trajectory optimization for QoE provisioning in energy-efficient UAV-enabled wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7634–7647, 2020.
- [68] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2020.
- [69] H. Qi, Z. Hu, H. Huang, X. Wen, and Z. Lu, "Energy efficient 3-D UAV control for persistent communication service and fairness: A deep reinforcement learning approach," *IEEE Access*, vol. 8, pp. 53 172–53 184, 2020.
- [70] H. V. Abeywickrama, Y. He, E. Dutkiewicz, B. A. Jayawickrama, and M. Mueck, "A reinforcement learning approach for fair user coverage using uav mounted base stations under energy constraints," *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 67–81, 2020.
- [71] S.-F. Chou, A.-C. Pang, and Y.-J. Yu, "Energy-aware 3D unmanned aerial vehicle deployment for network throughput optimization," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 563–578, 2020.
- [72] L. Liu, S. Zhang, and R. Zhang, "Comp in the sky: Uav placement and movement optimization for multi-user communications," *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5645–5658, 2019.
- [73] R. Ghanavi, E. Kalantari, M. Sabbaghian, H. Yanikomeroğlu, and A. Yongacoglu, "Efficient 3d aerial base station placement considering users mobility by reinforcement learning," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.
- [74] X. Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-UAV networks: Deployment and movement design," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8036–8049, 2019.

- [75] X. Liu, Y. Liu, Y. Chen, and L. Hanzo, "Trajectory design and power control for multi-uav assisted wireless networks: A machine learning approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7957–7969, 2019.
- [76] Etsi tr 136 942 v9.1.0. [Online]. Available: http://www.etsi.org/deliver/etsi_tr/136900_136999/136942/09.01.00_60/tr_136942v090100p.pdf
- [77] Y. Zhang, E. Pan, L. Song, W. Saad, Z. Dawy, and Z. Han, "Social network aware device-to-device communication in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 177–190, Jan. 2015.
- [78] Z. Wang and V. W. S. Wong, "A novel d2d data offloading scheme for lte networks," in *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2015, pp. 3107–3112.
- [79] M. Peer, V. A. Bohara, and A. Srivastava, "On the performance of network-assisted indoor device-to-device communication using location awareness and realistic path loss models," in *Proc. IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–7.
- [80] W. Xiang, K. Zheng, and X. S. Shen, Eds., *5G Mobile Communications*. Springer, 2017.
- [81] Lte tdd frame configuration. [Online]. Available: http://rfmw.em.keysight.com/wireless/helpfiles/n7625b/Content/Main/3_LTE_TDD_Frame_Configuration.htm
- [82] What a p-value tells you about statistical data. [Online]. Available: <http://www.dummies.com/education/math/statistics/what-a-p-value-tells-you-about-statistical-data/>
- [83] H. P. Keeler, "Notes on the poisson point process," University of Melbourne, Tech. Rep., 2015.
- [84] Linear models in statistics. [Online]. Available: [http://www.utstat.toronto.edu/~sim\\$brunner/books/LinearModelsInStatistics.pdf](http://www.utstat.toronto.edu/~sim$brunner/books/LinearModelsInStatistics.pdf)
- [85] The coefficient of determination, r-squared. [Online]. Available: <https://onlinecourses.science.psu.edu/stat501/node/255>
- [86] A. Orsino, A. Samuylov, D. Moltchanov, S. Andreev, L. Militano, G. Araniti, and Y. Koucheryavy, "Time-dependent energy and resource management in mobility-aware d2d-empowered 5g systems," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 14–22, Aug. 2017.
- [87] S. Hosny, A. Eryilmaz, and H. E. Gamal, "Impact of user mobility on d2d caching networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [88] M. Ahmed, Y. Li, M. Waqas, M. Sheraz, D. Jin, and Z. Han, "A survey on socially-aware device-to-device communications," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2018.
- [89] A. Konrad, B. Y. Zhao, A. D. Joseph, and R. Ludwig, "A markov-based channel model algorithm for wireless networks," *Wireless Networks*, vol. 9, no. 3, pp. 189–199, May 2003. [Online]. Available: <https://doi.org/10.1023/A:1022869025953>

- [90] E. B. Iversen, J. K. Møller, J. M. Morales, and H. Madsen, “Inhomogeneous markov models for describing driving patterns,” *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 581–588, Mar. 2017.
- [91] What the future holds: machine learning, navigating the great indoors and safer vehicles. [Online]. Available: <https://www2.deloitte.com/au/en/pages/media-releases/articles/what-the-future-holds-tmt-predictions-240117.html>
- [92] V. L. Erickson, M. A. Carreira-Perpiñán, and A. E. Cerpa, “Occupancy modeling and prediction for building energy management,” *ACM Trans. Sen. Netw.*, vol. 10, no. 3, pp. 42:1–42:28, May 2014. [Online]. Available: <http://doi.acm.org/10.1145/2594771>
- [93] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, “Cellular traffic offloading through opportunistic communications: A case study,” in *Proc. 5th ACM Workshop on Challenged Networks*, ser. CHANTS ’10. New York, NY, USA: ACM, 2010, pp. 31–38. [Online]. Available: <http://doi.acm.org/10.1145/1859934.1859943>
- [94] M. Kodialam and T. Nandagopal, “Characterizing achievable rates in multi-hop wireless mesh networks with orthogonal channels,” *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 868–880, 2005.
- [95] P. Dutta, V. Mhatre, D. Panigrahi, and R. Rastogi, “Joint routing and scheduling in multi-hop wireless networks with directional antennas,” in *IEEE INFOCOM*, 2010, pp. 1–5.
- [96] Jihui Zhang, Haitao Wu, Qian Zhang, and Bo Li, “Joint routing and scheduling in multi-radio multi-channel multi-hop wireless networks,” in *Proc. International Conference on Broadband Networks, 2005.*, Oct. 2005, pp. 631–640 Vol. 1.
- [97] S. Sarkar, M. H. R. Khouzani, and K. Kar, “Optimal routing and scheduling in multihop wireless renewable energy networks,” *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1792–1798, 2013.
- [98] S. Büyükçorak, G. K. Kurt, and A. Yongaçoğlu, “Uav assisted ground user localization,” in *2019 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)*, Oct. 2019, pp. 111–115.
- [99] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1990.
- [100] M. X. Cheng, X. Gong, Y. Xu, and L. Cai, “Link activity scheduling for minimum end-to-end latency in multihop wireless sensor networks,” in *IEEE Global Telecommunications Conference - GLOBECOM*, 2011, pp. 1–5.
- [101] H. He, S. Zhang, Y. Zeng, and R. Zhang, “Joint altitude and beamwidth optimization for uav-enabled multiuser communications,” *IEEE Communications Letters*, vol. 22, no. 2, pp. 344–347, Feb. 2018.
- [102] Q. Wu, L. Liu, and R. Zhang, “Fundamental trade-offs in communication and trajectory design for uav-enabled wireless network,” *IEEE Wireless Communications*, vol. 26, no. 1, pp. 36–44, Feb. 2019.

- [103] C. P. Hoffman, W. Cox, and T. J. Pack, “Dual-satellite emergency locator beacon and method for registering, programming and updating emergency locator beacon over the air,” US Patent, US9 031 497B1.
- [104] I. Bor-Yaliniz and H. Yanikomeroglu, “The new frontier in ran heterogeneity: Multi-tier drone-cells,” 2016.
- [105] C. T. Cicek, H. Gultekin, B. Tavli, and H. Yanikomeroglu, “Uav base station location optimization for next generation wireless networks: Overview and future research directions,” 2018.
- [106] B. Li, Z. Fei, and Y. Zhang, “Uav communications for 5g and beyond: Recent advances and future trends,” *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241–2263, Apr. 2019.
- [107] N. Aschenbruck, E. Gerhards-Padilla, M. Gerharz, M. Frank, and P. Martini, “Modelling mobility in disaster area scenarios,” in *Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, ser. MSWiM ’07. New York, NY, USA: ACM, 2007, pp. 4–12. [Online]. Available: <http://doi.acm.org/10.1145/1298126.1298131>
- [108] S. Burer and A. Saxena, “The MILP road to MIQCP,” in *Mixed Integer Nonlinear Programming*, J. Lee and S. Leyffer, Eds. New York, NY: Springer New York, 2012, pp. 373–405.
- [109] M. Peer, V. A. Bohara, and A. Srivastava, “Multi-UAV placement strategy for disaster-resilient communication network,” in *Proc. IEEE VTC Fall Workshop*, 2020.
- [110] M. Peer, V. A. Bohara, A. Srivastava, and G. Ghatak, “User mobility-aware time stamp for UAV-BS placement,” in *Proc. IEEE WCNC Workshop*, 2021.
- [111] M. Peer, V. A. Bohara, and A. Srivastava, “Cache selection in dynamic D2D multicast networks using inhomogeneous markov model,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3235–3245, 2020.
- [112] H. Tabassum, M. Salehi, and E. Hossain, “Fundamentals of mobility-aware performance characterization of cellular networks: A tutorial,” *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2288–2308, 2019.
- [113] M. Banagar and H. S. Dhillon, “Performance characterization of canonical mobility models in drone cellular networks,” *IEEE Transactions on Wireless Communications*, 2020.
- [114] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.
- [115] T. D. Humphries, R. D. Haynes, and L. A. James, “Simultaneous and sequential approaches to joint optimization of well placement and control,” *Computational Geosciences*, vol. 18, no. 3-4, p. 433–448, 2013.
- [116] M. Peer, V. A. Bohara, and A. Srivastava, “Real-world spatio-temporal behavior aware d2d multicast networks,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1675–1686, 2020.

- [117] ———, “Multi-hop d2d framework for disaster-resilient communication network,” in *Proc. IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 584–589.
- [118] ———, “Enabling disaster-resilient communication using multi-hop device-to-device framework,” Oct. 2020. [Online]. Available: <https://link.springer.com/article/10.1007/s11276-020-02481-2>
- [119] Y. Liu, K. Liu, J. Han, L. Zhu, Z. Xiao, and X.-G. Xia, “Resource allocation and 3-d placement for uav-enabled energy-efficient iot communications,” *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1322–1333, 2021.
- [120] N. Cherif, W. Jaafar, H. Yanikomeroglu, and A. Yongacoglu, “Disconnectivity-aware energy-efficient cargo-uav trajectory planning with minimum handoffs,” in *Proc. IEEE International Conference on Communications*, 2021, pp. 1–6.
- [121] L. Jiang and H. Jafarkhani, “Reconfigurable intelligent surface assisted mmwave uav wireless cellular networks,” in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [122] Introduction to transformation of random variables. [Online]. Available: http://nptel.ac.in/courses/108106083/lecture14_Introduction_to_Transformation_of_RVs.pdf