AFRL-RV-PS-
TR-2023-0037

AFRL-RV-PS-
TR-2023-0037

COMMUNICATIONS-THEORETIC FOUNDATIONS FOR ADVANCED EHF SATELLITE SYSTEMS

Sudharman K. Jayaweera

Department of Electrical Engineering and Computer Engineering
University of New Mexico
Albuquerque, NM 87131-0001

24 January 2023

Final Report

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776

DTIC COPY

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research which is exempt from public affairs security and policy review in accordance with AFI 61-201, paragraph 2.3.5.1. This report is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RV-PS-TR-2023-0037 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//	//SIGNED//
Dr. Steven A. Lane Program Manager, AFRL/RVB	For: Mark E. Roverse, Chief AFRL Geospace Technologies Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
24-01-2023	Final Report	07 Jan 2021 - 16 Mar 2023

4. TITLE AND SUBTITLE
Communications-Theoretic Foundations for Advanced EHF Satellite Systems

5a. CONTRACT NUMBER
FA9453-21-2-0019

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
63401F

6. AUTHOR(S)
Sudharman K. Jayaweera

5d. PROJECT NUMBER
3682

5e. TASK NUMBER
EF134102

5f. WORK UNIT NUMBER
V1MF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Electrical Engineering and Computer Engineering
University of New Mexico
Albuquerque, NM 87131-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Space Vehicles Directorate
3550 Aberdeen Avenue SE
Kirtland AFB, NM 87117-5776

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RVBYC

11. SPONSOR/MONITOR'S REPORT
AFRL-RV-PS-TR-2023-0037

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited (AFRL-2023-3416 dtd 31 Jul 2023).

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Work performed under this cooperative agreement between AFRL Space Vehicles Directorate, and the University of New Mexico focused on fundamental research to establish the communications-theoretic foundations for future Extremely High Frequency (EHF) satellite and terrestrial communications. This project started the investigation of two of the promising new concepts in wireless communications, namely the Intelligent Reflecting Surfaces (IRS) and 3D beamforming, in the context of millimeter wavelength satellite communications. Simulations were conducted to assess the performance of multiple user networks at higher frequencies. Two algorithms, a reward based greedy algorithm and a deep reinforcement learning algorithm implemented with a Deep Q-Network, were proposed for determining an optimal trajectory for a UAV base station to maximize the coverage of distinct ground users. Finally, simulations were conducted to assess performance impacts in an adversarial environment.

15. SUBJECT TERMS
space communication

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Dr. Steven A. Lane
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code),*
Unclassified	Unclassified	Unclassified	Unlimited	74	

This page is intentionally left blank.

TABLE OF CONTENTS

Section **Page**

LIST OF FIGURES ii
LIST OF TABLES v
1 SUMMARY 1
2 INTRODUCTION 2
3 METHODS, ASSUMPTIONS AND PROCEDURES 3
3.1 IRS-Aided mmWave Satellite Communications 3
3.2 IRS Orientation Optimization Problem for a Single Link 5
3.3 IRS Orientation Optimization Problem for Multiple Receiver Locations 16
3.4 3D Beamforming for mmWave UAV Communications 19
3.4.1 3D Beamforming Weight Vector Update Timing 15
3.4.2 Simplified Solution for 2D and 3D Beamforming Weights Update Period 18
3.5 Drone Trajectory Planning with Deep Reinforcement Learning and AI 25
3.5.1 Proposed Methods 1: Deep Reinforcement Learning using DQN 28
3.5.2 Proposed Method 2: Reward Based Greedy Algorithm 26
3.6 Security and Robustness of Machine Learning Based Trajectory Planning Systems 26
4 RESULTS AND DISCUSSION 32
4.1 IRS-Aided 3D Beamforming for mmWave Communications 32
4.2 Coverage area Optimization 35
4.3 Beamformer Weights Update Period Simulation 36
4.3.1 2D UAV Network Simulation 39
4.3.2 3D UAV Network Simulation 43
4.4 Optimal UAV Trajectory Planning 53
4.5 Security and Robustness of Machine Learning Based Trajectory Planning Systems 52
5 CONCLUSIONS 55
REFERENCES 57
LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS 61

LIST OF FIGURES

Figure **Page**

Figure 1. IRS Installation Around a Ground Receiver 4

Figure 2. IRS Angle and Distance Definitions 6

Figure 3. IRS Angle and Distance Definitions from the mx, my-th IRS 7

Figure 4. An IRS Deployed to Improve Reception Within an Area with Weak Signal Strength. 17

Figure 5. A Communications Link Between a UAV Transmitter and a UAV Receiver..... 14

Figure 6. The UAV Transmitter Antenna Beam Pattern and the HPBW 19

Figure 7. Comparison of Δd When the Beam is Directed to (θ, φ) 19

Figure 8. Illustration of UAV Hovering Geography with Assumed Location at (500, 500)..... 21

Figure 9. A GMM User Distribution with Four Clusters 22

Figure 10. A GMM User Distribution with Eight Clusters and High User Density..... 23

Figure 11. A State That Consists of the UAV's Location, the Remaining Energy and the Number of Users Covered at Each Hovering Point..... 24

Figure 12. A State That Consists of the UAV's Location, the Energy Remaining in the UAV, the Number of Users Covered for Time Instance 't' Along with Same Information for Nine Previous Time Steps 25

Figure 13. The DQN Algorithm..... 26

Figure 14. Detection of Ground Users by the UAV-BS 27

Figure 15. The State of the UAV-BS at t-th Time Step..... 29

Figure 16. A Square Area with Genuine and Fake Users..... 31

Figure 17. The State Matrix of the k-th Adversary at Epoch e 31

Figure 18. Variation of the Received Power in Watts from Tx-ITS-Rx Path with θn and φn 32

Figure 19. Variation of the Received Power from Tx-ITS-Rx Path for a Specific Location Tx and Rx With Respect to the rIr 33

Figure 20. Variation of the Received Power as a Function of Orientation Angles, $\theta n, \varphi n$ 34

Figure 21. The Optimal IRS Orientation Angles..... 34

Figure 22. Average Received Power in Selected Area from Tx-ITS-Rx Path for a Specific Tx Location 35

Figure 23. Average Received Power in Selected Area from Tx-ITS-Rx Path for a Specific Tx Location with Respect to the Average Distance rIr 36

Figure 24. Variation of Error Between Exact and Approximate Δt with Respect to the HPBW of an Antenna..... 37

LIST OF FIGURES (Continued)

Figure **Page**

Figure 25. Beamforming Update Time Period Variation with Frequency for Different Receiver Locations with Fixed Transmitter Aperture 100-cm^2, ω = 0.33 rad/s, and r = 30-m 38

Figure 26. Ratio of Overhead Bit Rate and Available Bit Rate 39

Figure 27. Variation of the Normalized Beamforming Gain with Time when Using $\Delta t*$ and $\Delta tapprox$ 40

Figure 28. Variation of the Normalize Beamforming Gain with Time When Using Approximate Equation in the Simulation Model when $\Delta t = E\{\Delta tapprox(b)\}$ 41

Figure 29. Variation of the Normalize Beamforming Gain with Time When Using Approximate Equation in the Simulation Model when $\Delta t = minb \in \mathcal{B}\{\Delta tapprox(b)\}$ 41

Figure 30. Histogram of Total Time Duration Where GB, tx Drops Below 0.5 When $\Delta t = E\Delta tapproxb$ 42

Figure 31. Histogram of Number of Beamformer Weights Update Times When $\Delta t = E\{\Delta tapprox(b)\}$ 43

Figure 32. Histogram of Number of Beamformer Weights Update Times When $\Delta t = minb \in \mathcal{B}\{\Delta tapprox(b)\}$ 43

Figure 33. Variation of GB, tx with Time When Using Approximation $\Delta tapprox$ and Optimal $\Delta t*$ 44

Figure 34. Variation of the Normalize Beamforming Gain GB, tx With Time When Using $\Delta t = \min a \in \mathcal{A}\{\Delta tapprox(a)\}$ in 3D Simulation Model 45

Figure 35. Variation of the Normalize Beamforming Gain GB, tx With Time When Using $\Delta t = E\{\Delta tapprox(a)\}$ in 3D Simulation Model 46

Figure 36. Histogram of Total Time Duration Where GB, tx Drops Below 0.5 When $\boldsymbol{\Delta t} = E\{\Delta tapprox(a)\}$ 47

Figure 37. Histogram of Number of Beamformer Weights Update Times When $\Delta t = E\{\Delta tapprox(a)\}$ 47

Figure 38. Histogram of Number of Beamformer Weights Update Times When $\boldsymbol{\Delta t} = mina \in A$ $\{\Delta tapprox\ (a)\}$ 48

Figure 39. Number of Users Covered for Varying Number of Total Ground Users Uniformly Distributed Over the Area of Interest 49

Figure 40. Average and the Variance of Number of Users Covered Using Proposed Methods When Users are Distributed According to Gaussian Mixtures 49

Figure 41. CCDF of Number of Users Covered When Users are Distributed According to Gaussian Mixtures 50

Figure 42. Number of Users Covered and the Variance in Number of Users Covered with Densely Packed Users Distributed According to Gaussian Mixtures 50

LIST OF FIGURES (Continued)

Figure **Page**

Figure 43. CCDF of Number of Users Covered When Users are Densely Packed and Distributed According to Gaussian Mixtures 51

Figure 44. Number of Users Covered When Users are Mobile and Distributed According to Gaussian Mixtures 51

Figure 45. CCDF of Number of Users Covered When Users are Mobile and Distributed According to Gaussian Mixtures 52

Figure 46. Trajectory of the UAV-BS in the (a) Absence and (b) Presence of Adversary 53

Figure 47. Time Spent at Each Location in Presence and Absence of Adversaries 53

LIST OF TABLES

Table **Page**

Table 1. Performance of the UAV-BS With Stationary Ground Users in the Presence and Absence of Adversaries 54

Table 2. Performance of the UAV-BS With Non-stationary Ground Users in the Presence and Absence of Adversaries 54

ACKNOWLEDGMENTS

This material is based on research sponsored by Air Force Research Laboratory under agreement number FA9453-21-2-0019. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

DISCLAIMER

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

1 SUMMARY

The objective of this project is to conduct fundamental research that will establish the communications-theoretic foundations for future Extremely High Frequency (EHF) satellite and terrestrial communications. Previous efforts in EHF band communications rely on existing communication theoretic approaches that were developed decades ago in the context of mobile wireless systems operating in lower frequencies in the Super High Frequency (SHF) spectrum, in particular, the so-called "sub-6 GHz" spectrum. This project plans to evaluate the validity and applicability of existing physical (PHY) and medium-access control (MAC) layer designs to EHF satellite and terrestrial communications systems and conduct fundamental research in identifying communications-theoretic foundations for future EHF band communications.

This report investigates three emerging concepts in wireless communications, intelligent reflecting surfaces (IRS), three-dimensional (3D) beamforming, and machine learning (ML) for the EHF band systems. All these are viewed as promising techniques to make the milli-meter wave (mmWave) terrestrial communications a reality. This project develops these techniques in the context of space and airborne communications in which platform mobility is an essential aspect. As a result, satellite to ground communications and unmanned aerial vehicle (UAV) communications with trajectory planning are investigated as potential applications.

2 INTRODUCTION

As the lower frequency bands in the Super High Frequency (SHF) spectrum (3 – 30 GHz) that are traditionally used for most terrestrial and satellite operations become ever more congested, there is a great incentive to migrate future protected tactical and strategic military satellite communications (MILSATCOM) systems to currently underutilized EHF spectrum (30 – 300 GHz), also called the mmWave spectrum. However, the mmWave spectrum is also already being explored for civilian applications including 5G and beyond 5G (5G/B5G) cellular mobile communications, Wireless Gigabit (WiGig) systems, rapidly expanding drone communications and vehicular radar and communications making spectrum efficiency and coexistence of importance. Add to that the fundamental difference in high attenuation of mmWaves compared to their microwave counterparts, there is a need for optimal techniques for power and spectrum utilization.

Not surprisingly, initial attempts in using mmWave band for wireless communications have relied on existing communication theoretic approaches that were developed decades ago in the context of mobile wireless systems operating in the very high frequency (VHF), ultra high frequency (UHF) and super high frequency (SHF) spectrum bands. For instance, the common multipath channel models of Rayleigh and Rician fading were primarily motivated through measurements in sub-6 GHz systems [1]. Similarly, most existing physical (PHY) and medium-access control (MAC) layer designs were motivated by the assumption of equally limited power and bandwidth constraints [2]. The mmWave spectrum, on the other hand, allows for potentially large bandwidth operation [3]. This raises the interesting question whether new techniques may allow better utilization of mmWave spectrum compared to legacy approaches.

In the case of satellite communications, there is an inherent constraint on power consumption even if wideband operation is possible in mmWave frequencies [4-6]. For example, the spectrum allocated for satellite services in 81 – 86 GHz and 71 – 76 GHz in W and V bands, respectively, allows for potentially GHz-wide signals to support high-rate communications. However, limited transmit power combined with the high attenuation of mmWave signals raise the question whether reasonable link budgets can be achieved, especially in geosynchronous earth orbit (GEO) operation. Even if a system design based on existing PHY and MAC layer designs were to perform acceptably, there is the possibility that it may not be optimally utilizing the enormous bandwidth available and may not be optimally exploiting the asymmetric bandwidth-power tradeoff.

Two of the promising new approaches that have gained significant attention in terrestrial wireless systems are the intelligent reflecting surfaces (IRS) and three-dimensional (3D), or full-dimension (FD), beamforming. The IRSs are passive reflecting arrays whose reflective properties can be controlled. Their placement in the channel between the transmitter and a receiver can help modify the wireless channel seen by the receiver. One way to make use of them is to control the IRS to result in passive beamforming. However, traditional beamforming is limited to a single plane (e.g., horizontal plane) and is called 2D beamforming. In contrast, 3D beamforming controls the signal beam in both azimuth and elevation. Such 3D beamforming combined with IRSs may be

particularly helpful in mmWave communications links due to the potential for focusing transmit energy more narrowly towards a desired receiver to counteract their high attenuation.

The proliferation of drone/UAV systems in recent years has made reliable drone communications an important emerging topic of interest. Their applications span a wide range of contexts including surveillance, remote sensing, disaster relief and other advanced aerial mobility (AAM) systems. Although many of these rely on lower frequencies for the moment, in future they may also be benefitted by mmWave communications. Depending on the application, the requirements on drone communications systems may vary widely.

Driven by the recent advances in computing hardware and software, artificial intelligence (AI) and machine learning (ML) based protocols have found increasing applications in optimizing the performance of many emerging and 5G mobile wireless systems. In this project, we explored AI/ML as a potential tool for developing efficient communications and networking protocols. In particular, such approaches are extremely appealing in the context of drone/UAV systems where autonomous trajectory learning and collaborative behavior can be highly desirable. However, online learning has the potential for exposing the system to malicious attacks. The many vulnerabilities of ML based communications and network optimization protocols have received relatively less attention. In this project, we investigate security and robustness of ML based protocols in mmWave communications systems by developing both spoofing techniques and possible counter-spoofing approaches.

3 METHODS, ASSUMPTIONS AND PROCEDURES

3.1 IRS-Aided mmWave Satellite Communications

The concept of IRS for wireless communications has gained considerable attention in recent years due to its ability to fundamentally change the way a communications channel is looked at in classical communications theory [7, 8]. Indeed, until now the basic assumption in communications system design has been that the channel is beyond the control of the designer and is the fundamental constraint [7]. The use of an IRS, however, allows one to control the wireless channel to some extent in ways that is not possible otherwise. In particular, a carefully placed and controlled IRS can modify the multipath profile seen by a receiver thereby improving the link performance [7, 9].

It appears that the IRSs can be an ideal technique to enable mmWave satellite communications for several reasons [10, 11, 12]. A set of IRSs, appropriately placed near a receiver, can be used to improve the received signal power through advanced signal processing techniques [13, 14]. The shorter wavelengths of mmWave signals make it possible to deploy compact IRSs consisting of a large number of elements, thereby leading to high overall gain. Moreover, the resulting narrow signal beamwidths possible with mmWave frequencies can help focus signal energy toward the receiver more efficiently [15, 16]. These advantages of IRSs can be instrumental in overcoming link budget constraints imposed by limited transmit power on satellite platforms.

Figure 1 shows a system made of a satellite transmitter, a ground receiver and a number of IRSs (N_{IRS}) located near the receiver. The orientation of an IRS at a particular location, defined by the normal vector to that IRS, is important to ensure reception of maximum power at the intended receiver. The orientation vectors given in Figure 1 (in black arrows) are defined with respect to the local tangent plane (LTP) in the ENU (East-North-Up) coordinate system where the origin is located at the middle of the IRS. Therefore, the i^{th} IRS orientation vector is denoted by $(1, \theta_n^i, \varphi_n^i)$. The position vectors given in Figure 1 (in blue symbols) are defined in the ENU coordinate system whose origin is considered at the receiver. Therefore, the i^{th} IRS position vector is defined as $(r_r^i, \theta_{rx}^i, \varphi_{rx}^i)$ where r_r^i, θ_{rx}^i, φ_{rx}^i are the distance between the i^{th} IRS and the receiver, the elevation angle of the i^{th} IRS and the azimuth angle of the i^{th} IRS, respectively. The transmitter position vector is defined as $(r_d^i, \theta_{tx}^i, \varphi_{tx}^i)$ where r_d^i, θ_{tx}^i, φ_{tx}^i are the distance between the transmitter and the receiver, the elevation angle of the transmitter and the azimuth angle of the transmitter, respectively.

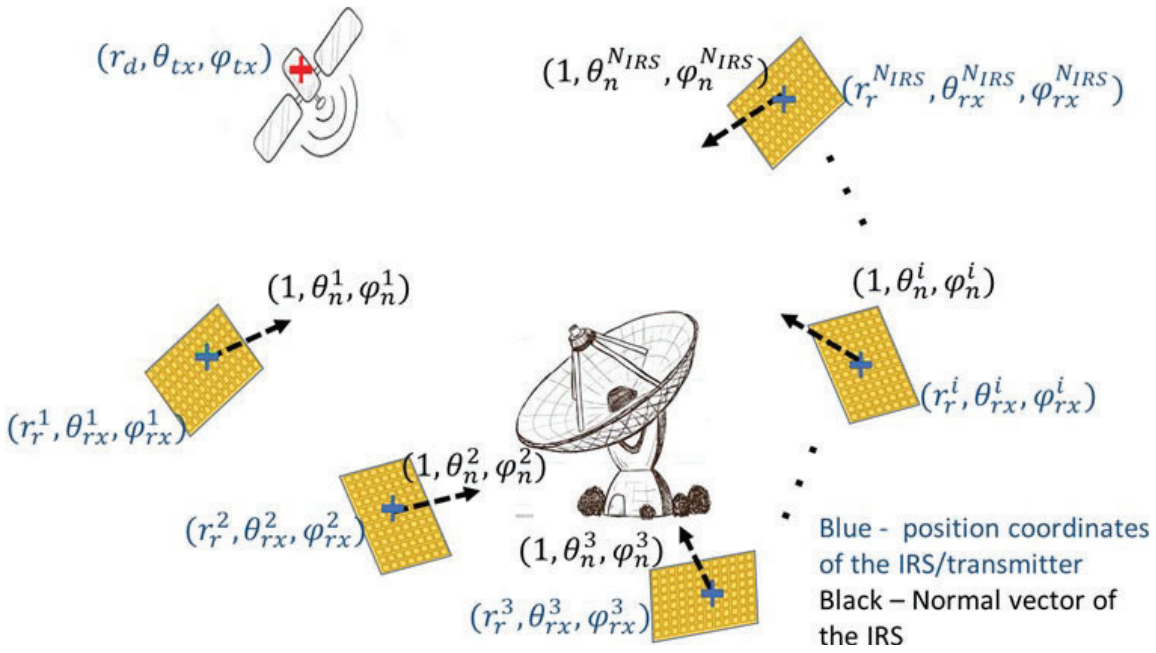


Figure 1. IRS Installation Around a Ground Receiver

The total power received at the ground receiver P_{tot} can be written as:

$$P_{tot} = P_d + \sum_{i=1}^{N_{IRS}} P_r^i \tag{1}$$

where P_d is the power received from the direct path between the transmitting satellite and the ground receiver, and P_r^i is the reflected power received from the i^{th} IRS [9]:

$$P_r^i = \frac{P_t G_{atmos} G_t G_r G d_x d_y \lambda^2}{64\pi^3} \left| \sum_{m=1-M_i/2}^{M_i/2} \sum_{n=1-N_i/2}^{N_i/2} \frac{\sqrt{{F_{i,nm}}^{combine}} A_{nm}^i}{r_{t,nm}^i r_{r,nm}^i} \; e^{-j\frac{2\pi\left(r_{t,nm}^i r_{r,nm}^i\right) - \lambda\phi_{nm}^i}{\lambda}} \right|^2 \quad (2)$$

where G_{atmos} is the atmospheric gain/loss due to gas and water vapour, G_t is the satellite antenna transmitter gain, G_r is the ground receiver antenna gain, G is the IRS element gain (assumed to be the same for all elements), λ is the wavelength of the transmitted signal, $\Gamma_{nm}^{\mathrm{i}} = A_{nm}^i \mathrm{e}^{j\phi_{nm}^i}$ is the reflectance coefficient of the (n, m)-th element of the i^{th} IRS with the magnitude A_{nm}^i and phase ϕ_{nm}^i, M_i and N_i are the number of elements of the i^{th} IRS along each axis, d_x and d_y are the dimensions of an IRS element and $r_{t,nm}^i$ and $r_{r,nm}^i$ are the distance from the (n, m)-th element of the i^{th} IRS to transmitter and receiver, respectively. Moreover,

$${F_{i,nm}}^{combine} = F^{tx}\left(\theta_{\mathrm{tx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{tx,nm}}^{\mathrm{i}}\right) F\left(\theta_{\mathrm{t,nm}}^{\mathrm{i}}, \varphi_{\mathrm{t,nm}}^{\mathrm{i}}\right) F\left(\theta_{\mathrm{r,nm}}^{\mathrm{i}}, \varphi_{\mathrm{r,nm}}^{\mathrm{i}}\right) F^{rx}\left(\theta_{\mathrm{rx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{rx,nm}}^{\mathrm{i}}\right) \quad (3)$$

where $F\left(\theta_{\mathrm{t,nm}}^{\mathrm{i}}, \varphi_{\mathrm{t,nm}}^{\mathrm{i}}\right)$ and $F\left(\theta_{\mathrm{r,nm}}^{\mathrm{i}}, \varphi_{\mathrm{r,nm}}^{\mathrm{i}}\right)$ are the normalized radiation patterns of the IRS and $F^{tx}\left(\theta_{\mathrm{tx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{tx,nm}}^{\mathrm{i}}\right)$ and the $F^{rx}\left(\theta_{\mathrm{rx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{rx,nm}}^{\mathrm{i}}\right)$ are the normalized radiation patterns of the transmitter and receiver respectively. As an example, normalized radiation function of each element of an IRS can be modeled as [9],

$$F(\theta, \varphi) = \begin{cases} \cos^3 \theta & for \quad 0 \le \theta \le \frac{\pi}{2} \\ 0 & else \end{cases} \quad (4)$$

The IRSs are assumed to be placed closer to the ground receiver antenna. In general, the radiation pattern $F^{tx}\left(\theta_{\mathrm{tx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{tx,nm}}^{\mathrm{i}}\right)$ can vary with the location of IRS. However, since the location of the IRS, satellite transmitter (GEO) and ground receiver are fixed and the distance r_d is large, $F^{tx}\left(\theta_{\mathrm{tx,nm}}^{\mathrm{i}}, \varphi_{\mathrm{tx,nm}}^{\mathrm{i}}\right)$ can be assumed to be constant for the optimization problem.

3.2 IRS Orientation Optimization Problem for a Single Link

Consider a communication system with a transmitter having M_{tx} number of transmit antenna elements where $M_{tx} = M_{tx}^x M_{tx}^y$ and a receiver having M_{rx} number of receiver antenna elements where $M_{rx} = M_{rx}^x M_{rx}^y$. Let us denote by M_I the number of elements in the deployed IRS where $M_I = M_I^x M_I^y$. The numbers of elements in the transmitter antenna's horizontal and vertical directions are assumed to be M_{tx}^x and M_{tx}^y, respectively. Similarly, M_{rx}^x, M_{rx}^y and M_I^x, M_I^y are defined for the receiver antenna elements and IRS elements, respectively. Receiver antenna elements are indexed with the ℓ^x, ℓ^y for $0 < l^x < M_{rx}^x - 1$ and $0 < l^y < M_{rx}^y - 1$. Let $\underline{l}^x = [0 \quad \cdots \quad M_{rx}^x - 1]$ and $\underline{l}^y = [0 \quad \cdots \quad M_{rx}^y - 1]$ denote the vectors that define the indices of the receiver antenna elements in each directions. The received complex baseband signal $Y_r \in \mathbb{C}^{M_{rx} \times 1}$ can be expressed as:

$$Y_r = Y_r^d + Y_r^I + Z_r \quad (5)$$

where $Y_r^d \in \mathbb{C}^{M_{rx} \times 1}$ is the received complex baseband signal from the line of sight (LOS) path, $Y_r^I \in \mathbb{C}^{M_{rx} \times 1}$ is the received complex baseband signal from the Tx-IRS-Rx path and $Z_r \in \mathbb{C}^{M_{rx} \times 1}$ is

the additive noise. Transmitter and IRS elements are indexed with the (k^x, k^y) and (m^x, m^y) respectively as in Figure 2, for $0 < k^x < M_{tx}^x - 1$, $0 < k^y < M_{tx}^y - 1$, $0 < m^x < M_I^x - 1$ and $0 < m^y < M_I^y - 1$. Let $\underline{k}^x = [0 \quad \cdots \quad M_{tx}^x - 1]$ and $\underline{k}^y = [0 \quad \cdots \quad M_{tx}^y - 1]$ denote the vectors of indices of the transmitter antenna elements in each direction. Similarly, let, $\underline{m}^x = [0 \quad \cdots \quad M_I^x - 1]$ and $\underline{m}^y = [0 \quad \cdots \quad M_I^y - 1]$ denote the vectors of indices of the IRS elements in each direction as in Figure 2.

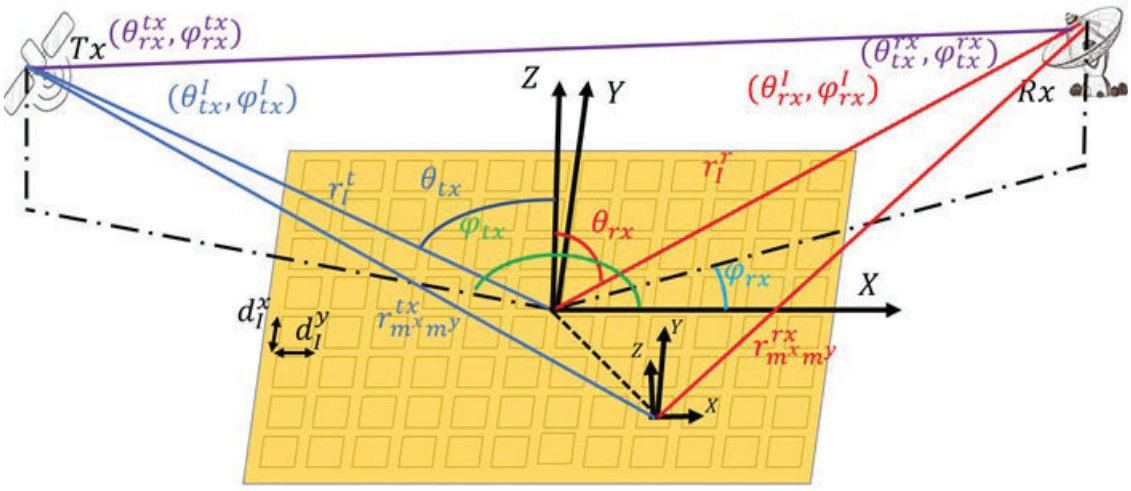


Figure 2. IRS Angle and Distance Definitions

We may write:

$$Y_r^d = \sqrt{p_s} H_d X \tag{6}$$

and

$$Y_r^I = \sqrt{p_s} H_{rx} diag(\Gamma) H_{tx} X \tag{7}$$

where $X \in \mathbb{C}^{M_{tx} \times 1}$ is the transmit signal vector, $\Gamma \in \mathbb{C}^{M_I \times 1}$ is the IRS reflecting coefficient vector, $H_d \in \mathbb{C}^{M_{rx} \times M_{tx}}$ is the LOS channel coefficients matrix from the transmitter to the receiver, p_s is the transmit symbol power, $H_{tx} \in \mathbb{C}^{M_I \times M_{tx}}$ is the channel coefficients matrix from the transmitter to the IRS and $H_{rx} \in \mathbb{C}^{M_{rx} \times M_I}$ is the channel coefficients matrix from the IRS to the receiver.

The IRS orientation is important specifically if the locations of the transmitter and the receiver of the communication system are not varying with the time (e.g., GEO satellite communication [17]). When there is a direct line-of-sight (LOS) link between the receiver and transmitter, received power can be further improved by installing an IRS. However, when a direct LOS link between

the receiver and transmitter is not available, the effectiveness of the communication system can be much higher with the installation of the IRS, since it provides a link between the receiver and transmitter (e.g., mobile communication [13]). For both scenarios, improving received signal power from the IRS can improve the performance of the communication system. Therefore, the LOS direct link between the receiver and transmitter is ignored in the rest of this discussion.

Let channel coefficient matrix H_{tx} be expressed as:

$$H_{tx} = \alpha_{tx} \underline{a}_I^{tx} \otimes \left(\underline{a}_{tx}^I\right)^H \tag{8}$$

where $\alpha_{tx} \in \mathbb{C}^{1\times 1}$ is the LOS path gain from the transmitter antenna to the IRS, $\underline{a}_{tx}^I \in \mathbb{C}^{M_{tx}\times 1}$ is the array response vector at the transmitter antenna based on the IRS location, and $\underline{a}_I^{tx} \in \mathbb{C}^{M_I\times 1}$ is the array response vector at the IRS based on receive signal from the transmitter.

The LOS path gain α_{tx} can be expressed as:

$$\alpha_{tx} = \sqrt{\frac{g_t f^{tx}(\theta_{tx}^I, \varphi_{tx}^I) f(\theta_{m^x m^y}^{tx}, \varphi_{m^x m^y}^{tx}) d_I^x d_I^y}{4\pi {r_I^t}^2}} \tag{9}$$

where g_t is the transmitter antenna element gain, θ_{tx}^I and φ_{tx}^I are the zenith and azimuth angles of the IRS measured from the transmitter as in Figure 2, $\theta_{m^x m^y}^{tx}$ and $\varphi_{m^x m^y}^{tx}$ are the zenith and azimuth angles of the transmitter measured from the (m^x, m^y)-th IRS as in Figure 3, $f^{tx}(\theta_{tx}^I, \varphi_{tx}^I)$ is the normalized beam pattern of the transmitter antenna element, r_I^t is the distance between the IRS and transmitter antenna, $f\left(\theta_{m^x m^y}^{tx}, \varphi_{m^x m^y}^{tx}\right)$ is the normalized beam pattern of the (m^x, m^y)-th IRS element, and d_I^x, d_I^y are the IRS element sizes in each directions, respectively.

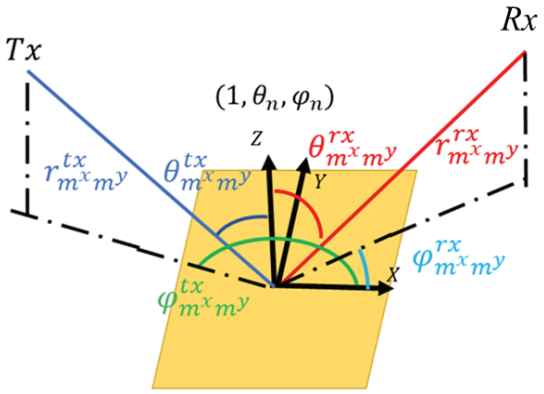


Figure 3. IRS Angle and Distance Definitions from the (m^x, m^y)-th IRS

The array response vectors $\underline{a}_I^{tx}$ and $\underline{a}_{tx}^I$ can be expressed as:

$$\underline{a}_I^{tx} = e(\theta_{m^x m^y}^{tx}, \varphi_{m^x m^y}^{tx}, M_I^x, M_I^y, d_I^x, d_I^y) \tag{10}$$

and

$$\underline{a}_{tx}^{I} = e(\theta_{tx}^{I}, \varphi_{tx}^{I}, M_{tx}^{x}, M_{tx}^{y}, d_{tx}^{x}, d_{tx}^{y}) \tag{11}$$

where d_{tx}^{x} and d_{tx}^{y} are the transmitter element spaces in each direction, respectively, and we have defined the function e(.) as:

$$e(\theta, \varphi, M^{x}, M^{y}, d^{x}, d^{y}) = \underline{p}^{y} \otimes \underline{p}^{x} \tag{12}$$

where

$$\underline{p}^{y} = \left[1, e^{j2\pi(1)(sin\theta sin\varphi)d^{y}/\lambda}, \dots, e^{j2\pi(M^{y}-1)(sin\theta sin\varphi)d^{y}/\lambda}\right]^{T} \tag{13}$$

θ and φ are the zenith and azimuth angles, respectively,

$$\underline{p}^{x} = \left[1, e^{j2\pi(1)(sin\theta cos\varphi)d^{x}/\lambda}, \dots, e^{j2\pi(M^{x}-1)(sin\theta cos\varphi)d^{x}/\lambda}\right]^{T} \tag{14}$$

M^{x} and M^{y} are the number of elements in the vertical and horizontal directions of the IRS/transmitter/receiver, respectively, and d^{x} and d^{y} are the element spacing along the vertical and horizontal directions of the IRS/transmitter/receiver, respectively.

Similarly, let channel coefficient matrix H_{rx} be expressed as:

$$H_{rx} = \alpha_{rx}\underline{a}_{rx}^{I} \otimes \left(\underline{a}_{I}^{rx}\right)^{H} \tag{15}$$

where $\alpha_{rx} \in \mathbb{C}$ is the LOS path gain from the receiver antenna to the IRS, $\underline{a}_{rx}^{I} \in \mathbb{C}^{M_{rx}\times 1}$ is the array response vector at the receiver antenna based on receive signal from the IRS, and $\underline{a}_{I}^{rx} \in \mathbb{C}^{M_{I}\times 1}$ is the array response vector at the IRS based on the receiver location.

The LOS path gain α_{rx} can be expressed as:

$$\alpha_{rx} = \sqrt{\frac{Gg_{r}\lambda^{2}f^{rx}(\theta_{rx}^{I}, \varphi_{rx}^{I})f(\theta_{m^{x}m^{y}}^{rx}, \varphi_{m^{x}m^{y}}^{rx})}{16\pi^{2}{r_{I}^{t}}^{2}}} \tag{16}$$

where G is the IRS element gain, g_{r} is the receiver antenna element gain, θ_{rx}^{I} and φ_{rx}^{I} are the zenith and azimuth angle of the IRS measured from the receiver as in Figure 2, $\theta_{m^{x}m^{y}}^{rx}$ and $\varphi_{m^{x}m^{y}}^{rx}$ are the zenith and azimuth angles of the receiver measured from the (m^{x}, m^{y})-th IRS element as in Figure 3, $f^{rx}(\theta_{rx}^{I}, \varphi_{rx}^{I})$ is the normalized beam pattern of the receiver antenna element, r_{I}^{t} is the distance between the IRS and receiver antenna, $f\left(\theta_{m^{x}m^{y}}^{rx}, \varphi_{m^{x}m^{y}}^{rx}\right)$ is the normalized beam patterns of the (m^{x}, m^{y})-th IRS element, and d_{I}^{x}, d_{I}^{y} are the IRS element sizes in each directions, respectively.

As before, the array response vector $\underline{a}_{I}^{rx}$and $\underline{a}_{rx}^{I}$ can be expressed as:

$$\underline{a}_{I}^{rx} = e(\theta_{m^{x}m^{y}}^{rx}, \varphi_{m^{x}m^{y}}^{rx}, M_{I}^{x}, M_{I}^{y}, d_{I}^{x}, d_{I}^{y}) \tag{17}$$

and

$$\underline{a}_{rx}^{I} = e(\theta_{rx}^{I}, \varphi_{rx}^{I}, M_{rx}^{x}, M_{rx}^{y}, d_{rx}^{x}, d_{rx}^{y}) \tag{18}$$

where d_{rx}^{x} and d_{rx}^{y} are the receive antenna element spacing in each direction, respectively.

When the transmitter, receiver and IRS are placed adequately far away, the spherical wave emitted from the antenna/IRS can be approximated as a plane wave. Hence, the angles defined in Figure 3 for (m^x, m^y)-th IRS element can assumed to be the same for each element in IRS. The reflection coefficient of the (m^x, m^y)-th IRS element can be expressed as:

$$\Gamma_{m^x m^y} = A_{m^x m^y} e^{j\phi_{m^x m^y}} \tag{19}$$

where $A_{m^x m^y}$ is the magnitude of the (m^x, m^y)-th IRS element and $\phi_{m^x m^y}$ is the reflection phase shift of the (m^x, m^y)-th element of the IRS. This enables IRS to steer the reflected beam towards the receiver antenna. To direct the reflected beam towards the receiver antenna, the reflection phase shift of the (m^x, m^y)-th IRS element need to be set as [9]:

$$\phi_{m^x m^y} = \frac{2\pi}{\lambda}\left((m^x - 1) d_I^x p + (m^y - 1) d_I^y q \right) \tag{20}$$

where

$$p = \sin\left(\theta_{m^x m^y}^{tx}\right) \cos\left(\varphi_{m^x m^y}^{tx}\right) - \sin\left(\theta_{m^x m^y}^{rx}\right) \cos\left(\varphi_{m^x m^y}^{rx}\right) \tag{21}$$

and

$$q = \sin\left(\theta_{m^x m^y}^{tx}\right) \sin\left(\varphi_{m^x m^y}^{tx}\right) - \sin\left(\theta_{m^x m^y}^{rx}\right) \sin\left(\varphi_{m^x m^y}^{rx}\right). \tag{22}$$

As seen, the angles used to calculate the p and q depend on the orientation of the IRS. However, the optimal phase shifts above result in canceling out each IRS element reflection coefficient with the array response of the IRS. Hence, when optimal reflection coefficients are used, the optimal IRS orientation does not depend on neither the IRS array response nor the IRS element reflection coefficient.

The beam pattern of an IRS element depends on the material, size, and structure [9, 18]. The normalized beam pattern of the IRS element can be represented as $f(\theta, \varphi)$ in the spherical coordinate system where θ and φ are the zenith and azimuth angles from the IRS element to a particular transmitting/receiving direction. While there are several possible approximations that have been considered for the IRS beam pattern $f(\theta, \varphi)$ one of the most commonly used is the following [19]:

$$f(\theta, \varphi) \; = \; 10^{-1.2A(\theta,\varphi)} \tag{23}$$

where

$$A(\theta, \varphi) = \begin{cases} \left(\frac{\theta_n - \theta}{\theta_{3dB}}\right)^2 + \left(\frac{\varphi_n - \varphi}{\varphi_{3dB}}\right)^2 & if \quad \|\varphi_n - \varphi\| \le \pi \\ \left(\frac{\theta_n - \theta}{\theta_{3dB}}\right)^2 + \left(\frac{\varphi_n - \varphi}{\varphi_{3dB}}\right)^2 & \quad\;\; \|\varphi_n - \varphi\| > \pi \end{cases} \tag{24}$$

for $0 \le \theta < \pi$ and $0 \le \varphi < 2\pi$. Note that, θ_{3dB} and φ_{3dB} are the vertical and horizontal 3 dB beamwidths, θ and φ angles are measured from the IRS to the transmitter/receiver with respect to a coordinate system whose origin is at the IRS. The normalized IRS beam pattern is maximized

when $\theta = \theta_n$ and $\varphi = \varphi_n$. The gain of the IRS element is the amount of power transmitted/received from in the direction of peak radiation relative to that of an isotropic radiator.

Far-field approximation of the angles shown in the Figure 1 are measured with respect to the local tangent plane (LTP) east-north-up (ENU) coordinate system where the origin is at the middle of the IRS. Figure 3 illustrates the angle definitions of the (m^x, m^y)-th IRS element. The angles of the beam patterns, $f\left(\theta^{tx}_{m^x m^y}, \varphi^{tx}_{m^x m^y}\right)$ and $f\left(\theta^{rx}_{m^x m^y}, \varphi^{rx}_{m^x m^y}\right)$ are measured with respect to the coordinate system whose Z-axis is normal to the IRS surface. As a result, the angles in Figure 2 depend on the orientation vector of the IRS. Therefore, an optimization problem can be formulated to find the optimal orientation of the IRS to maximize the total received power at the receiver.

The beamformed received signal at the receiver can be written as:

$$\mathrm{y} = W_r^H Y_r \tag{25}$$

where $W_r \in \mathbb{C}^{M_{rx} \times 1}$ is the receiver beamforming vector. The objective is to maximize the received power at the receiver by optimizing the IRS orientation vector, $(1, \theta_n, \varphi_n)$ where θ_n is the zenith angle of the IRS orientation and φ_n is the azimuth angle of the IRS orientation. However, the orientation of the IRS does not influence the received signal from the transmitter-receiver direct path. Therefore, maximizing the received signal from the transmitter-IRS-receiver path eventually maximizes the total received signal power. Hence, the IRS orientation optimization problem can be expressed as:

$$\max_{\theta_n, \varphi_n} |W_r^H Y_r|^2 \tag{26}$$

$$\max_{\theta_n, \varphi_n} \left| \sqrt{\mathrm{p}}\, W_r^T \alpha_{\mathrm{rx}} \underline{\mathrm{a}^{\mathrm{I}}_{\mathrm{rx}}} \otimes \left(\underline{\mathrm{a}^{\mathrm{rx}}_{\mathrm{I}}}\right)^{\mathrm{H}} \mathrm{diag}(\Gamma) \alpha_{\mathrm{tx}} \underline{\mathrm{a}^{\mathrm{tx}}_{\mathrm{I}}} \otimes \left(\underline{\mathrm{a}^{\mathrm{I}}_{\mathrm{tx}}}\right)^{\mathrm{H}} \mathrm{X} \right|^2 \tag{27}$$

The above optimization function can be simplified and expressed in terms of the angle of arrival (AOA) of the impinging signal on IRS elements and angle of departure (AOD) of the reflecting signals from the IRS elements. By using maximal ratio combining (MRC) at the receiver and 3D beamforming at the IRS and the transmitter, the above maximization of the received power at the receiver over orientation angles can be reduced to the following maximization:

$$\max_{\theta_n, \varphi_n} \left\| \sqrt{p_s} \alpha_{rx} \alpha_{tx} \|\Gamma\| \right\|^2 \tag{28}$$

where

$$p_r(s_{rx}, \theta_n, \varphi_n) = \left\| \sqrt{p_s} \alpha_{rx} \alpha_{tx} \|\Gamma\| \right\|^2 \tag{29}$$

can be identified as the maximum received power at a location for a given IRS orientation.

For a specific transmitter and receiver location, it can be shown that there exists a closed-form solution for the IRS orientation angle based on the transmitter and receiver locations with respect to the IRS. Indeed, the exact solution for the IRS orientation optimization problem (i.e., optimal IRS orientation angle) can be derived to be:

$$\varphi_n = \left(\frac{\varphi_{tx}+\varphi_{rx}}{2}\right) \tag{30}$$

and

$$\theta_n = \left(\frac{\theta_{tx}+\theta_{rx}}{2}\right). \tag{31}$$

Since in the high signal to noise ratio (SNR) region the achievable data rate at a location s_{rx} can be expressed as:

$$R(s_{rx},\theta_n,\varphi_n) = \log_2(1+SNR(s_{rx},\theta_n,\varphi_n)) \approx \log_2(SNR(s_{rx},\theta_n,\varphi_n)) \tag{32}$$

where

$$SNR(s_{rx},\theta_n,\varphi_n) = p_r(s_{rx},\theta_n,\varphi_n)/\sigma^2, \tag{33}$$

the above power maximizing solution is also equivalent to maximizing the achievable rate.

From both numerical simulations and above analytical expression, it is seen that the optimal IRS orientation vector always lies between the location vectors of the transmitter and the receiver. Hence, the optimal azimuth and elevation angles of the IRS are always in between the corresponding angles of the location vectors of the transmitter and the receiver (modulo 2π). For example, let us denote by Ω the set of azimuth and elevation angles in between the corresponding angles of the location vectors of the transmitter and the receiver. Then, the expected received power $\mathrm{E}\{p_\Omega\}$ when the IRS is oriented in a random direction between the transmitter and receiver location vectors, can be calculated as:

$$\mathrm{E}\{p_\Omega\} = \frac{1}{\|\theta_{tx}-\theta_{rx}\|\|\varphi_{tx}-\varphi_{rx}\|}\iint_{\Omega_p}\left|\sqrt{p}\alpha_{rx}\alpha_{tx}\,\|\Gamma\|\right|^2 d\theta_n d\varphi_n \tag{34}$$

This can be simplified as:

$$\mathrm{E}\{p_\Omega\} = \frac{\widetilde{B}}{\|\theta_{tx}-\theta_{rx}\|\|\varphi_{tx}-\varphi_{rx}\|}\int_{\theta_{tx}}^{\theta_{rx}}\int_{\varphi_{tx}}^{\varphi_{rx}} f\left(\theta^{tx}_{m^x m^y},\varphi^{tx}_{m^x m^y}\right) f\left(\theta^{rx}_{m^x m^y},\varphi^{rx}_{m^x m^y}\right) d\theta_n d\varphi_n \tag{35}$$

where

$$\widetilde{B} = \frac{p g_t G g_r \lambda^2 f^{tx}(\theta^I_{tx},\varphi^I_{tx})\, f^{rx}(\theta^I_{rx},\varphi^I_{rx})\, d_I^x d_I^y \|\Gamma\|^2}{64\pi^3 {r_I^t}^2 {r_I^t}^2}. \tag{36}$$

3.3 IRS Orientation Optimization Problem for Multiple Receiver Locations

In many practical situations, it might be of interest that the IRS provide improved signal quality to users distributed in a desired geographical area. Consider an IRS placed in a particular location to improve the receive signal power of users located inside a weak signal strength area, A_ϵ (due to obstacles, for example). We may divide the area A_ϵ into a total of N_{rx} number of small areas where each small area is approximately equal to the a_ϵ as shown in Figure 4.

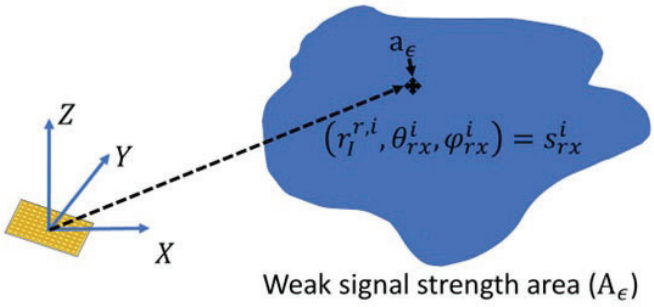


Figure 4. An IRS Deployed to Improve Reception Within an Area with Weak Signal Strength

Consider a set of receiver locations S_{rx} where $s_{rx}^i \in S_{rx}$ is defined as $(r_I^{r,i}, \theta_{rx}^i, \varphi_{rx}^i)$ for $i = 1, .., N_{rx}$. Let $r_I^{r,i}$, θ_{rx}^i and φ_{rx}^i be defined with similar definition as r_I^r, θ_{rx} and φ_{rx}, respectively, except the receiver location is now defined as the i^{th} receiver location. The received power at the i^{th} receiver location (from the transmitter-IRS-receiver path) can be expressed in terms of the parameters $r_I^{r,i}, \theta_{rx}^i, \varphi_{rx}^i, \theta_n$ and , φ_n as:

$$p(s_{rx}^i, \theta_n, \varphi_n) = \left| \sqrt{p} W_{r,i}^T \alpha_{rx,i} \underline{a}_{rx,i}^I \otimes \left(\underline{a}_I^{rx,i}\right)^H diag(\Gamma) \alpha_{tx,i} \underline{a}_I^{tx,i} \otimes \left(\underline{a}_{tx,i}^I\right)^H X \right|^2 \quad (37)$$

where $\alpha_{rx,i}$, $\underline{a}_{rx,i}^I$, $\underline{a}_I^{rx,i}$, $\alpha_{tx,i}$, $\underline{a}_I^{tx,i}$ and $\underline{a}_{tx,i}^I$ coefficients correspond to α_{rx}, $\underline{a}_{rx}^I$, $\underline{a}_I^{rx}$, α_{tx}, $\underline{a}_I^{tx}$ and $\underline{a}_{tx}^I$ coefficients, respectively, except the receiver location is defined as i^{th} receiver location.

The above equation calculates the received power at the i^{th} receiver location, $p(s_{rx}^i, \theta_n, \varphi_n)$ when the IRS orientation is $(1, \theta_n, \varphi_n)$. The optimization problem to maximize the received power at each location can be expressed as:

$$\max_{\theta_n, \varphi_n} w_i \, R(s_{rx}^i, \theta_n, \varphi_n) \quad (38)$$

where w_i is the weight factor for the i^{th} receiver based on the user population density and the direct link received signal power of the i^{th} receiver location.

Let $\underline{w}$ be the normalized weight vector and $\underline{R} \in \mathbb{C}^{N_{rx} \times 1}$ be the achievable rate vector whose elements correspond to the w_i and $R(s_{rx}^i, \theta_n, \varphi_n)$, respectively. Then the IRS orientation optimization problem for multiple receiver locations can be written as:

$$\max_{\theta_n, \varphi_n} \underline{w}^T R \quad (39)$$

Similar to the single receiver location problem earlier, we may also consider the expected received power at receiver locations if the IRS were to be oriented in a random direction in between the transmitter location and any point inside the weak signal strength area. Let $R_{\Omega_i} = \mathrm{E}_\Omega\{R_i\}$ be

defined as the expected received power corresponds to the i-th receiver location when the IRS is oriented in a random direction in between the transmitter location and any point inside the weak signal strength area. Therefore, the IRS orientation optimization problem can be formulated to maximize the average achievable rate at locations inside the desired area, A_ϵ:

$$(\theta_n^*, \varphi_n^*) \approx \arg\max_{\theta_n, \varphi_n} \int_{A_\epsilon} w(s_{rx})\, R(s_{rx}, \theta_n, \varphi_n) \mathrm{d}s_{rx} \tag{40}$$

where $w(s_{rx})$ is the weight distribution over the receiver locations based on, for example, the user population density and the strength of the direct link.

We assume that within A_ϵ users are uniformly distributed over the three-dimensional space defined in a cylindrical coordinate system with $\rho_0 \leq \rho \leq \rho_m$, $\varphi_0 \leq \varphi_l^{rx} \leq \varphi_m$ and $z_0 \leq z \leq z_m$ are the cylindrical coordinates of receiver locations with respect to the IRS, ρ_0 and ρ_m are the minimum and maximum coverage radius of the A_ϵ area, φ_0 and φ_m are the range of φ_l^{rx} and z_0 and z_m are the minimum and maximum of the A_ϵ area height with respect to the IRS. For simplicity, let us assume that transmitter and receiver locations satisfy $\|\varphi_{tx} - \varphi_{rx}\| < \pi$, which can easily be generalized (at the expense of notational complexity). If we were to assume that the number of users with distance ρ is proportional to ρ and but independent of z and φ, the user distribution probability density function (PDF) can be derived to be:

$$w(\rho, \varphi, z) = \frac{2\rho}{(\rho_m^2 - \rho_0^2)(\varphi_m - \varphi_0)(z_m - z_0)}. \tag{41}$$

Since the second derivative of optimal solution equation is negative, the optimal IRS orientation can be obtained as the solution to:

$$\frac{d}{d\theta_n} \mathrm{E}_{\mathrm{P},\Phi,\mathrm{Z}}\{\log_2(SNR(\rho, \varphi, z, \theta_n, \varphi_n))\} = 0 \tag{42}$$

and

$$\frac{d}{d\varphi_n} \mathrm{E}_{\mathrm{P},\Phi,\mathrm{Z}}\{\log_2(SNR(\rho, \varphi, z, \theta_n, \varphi_n))\} = 0\,. \tag{43}$$

It can be shown that the solution to above equations are:

$$\theta_n^* = \frac{\mathrm{E}_{\mathrm{P},\Phi,\mathrm{Z}}\{\theta_{rx}\} + \theta_{tx}}{2} \tag{44}$$

and

$$\varphi_n^* = \frac{\varphi_m + \varphi_0 + 2\varphi_{tx}}{4} \tag{45}$$

where

$$\mathrm{E}_{\mathrm{P},\Phi,\mathrm{Z}}\{\theta_{rx}\} = \frac{\pi}{2} + \left(\rho_0^3 \ln\left(\frac{\rho_0^2 + z_0^2}{\rho_0^2 + z_m^2}\right) - \rho_m^3 \ln\left(\frac{\rho_m^2 + z_0^2}{\rho_m^2 + z_m^2}\right) - (\rho_m - \rho_o)(z_m^2 - z_0^2) + \right.$$

$$3z_0\rho_m^2 \tan^{-1}\left(\frac{z_0}{\rho_m}\right) - 3z_0\rho_0^2 \tan^{-1}\left(\frac{z_0}{\rho_0}\right) - 3z_m\rho_m^2 \tan^{-1}\left(\frac{z_m}{\rho_m}\right) + 3z_m\rho_0^2 \tan^{-1}\left(\frac{z_m}{\rho_0}\right) -$$

$$z_m^3\left(\tan^{-1}\left(\frac{z_m}{\rho_m}\right)-\tan^{-1}\left(\frac{z_m}{\rho_0}\right)\right)+z_m^3\left(\tan^{-1}\left(\frac{z_0}{\rho_m}\right)-\tan^{-1}\left(\frac{z_0}{\rho_0}\right)\right)\Big)/(3(\rho_m^2-\rho_o^2)(z_m-z_0)) \tag{46}$$

We may also consider the expected achievable rate at an arbitrary receiver location within A_ϵ if the IRS were to be oriented in a random direction in between the transmitter location and any point inside A_ϵ, which can be defined as $R_\Omega = \mathrm{E}_\Omega\{\mathrm{E}_{\mathrm{S_{rx}}}\{R\}\}$ where $\Omega(\theta_n, \varphi_n)$ is the set of IRS orientation angles considered for averaging:

$$R_\Omega = \iint_\Omega \frac{1}{\|\theta_{tx}-\theta_{rx}\|\|\varphi_{tx}-\varphi_{rx}\|}\iint_{\Omega_p} w(\rho,\varphi,z)R(s_{rx},\theta_n,\varphi_n)\,d\rho d\varphi dz \;\; d\theta_n d\varphi_n \tag{47}$$

This can be simplified as:

$$\mathrm{E}\{p_\Omega\} = \frac{\widetilde{B}}{\|\theta_{tx}-\theta_{rx}\|\|\varphi_{tx}-\varphi_{rx}\|}\int_{\theta_{tx}}^{\theta_{rx}}\int_{\varphi_{tx}}^{\varphi_{rx}} f\left(\theta^{tx}_{m^x m^y}, \varphi^{tx}_{m^x m^y}\right) f\left(\theta^{rx}_{m^x m^y}, \varphi^{rx}_{m^x m^y}\right) d\theta_n d\varphi_n \tag{48}$$

where

$$\widetilde{B} = \frac{p g_t G g_r \lambda^2 f^{tx}(\theta^I_{tx},\varphi^I_{tx})\, f^{rx}(\theta^I_{rx},\varphi^I_{rx})\, d_I^x d_I^y \|\Gamma\|^2}{64\pi^3 r_I^{t^2} r_I^{t^2}}. \tag{49}$$

3.4 3D Beamforming for mmWave UAV Communications

Figure 5 shows a communications link between two Unmanned Aerial Vehicles (UAVs). The UAV transmitter is equipped with M_{tx} number of transmit antenna elements where $M_{tx} = M_{tx}^x M_{tx}^y$ and the UAV receiver is equipped with M_{rx} number of receiver antenna elements where $M_{rx} = M_{rx}^x M_{rx}^y$.

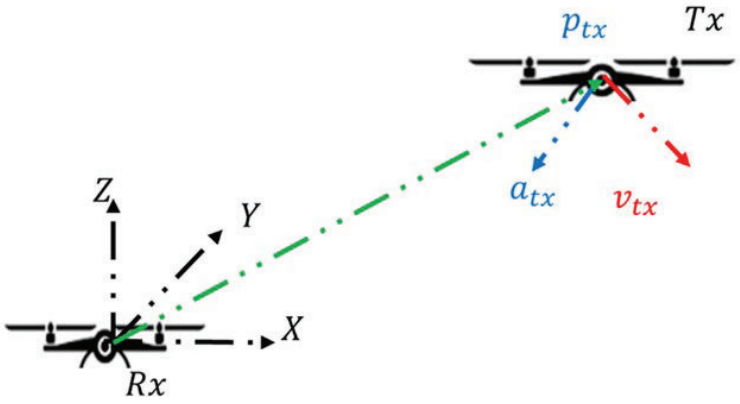


Figure 5. A Communications Link Between a UAV Transmitter and a UAV Receiver

The numbers of elements in the transmitter UAV antenna's horizontal and vertical directions are assumed to be M_{tx}^{x} and M_{tx}^{y}, respectively. Similarly, M_{rx}^{x} and M_{rx}^{y} are the numbers of elements in the receiver UAVs array in each direction. These receiver antenna elements are indexed as (ℓ^x, ℓ^y) for $0 < l^x < M_{rx}^{x} - 1$ and $0 < l^y < M_{rx}^{y} - 1$. Let $\underline{\ell}^x = [0 \quad \cdots \quad M_{rx}^{x} - 1]$ and $\underline{\ell}^y = [0 \quad \cdots \quad M_{rx}^{y} - 1]$ denote the vectors of indices of the receiver antenna elements in each direction. The received complex baseband signal $Y_r \in \mathbb{C}^{M_{rx} \times 1}$ can be expressed as:

$$Y_r = \sqrt{p} H_{tx}^{rx} X + Z \tag{50}$$

where $X \in \mathbb{C}^{M_{tx} \times 1}$ is the UAV transmit signal vector, $H_{tx}^{rx} \in \mathbb{C}^{M_{rx} \times M_{tx}}$ is the LOS channel coefficients matrix from the transmitter to the receiver, $Z \in \mathbb{C}^{M_{rx} \times 1}$ is the receiver noise vector, and p is the transmit symbol power.

The channel coefficient matrix H_{tx}^{rx} can be expressed as:

$$H_{tx}^{rx} = \alpha \underline{a}_{rx}^{tx} \left(\underline{a}_{tx}^{rx}\right)^H \tag{51}$$

where $\alpha \in \mathbb{C}^{1 \times 1}$ is the LOS path gain from the UAV transmitter antenna to the UAV receiver, $\underline{a}_{tx}^{rx} \in \mathbb{C}^{M_{tx} \times 1}$ is the array response vector at the UAV transmitter antenna based on the UAV receiver location, and $\underline{a}_{rx}^{tx} \in \mathbb{C}^{M_{rx} \times 1}$ is the array response vector at the UAV receiver based on received signal from the UAV transmitter.

The beamformed received signal at the UAV receiver can be expressed as:

$$\mathrm{y} = W_r^H Y_r \tag{52}$$

where $W_r \in \mathbb{C}^{M_{rx} \times 1}$ is the receiver beamforming vector. Substituting for the channel coefficient matrix gives:

$$\mathrm{y} = \alpha \sqrt{p} W_r^H \underline{a}_{rx}^{tx} \left(\underline{a}_{tx}^{rx}\right)^H X + W_r^H Z \tag{53}$$

where LOS path loss α can be represented as:

$$\alpha = \sqrt{\frac{g_t g_r \lambda^2 f^{rx}(\theta_{rx}^{tx}, \varphi_{rx}^{tx}) f^{tx}(\theta_{tx}^{rx}, \varphi_{tx}^{rx})}{16\pi^2 r^2}}. \tag{54}$$

We define the normalized receiver beamforming gain $\mathrm{G_{B,rx}}$ as:

$$\mathrm{G_{B,rx}} = f^{rx}(\theta_{rx}^{tx}, \varphi_{rx}^{tx}) \left\| \frac{1}{M_{rx}} W_r^H \underline{a}_{rx}^{tx} \right\|^2 \tag{55}$$

3.4.1 3D Beamforming Weight Vector Update Timing

By supporting a large number of elements in a smaller aperture, mmWave antennas can produce highly directional narrow beams. However, even a slight misalignment in such narrow beams could result in significant signal loss at the receiving UAV. As a result, due to the mobility of both receiving and transmitting UAVs, it is important to maintain the alignment of both beams in order to ensure a reliable connection. To accomplish this, beamforming weights have to be updated sufficiently fast in accordance with relative movement of the two UAVs.

Without loss of generality, let us consider the relative motion of a UAV transmitter with respect to the UAV receiver as illustrated in Figure 5 (if both are indeed mobile, one takes the relative orientation and the velocity). Let $\theta_{rx}^{tx}(t), \varphi_{rx}^{tx}(t)$, $\theta_{tx}^{rx}(t)$ and $\varphi_{tx}^{rx}(t)$ be the zenith and azimuth angles of the transmitter from the receiver and the zenith and azimuth angles of the receiver from the transmitter at time *t*, respectively, based on coordinate systems placed normal to the receiver or transmitter UAV planner antenna array. Similarly, let $\theta_{rx}^{tx}(t+\Delta t)$, $\varphi_{rx}^{tx}(t+\Delta t)$, $\theta_{tx}^{rx}(t+\Delta t)$ and $\varphi_{tx}^{rx}(t+\Delta t)$ be the corresponding angles of the transmitter and receiver locations at time $t+\Delta t$. When the receiver beamforming is performed based on the transmitter location at time *t*, W_r can be expressed as:

$$W_r = e(\theta_{rx}^{tx}(t), \varphi_{rx}^{tx}(t), M_{rx}^x, M_{rx}^y, d_{rx}^x, d_{rx}^y) \tag{56}$$

However, the instantaneous $\underline{a}_{rx}^{tx}$ at $t+\Delta t$ depends on the time $t+\Delta t$:

$$\underline{a}_{rx}^{tx} = e(\theta_{rx}^{tx}(t+\Delta t), \varphi_{rx}^{tx}(t+\Delta t), M_{rx}^x, M_{rx}^y, d_{rx}^x, d_{rx}^y) \tag{57}$$

so that

$$\mathrm{G_{B,rx}} = f^{rx}(\theta_{rx}^{tx}, \varphi_{rx}^{tx}) \left\| \frac{1}{M_{rx}} e\left(\theta_{rx}^{tx}(t), \varphi_{rx}^{tx}(t), M_{rx}^x, M_{rx}^y, d_{rx}^x, d_{rx}^y\right)^H e\left(\theta_{rx}^{tx}(t+\Delta t), \varphi_{rx}^{tx}(t+\Delta t), M_{rx}^x, M_{rx}^y, d_{rx}^x, d_{rx}^y\right) \right\|^2. \tag{58}$$

Suppose that the minimum receiver beamforming gain required at the receiver is $G_{\mathrm{B,rx,min}}$. Then, the maximum tolerable update period Δt can be obtained based on the receiver gain as:

$$\max_{s.t\ \mathrm{G_{B,rx}} > G_{\mathrm{B,rx,min}}} \Delta t. \tag{59}$$

Similarly, we define the normalized transmitter beamforming gain $\mathrm{G_{B,tx}}$ as:

$$\mathrm{G_{B,tx}} = f^{tx}(\theta_{tx}^{rx}, \varphi_{tx}^{rx}) \left\| \frac{1}{M_{tx}} W_t^H \underline{a}_{tx}^{rx} \right\|^2 \tag{60}$$

When the transmitter beamforming is performed based on the receiver location at time *t*, W_t can be expressed as:

$$W_t = e(\theta_{tx}^{rx}(t), \varphi_{tx}^{rx}(t), M_{tx}^x, M_{tx}^y, d_{tx}^x, d_{tx}^y) \tag{61}$$

As before, the instantaneous $\underline{a}_{rx}^{tx}$ at $t+\Delta t$ depends on the time $t+\Delta t$:

$$\underline{a}_{tx}^{rx} = e(\theta_{tx}^{rx}(t+\Delta t), \varphi_{tx}^{rx}(t+\Delta t), M_{tx}^x, M_{tx}^y, d_{tx}^x, d_{tx}^y) \tag{62}$$

so that

$$\mathrm{G_{B,tx}} = f^{tx}\left(\theta_{tx}^{rx}(t+\Delta t), \varphi_{tx}^{rx}(t+\Delta t)\right) \left\| \begin{matrix} \frac{1}{M_{tx}} e\left(\theta_{tx}^{rx}(t), \varphi_{tx}^{rx}(t), M_{tx}^x, M_{tx}^y, d_{tx}^x, d_{tx}^y\right)^H \\ e\left(\theta_{tx}^{rx}(t+\Delta t), \varphi_{tx}^{rx}(t+\Delta t), M_{tx}^x, M_{tx}^y, d_{tx}^x, d_{tx}^y\right) \end{matrix} \right\|^2. \tag{63}$$

Suppose that the minimum beamforming gain required at the transmitter is $G_{\mathrm{B,tx,min}}$. Then, the maximum tolerable update period Δt can be obtained based on the transmitter gain as:

$$\max_{s.t\ \mathrm{G_{B,tx}} > G_{\mathrm{B,tx,min}}} \Delta t\ . \tag{64}$$

The transmitter beamforming gain can be simplified as:

$$\mathrm{G_{B,tx}} = f^{tx}\left(\theta_{tx}^{rx}(t+\Delta t), \varphi_{tx}^{rx}(t+\Delta t)\right)\left\| \frac{\sin\left(\frac{M_{tx}^{x}\phi_x}{2}\right)\sin\left(\frac{M_{tx}^{y}\phi_y}{2}\right)}{M_{tx}^{x}\sin\left(\frac{\phi_x}{2}\right)M_{tx}^{y}\sin\left(\frac{\phi_y}{2}\right)} \right\|^2 \tag{65}$$

where

$$\phi_x = \frac{2\pi d_{tx}^{x}}{\lambda}\left(\sin\left(\theta_{tx}^{rx}(t)+\Delta\theta_{tx}^{rx}(t)\right)\cos\left(\varphi_{tx}^{rx}(t)+\Delta\varphi_{tx}^{rx}(t)\right) - \sin\left(\theta_{tx}^{rx}(t)\right)\cos\left(\varphi_{tx}^{rx}(t)\right)\right) \tag{66}$$

$$\phi_y = \frac{2\pi d_{tx}^{y}}{\lambda}\left(\sin\left(\theta_{tx}^{rx}(t)+\Delta\theta_{tx}^{rx}(t)\right)\sin\left(\varphi_{tx}^{rx}(t)+\Delta\varphi_{tx}^{rx}(t)\right) - \sin\left(\theta_{tx}^{rx}(t)\right)\sin\left(\varphi_{tx}^{rx}(t)\right)\right) \tag{67}$$

$$\Delta\theta_{tx}^{rx}(t) = \int_{t}^{t+\Delta \mathrm{t}} \dot{\theta}_{tx}^{rx}(\tau)\, d\tau \tag{68}$$

and

$$\Delta\varphi_{tx}^{rx}(t) = \int_{t}^{t+\Delta \mathrm{t}} \dot{\varphi}_{tx}^{rx}(\tau)\, d\tau. \tag{69}$$

Based on the relative angular velocities $\dot{\theta}_{tx}^{rx}(t)$ and $\dot{\varphi}_{tx}^{rx}(t)$ of receive UAV with respect to the transmit UAV, the optimal Δt denoted as Δt^* to maintain a minimum required beamforming gain, $G_{\mathrm{B,tx,min}}$ can be determined. However, due to the form of $\mathrm{G_{B,tx}}$, in general the solution for Δt^* can not be expressed in closed-form but can only be determined numerically. Moreover, calculation of Δt^*requires the knowledge of the receive UAV trajectory from time t to $t+\Delta t$, requiring the complete knowledge of the relative flight paths of the UAVs.

As a special case, let us assume that the receiver UAV is moving only in the elevation plane (i.e., $\dot{\varphi}_{tx}^{rx}(t) = 0$). In addition, we will assume that $\varphi_{tx}^{rx}(t) = 0$ and constant relative angular velocity of receive UAV with respect to transmit UAV is ω_{rx}. Let Δt_{el}^* be the Δt^*corresponds to this scenario. In this case, the Δt_{el}^* to maintain a required minimum transmit beamforming gain of $G_{\mathrm{B,tx,min}}$ can be expressed in closed-form as:

$$\Delta t_{el}^* = \frac{\sin^{-1}\left(\frac{\lambda sinc^{-1}\left(\sqrt{G_{\mathrm{B,tx,min}}}\right)}{M_{tx}^{x}\pi d_{tx}^{y}} + sin(\theta_{tx}^{rx}(t))\right) - \theta_{tx}^{rx}(t)}{\omega_{rx}} \tag{70}$$

In general, however, the relative UAV movement may not be confined to only the elevation plane in a real UAV to UAV (U2U) communication system. When that is the case, even if the complete relative flight path of the receive UAV is available, calculating Δt^* numerically can still be computationally demanding.

In some situations, receivers are required to maintain a minimum received signal power to operate. Assume that the required minimum received power is p_r^{min} and the transmit UAV power is P_{tx} . The free space path loss can be expressed as:

$$\frac{p_r^{min}}{P_{tx}} = G_r G_t G_{B,tx} \frac{\lambda^2}{16\pi^2 r^2}. \tag{71}$$

For a 2-D antenna array, $G_t = 4\pi A_{e,t}/\lambda^2$ where $A_{e,t}$ is the effective aperture of transmit antenna, so that

$$\frac{p_r^{min}}{P_{tx}} = G_r G_{B,tx} \frac{A_{e,t}}{4\pi r^2} \tag{72}$$

and

$$G_{B,tx} = \frac{p_r^{min} 4\pi r^2}{P_{tx} G_r A_{e,t}}. \tag{73}$$

From this the threshold value of $G_{B,tx}$ can be determined based on the UAV receiver location. Then, the maximum allowable Δt^* can be determined to maintain a required minimum received power at the UAV receiver.

3.4.2 Simplified Solution for 2D and 3D Beamforming Weights Update Period

The formulation in the preceding section requires the knowledge of the relative trajectories of the transmitter and receiver from time *t* to $t + \Delta t$ to determine the Δt^* at time *t*, which may not be realistic in some scenarios. However, since the AOA of signals can be determined with high accuracy, in the following we derive an approximation for beamforming weight vector update period that can be calculated based on the present estimates of the receiving UAV's location and speed [20].

Let us assume that the transmit UAV antenna is directed at the receiver UAV in bore-sight direction as shown in Figure 6. We assume that the transmit beamforming weight vector needs to be changed before the receiving UAV moves outside of half-power beamwidth (HPBW) region of the transmit antenna array. Figure 7 shows the transmitter beam pattern when the antenna is pointed in $\theta_{tx}^{rx}(t) = 0$ direction.

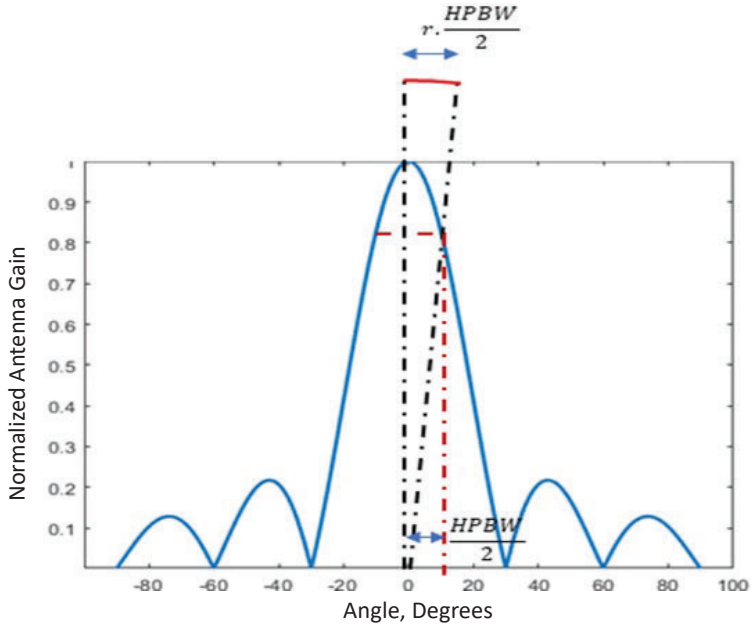


Figure 6. The UAV Transmitter Antenna Beam Pattern and the HPBW

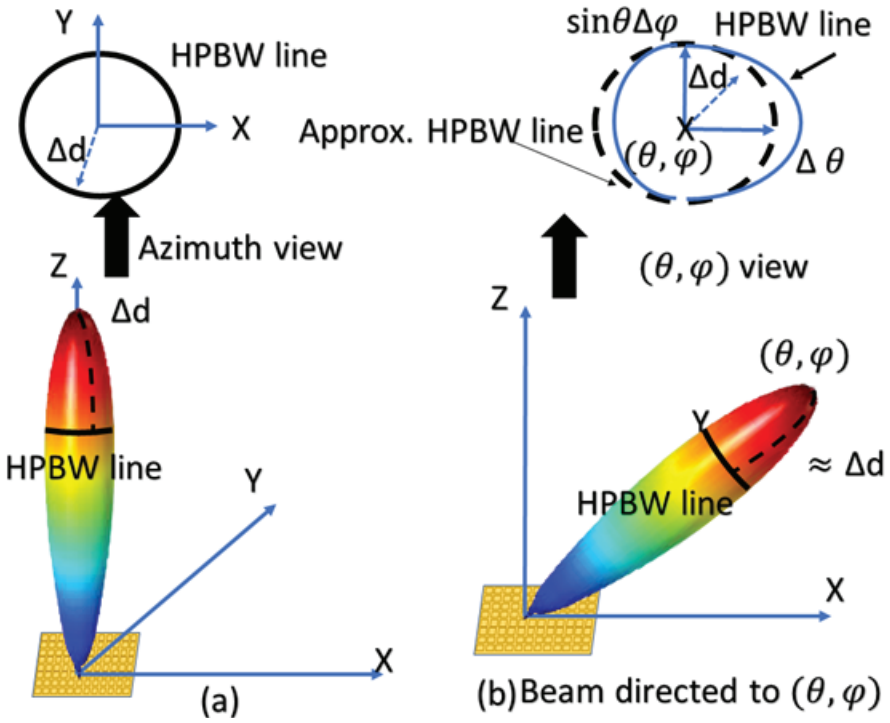


Figure 7. Comparison of Δd When the Beam is Directed to (θ, φ)

For a square planner array, in this case the boundary of the half-power beamwidth in 3D signal beam will be a circle around the beam-axis (which coincides with the *z*-axis), as shown in Figure 7(a). The distance Δd to move beyond the HPBW line in this case can be expressed as:

$$\Delta d = \frac{r.HPBW}{2}. \tag{74}$$

Thus, an approximate update period can be derived as:

$$\Delta t_{approx} = \frac{r.HPBW}{2v_t} \tag{75}$$

where $\boldsymbol{v_t}$ is the relative receive UAV velocity component in tangential direction at time $\boldsymbol{t}$. It can be seen that:

$$v_t = r.\sqrt{\dot{\theta}_{tx}^{rx}(t)^2 + \dot{\varphi}_{tx}^{rx}(t)^2 \sin^2 \theta_{tx}^{rx}(t)}. \tag{76}$$

When signal beam is directed to a desired (θ, φ) direction by using 3D beamforming, the shape of the boundary line of the HPBW is altered as shown in Figure 7(b). If we were to neglect this effect, we may assume that the distance Δd required to move beyond the HPBW line is approximately equal in both scenarios so that

$$\frac{r.HPBW}{2v} = r.\Delta t_{approx}\sqrt{\dot{\theta}_{tx}^{rx}(t)^2 + \sin^2 \theta_{tx}^{rx}(t)\, \dot{\varphi}_{tx}^{rx}(t)^2}. \tag{77}$$

The update period of 3D beamforming weights, Δt, can thus be simplified to:

$$\Delta t_{approx} = \frac{HPBW}{2\sqrt{\dot{\theta}_{tx}^{rx}(t)^2 + \sin^2 \theta_{tx}^{rx}(t)\dot{\varphi}_{tx}^{rx}(t)^2}} \tag{78}$$

or

$$\Delta t_{approx} = \frac{r.HPBW}{2\sqrt{v_\theta(t)^2 + \sin^2 \theta_{tx}^{rx}(t)v_\varphi(t)^2}} \tag{79}$$

where $v_\theta(t)$ and $v_\varphi(t)$ are the linear velocities corresponding to $\dot{\theta}_{tx}^{rx}(t)$ and $\dot{\varphi}_{tx}^{rx}(t)$, respectively.

3.5 Drone Trajectory Planning with Deep Reinforcement Learning and AI

With the deployment of 5G cellular systems that use millimeter waves and consequently have smaller cell sizes, UAV mounted base stations are a viable alternative to fixed base stations. However, owing to their limited radius of coverage and battery constraints, it is not realistic for UAVs to hover at a single location and cover all users. Ideally, the UAV must traverse a trajectory covering as many users as possible. For a general application scenario, it may not be able to find such a trajectory that satisfies a specified optimality criterion. We therefore propose two AI based approaches: (1) Deep Reinforcement Learning (DRL) implemented with a Deep Q-Network (DQN), and (2) a reward-based "greedy" algorithm.

As in [21], a square geographical area of length L is divided into M number of square cells, each of identical size, as shown in Figure 8. An arbitrary, and unknown, number of users are assumed to be distributed across this area. The UAV coverage radius is represented by the green circle. The blue dots indicate the users. The 41 red crosses represent the possible hovering points. We assume that the UAV maintains a fixed altitude precluding the actions of moving up or down. In many situations, this may make sense to provide uniform coverage while simplifying interference management among multiple UAVs and subscribers. For simplicity, we restrict the UAV to hover only in the center and corners of the cells, so that there are $2\sqrt{M}\left(\sqrt{M}+1\right)+1$ possible hovering points. If we assume that the coverage radius of the UAV base station is $l/2$, where l is the side-length of a square cell, then the UAV base station will be able to provide coverage to any user from one of the possible hovering points.

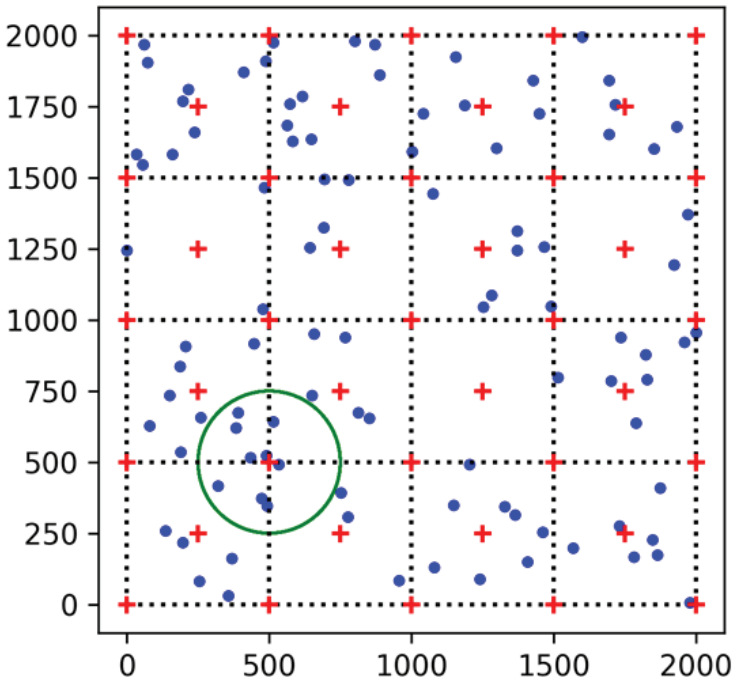


Figure 8. Illustration of UAV Hovering Geography with Assumed Location at (500, 500)

We use three approaches to model how users are spread out in the area of interest. First, we assume that the users are distributed uniformly and remain static throughout. Second, we assume that the users are clustered around certain fixed hotspots. Third, we assume that the users are clustered and the cluster means are time-varying. Note that Figure 8 is an example case with 100 users that are distributed uniformly over a square area of 2000-m × 2000-m. However, the latter models would make sense, since users are more likely to be around places like residential areas, offices and universities. We model these clusters using the Gaussian Mixture Model (GMM). Each cluster in the mixture is parameterized by a cluster weight π_k, cluster mean μ_k and variance σ_k^2. The cluster

weights represent the probability of users being in each cluster and satisfy the following normalization condition:

$$\sum_{k=1}^{K} \pi_k = 1 \tag{80}$$

where K is the number of clusters. The cluster means $\mu_k = [\mu_x, \mu_y]$ represent the mean X and Y coordinates of each cluster while variances $\sigma_k^2 = [\sigma_x^2, \sigma_y^2]$ represent the spread along X and coordinates. Figure 9 shows an example in which the users are distributed according to a GMM with four clusters with an equal variance of 100 along each coordinate.

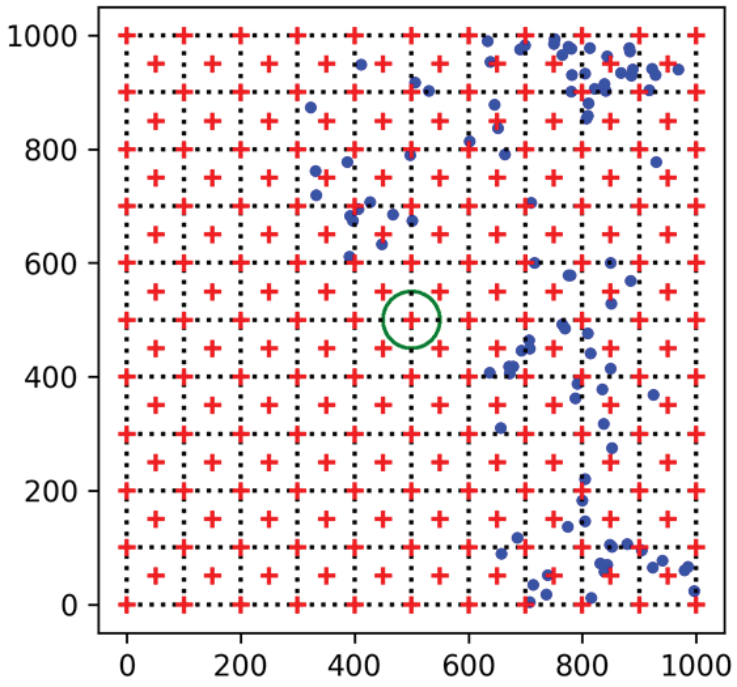


Figure 9. A GMM User Distribution with Four Clusters

There could also be scenarios in which users are densely clustered around hotspots. We may capture this situation by using GMM with clusters having smaller variances. Figure 10 shows the distribution of users around eight hotspots.

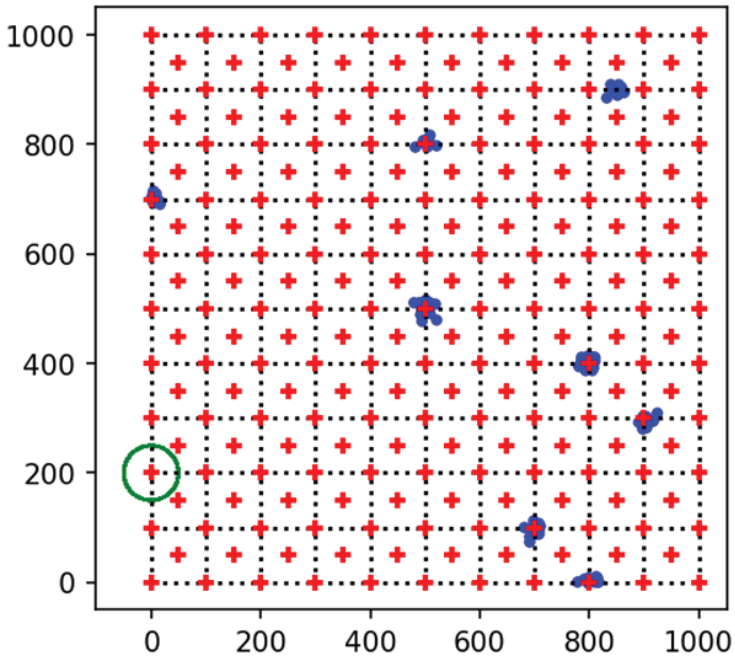


Figure 10. A GMM User Distribution with Eight Clusters and High User Density

However, the GMM with fixed clusters become unsuitable when ground users correspond to people commuting to work or school from their homes or troops on the move. To capture mobility of users in such scenarios, in our third approach we allow the clusters to move in predetermined directions at predetermined speeds. This model can later be generalized to capture more complex mobility patterns.

3.5.1 Proposed Methods 1: Deep Reinforcement Learning using DQN

A reinforcement learning algorithm consists of an agent in an environment that takes an action a_t at each time step t depending on the state of the environment at that time instance represented by S_t. The state S_t is the input to the agent and it describes the environment. Based on the action, for each time step, the agent receives a scalar reward r_t. A favorable action yields a higher reward. The objective of the agent is to take actions that maximizes the sum of discounted future rewards.

One way to find such good actions is to use machine learning. According to Q-learning, for example, such an action strategy can be learnt by using a look up table called the Q-Table. The rows of the Q-Table represent the states and the columns represent the actions. The number of columns is equal to the number of actions. However, when the possible space of state-action pairs is large, the Q-Table will be too large requiring a larger memory footprint. To mitigate this, neural networks (NNs) can be used to approximate the Q-table. However, until recently this approach has had only limited success due to convergence issues. A breakthrough was achieved by DeepMind in which a deep Q-network (DQN) was used as the function estimator [22]. By utilizing an

experience replay memory and a target network, the DQN approach has been shown to successfully learn a Q-function approximation in many reinforcement learning applications [23, 24].

Since the state is the input to an NN, the state must be chosen carefully to encapsulate the information about the environment. We propose two candidates: The first state, shown in Figure 11, consists of a matrix whose elements consist of number of users covered at each hovering point along with the location of the UAV and the energy remaining in the UAV. The second candidate for state, shown in Figure 12, has the UAV's location, number of users covered at that time instance and the energy remaining along with the same information from nine previous time steps. The former was used as the state when the user distribution was random and static whereas the latter was used when the user distribution was modelled as a Gaussian mixture.

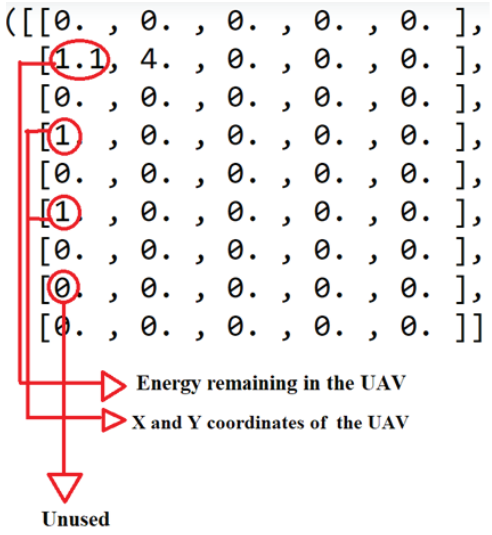


Figure 11. A State That Consists of the UAV's Location, the Remaining Energy and the Number of Users Covered at Each Hovering Point

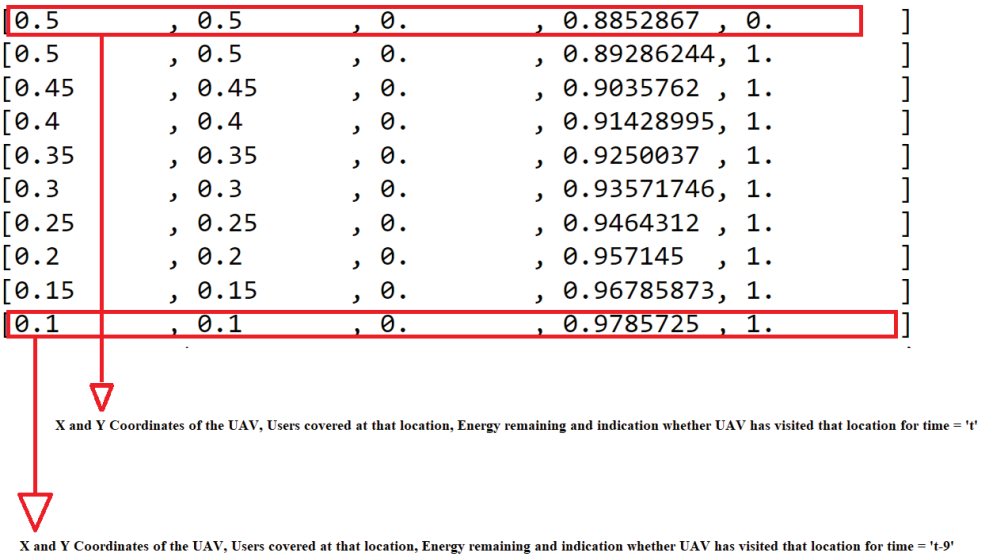


Figure 12. A State That Consists of the UAV's Location, the Energy Remaining in the UAV, the Number of Users Covered for Time Instance '*t*' Along with Same Information for Nine Previous Time Steps

We propose two candidates for the reward function. The first reward function, which was used for the case with uniform user distribution, is given below [25]:

$$r_t = p_t + \begin{cases} \sum_{n=1}^{N} Cov_n^t \times (F_t - F_{t-1}), & \text{if } F_{t-1} \neq 1 \\ \sum_{n=1}^{N} Cov_n^t \times F_t, & \text{otherwise} \end{cases} \tag{81}$$

where $Cov_n^t \in \{0,1\}$ for $n \in \{1, \ldots . N\}$ and F_t is the Jain's fairness index [26] given by:

$$F_t = \frac{(\sum_{n=1}^{N} \sum_{i=1}^{t} Cov_n^i)^2}{N(\sum_{n=1}^{N} (\sum_{i=1}^{t} Cov_n^i)^2)} \tag{82}$$

and p_t is a penalty that is added to discourage the agent from taking an action that results in the UAV flying out of the square area.

The second reward function that was developed during this project is used when the user distribution was modelled as a GMM:

$$r_t = p_t + (|N_t| - |N_{t-1}|) + g * V + F_t \tag{83}$$

where N_t denotes the set of indices of the users that have been covered up to time t so that $(|N_t| - |N_{t-1}|)$ is the number of new users covered during the t-th time step, $V \in \{0,1\}$ denotes whether the hovering point at (x_t, y_t) was visited previously or not and $g \in \mathbb{R}$ is a scaling factor. The reinforcement learning algorithm for updating the neural network used to approximate the Q-function is shown in Figure 13 [22].

Algorithm 1 DQN Algorithm

Initialize policy network $Q(s, a)$ with random weights
Set target network $Q'(s, a)$ with the same weights
Set $\varepsilon = 1$
Create a replay memory with
for episode = 1 to M **do**
    for time t = 1 to N **do**
        Generate a random number w between 0 and 1
        if $\varepsilon > w$ **then**
            Take a random action a_t
        else
            Take action: $a_t = \text{argmax}_a\, Q(S_t, a)$
        Store transition: $(S_t\ ,\ a_t\ ,\ r_t\ ,\ S_{t+1})$ into buffer
        Sample random batch of transitions from the buffer
        if Episode ends at $t + 1$ **then**
            Set $y_t = r_t$
        else
            Set $y_t = r_t + \gamma \max_{a'} Q'(S_{t+1}, a')$
        Fit the policy network on the data: (S_t, y_t).
        Reduce the value of ε
        Every K steps set weights of target equal to policy network.

Figure 13. The DQN Algorithm

3.5.2 Proposed Method 2: Reward Based Greedy Algorithm

The second approach for action selection uses a reward-based greedy algorithm: at each time step, the UAV takes an action that yields the maximum instantaneous reward. This greedy action selection does not consider whether the selected action can have negative consequences on the long run. However, compared to DQN based ML approach, the greedy algorithm maybe implemented with much lower computational complexity. Future research will focus on further development and implementation of the greedy algorithm.

3.6 Security and Robustness of Machine Learning Based Trajectory Planning Systems

Although reinforcement learning is an attractive approach for solving optimization problems such as resource allocation and trajectory planning for which analytical solutions are hard to obtain, these systems can be vulnerable to adversarial attacks. In particular, when the inputs to these algorithms are modified either by accident or with a malicious intent, the performance could be degraded significantly. It is of interest to investigate such vulnerabilities and develop mitigation methods to counteract them.

In this section, we assess the robustness of reinforcement learning (RL) based resource allocation algorithms to a particular class of adversarial attack known as the Sybil attack. Sybil attack aims to undermine the fairness of a system by creating large pseudonymous identities to gain a disproportionate advantage. We consider a similar setup as in the previous section with addition of data requirements for ground users in addition to just coverage. We also impose stricter conditions on the number of users the UAV Base Station (UAV-BS) can simultaneously cover and introduce path-loss models to account for losses introduced in the wireless medium.

Consider a finite ground area with an unknown but fixed number of distributed users. A UAV-BS is assumed to fly at a fixed altitude and can take nine possible actions (i.e., moving in directions S, N, E, W, NE, NW, SE, SW or remain stationary). Once a direction is selected, the distance the UAV travels in that direction is assumed to be fixed at each time step. The objective of the UAV-BS is to discover as many ground users as possible and relay their data. The maximum number of users the UAV-BS can handle simultaneously is assumed to be fixed. The maximum number of users it can simultaneously handle is assumed to be fixed. At any given location, if the number of users exceeds this limit, the user requests are held in a first-in-first-out (FIFO) queue. To discover the presence of subscribers, the UAV-BS transmits a beacon signal requesting users in the vicinity to announce their presence with a response signal which consists of a user ID, and optionally the type of service they are requesting (i.e., data and voice). At any location the UAV will spend a maximum of T_w time units until hearing a response from at least one ground user. If at least one ground user does respond within T_w time limit, the UAV stays at that location for a maximum of T_s time units, where $T_s \geq T_w$, providing coverage to users within its footprint on the ground. This handshake mechanism between the UAV-BS and user is illustrated in Figure 14.

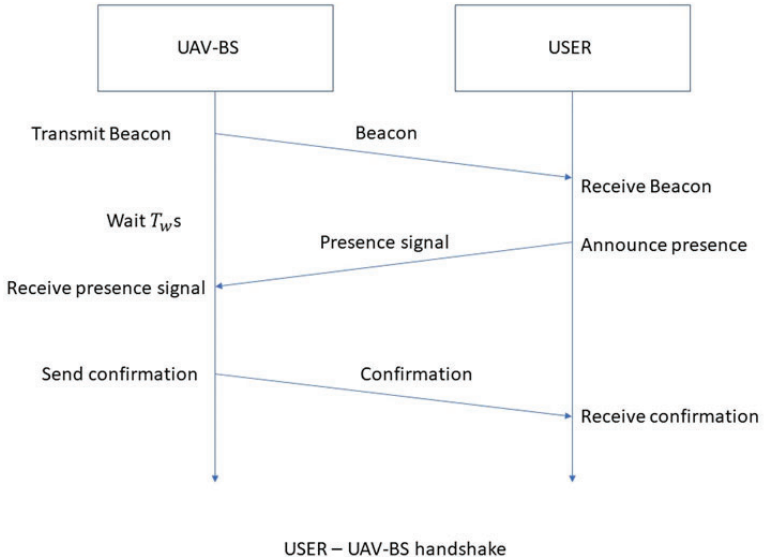


Figure 14. Detection of Ground Users by the UAV-BS

To ensure fairness, even if the first ground user responds at time $T_w - 1$, the UAV will stay at that location only for $T_s - T_w + 1$ time units. We model the response time of ground users as an exponential distribution with mean λ. Note that the exponential distribution has been shown to be a good probabilistic model to estimate packet arrival times [27].

The high throughput, directivity and low latency properties of the mmWave frequency band makes it an excellent choice for communications between the UAV-BS and the ground users. For example, a 5G link at 28 GHz has been shown to achieve a data rate of 1 Gbps and sub-millisecond (ms) latency [28]. Moreover, the mmWave spectrum may be better suited for UAV communications due to the ability to support LOS links with narrow signal beams. We assume a simplified path loss model for communication between the UAV-BS and the ground users which has been used for mmWave frequency bands at 28 GHz and 73 GHz:

$$P_r = P_t K \left(\frac{d_0}{d}\right)^{\gamma} \tag{84}$$

where P_r and P_t are received and transmit power by the UAV and user equipment, respectively, K is a unitless constant that depends on the antenna characteristics and the average channel attenuation, d_0 is a reference distance for antenna far field, and γ is the path loss exponent [29].

The required SNR for the user equipment to operate is modelled by assuming noise to be Gaussian with unit variance so that the SNR can be written as:

$$SNR\ [dB] = 10 \log_{10} P_r \tag{85}$$

For simplicity, all ground users are assumed to use similar equipment for transmitting and receiving data, so that transmit powers of all devices can be assumed to be the same. With this and the free space path loss model given by the equation above, the UAV is able to calculate the approximate distance between itself and the ground user. We assume that the total system bandwidth is divided into a set of frequency channels so that each user gets a fixed bandwidth. This imposes a limit on the maximum number of users who can be serviced simultaneously by the UAV-BS. The users are serviced based on their response times so that users who responded first to the beacon signal are given priority. There is a maximum limit on the amount of data each user can transmit to prevent misuse and cause unintentional denial of service (DoS) attacks similar to how Internet Service Providers (ISPs) throttle bandwidth after exceeding a threshold [30]. Users whose responses were late, users who still have data to transmit and users who have exhausted their maximum bandwidth are all placed in a queue. Requests from users who have exhausted the bandwidth, are entertained only if the UAV-BS is operating below its maximum capacity. The battery decay of the UAV-BS is assumed to be proportional to the flying time since more power is required to keep the UAV in the air [31]. The amount of data each user generates may be modelled as a gamma distribution or as a Generalized Extreme Value (GEV) distribution whose PDF is [32]:

$$f(x; \mu, \sigma, \xi) = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \tag{86}$$

where $t(x)$ is given by

$$t(x) = \begin{cases} (1 + \xi\left(\frac{x-\mu}{\sigma}\right))^{-\frac{1}{\xi}}, & \text{if } \xi \neq 0 \\ e^{-\frac{(x-\mu)}{\mu}}, & \text{if } \xi = 0 \end{cases} \quad (87)$$

The UAV-BS utilizes the user density learned based on the amount of received responses and the bandwidth demand at locations to determine a trajectory that maximizes distinct user coverage using a DQN-implemented reinforcement learning algorithm. The objective of the adversaries in this system is to disrupt and degrade the coverage provided by the UAV-BS. To accomplish this, the adversaries create groups of fake users and high bandwidth demands from these nonexistent users to trick the learning algorithm into wasting the UAV-BS's energy on these counterfeits. The attackers are assumed to be capable of transmitting data at a range of power levels while the UAV-BS assumes ground users transmit signals at a fixed power. This allows the adversaries to spoof their locations by simply altering their transmit power. During the training phase of determining an optimal trajectory, the UAV-BS is unaware of the locations of the ground users. The adversaries exploit this by placing fake users at strategic locations. At the beginning of each training epoch, adversaries place fake users and observe the flight path of the UAV-BS and then use this information in turn to place fake users at locations that are more likely to be visited by the UAV in future, rendering the learnt trajectories by the UAV-BS ineffective. It should be noted that each adversary decides the position of fake users (which it generates) independently of others.

The UAV uses a DQN-implemented RL to determine an optimal strategy that discovers as many ground users as possible while simultaneously maximizing the total amount of data relayed. The state input to the DQN at time instant t is made of the coordinates of the UAV locations (X_t, Y_t), the number of users detected at each location (N_t), the battery life of the UAV (B_t), a value that indicates whether that location was visited previously (V_t) and the number of users waiting in queue at that location (Q_t) along with the same information for a fixed number of previous time instances as shown in Figure 15.

$$\begin{pmatrix} X_t & Y_t & B_t & N_t & V_t & Q_t \\ X_{t-1} & Y_{t-1} & B_{t-1} & N_{t-1} & V_{t-1} & Q_{t-1} \\ & \cdot & \cdot & \cdot & & \\ X_{t-m} & Y_{t-m} & B_{t-m} & N_{t-m} & V_{t-m} & Q_{t-m} \end{pmatrix}$$

Figure 15. The State of the UAV-BS at *t*-th Time Step

At any given instance t, the UAV-BS can choose one of the nine possible actions in the action space. The actions available to the UAV-BS correspond to staying at the same location and the fixed directions in which the UAV can move: {N, S, E, W, N E, N W, SE, SW, be stationary}. The reward function for time step$t$ for the UAV which quantifies the quality of the selected action is proposed to be:

$$r_t = \begin{cases} (|N_t| - |N_{t-1}| + 1)\frac{D_t}{DN_t}, & \text{if } N_t \neq 0 \\ \eta V + J_t & \text{otherwise} \end{cases} \quad (88)$$

where $|N_t|$ is the set of users covered until the t-th time instant. Thus $|N_t| - |N_{t-1}|$ is the number of new users who had not been previously covered, D_t is the amount of data handled at the t-th instant for all users, D is the maximum allocated bandwidth limit for each user, η is a scaling factor, $V \in \{0,1\}$ denotes whether the location was visited previously, and J_t is a penalty for moving out of the coverage area of interest.

The objective of the adversaries is to reduce the number of genuine users discovered and serviced by the UAV-BS. To accomplish this, each adversary creates F_{max} number of fake users with spoofed locations which in turn is achieved by changing the transmit power appropriately. The spawning radius of an adversary is defined to be the maximum distance within which fake users can be placed. To define the action set of an adversary, the circular area determined by the spawning radius is divided into square grids as shown in Figure 16. At each time instant, the adversary's decision problem is to determine the grid cell in which to place the fake users. Once the grid cell is selected, the fake users within it are to be distributed uniformly. Thus, the action space for the adversary is the possible locations where the fake users can be placed and is equal to the number of grid cells within its spawning radius. Note that, the actions are scalars meaning that each value denotes a corresponding grid cell. In Figure 16, the blue dots indicate genuine users, the four black hexagons represent the four adversaries, the four orange circles, each of radius 250 units, correspond to the spawning radius of each adversary, the green squares denote the grid cells and the red stars represent the created fake users. In this case, each adversary can choose one of the 25 possible grid cells in which to place fake users as its action. It is assumed that the position of an adversary is fixed and that it always spawns F_{max} number of fake users. Note that, an adversary may estimate the coordinates of the UAV by combining the *a priori* information of its altitude with an estimate of signal angle-of-arrival (AOA). When an adversary is too far from the UAV-BS, we assume that it may receive this information from other adversaries via information sharing.

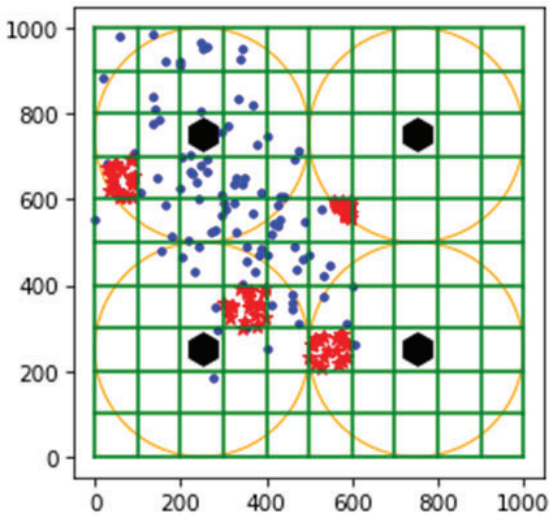


Figure 16. A Square Area with Genuine and Fake Users

Figure 17 shows the proposed state of an adversary at epoch e for the k-th adversary. It consists of information about the UAV-BS trajectory and the location of fake users during epoch $e-1$. The first column contains the location centered around which the fake users were placed (X_{e-1}^k, Y_{e-1}^k) and the corresponding grid cell index G_{e-1}^k. In the example scenario of Figure 16, $G_{e-1}^k \in \{0, \dots, 24\}$. Columns 2 to $t+2$ consists of the locations visited by the UAV-BS (X_i, Y_i) and the time spent at each of those locations (τ_i) where $i \in \{0, \dots, t\}$ if the epoch ended after t time steps. Note that, since energy consumed by the UAV-BS is assumed to be proportional to the flying time and the UAV-BS may spend anywhere between T_s and T_w time units at any given location, the number of time steps in an epoch is not fixed. However, since the UAV-BS spends at least T_w time units at any given location, the maximum number of time steps in an epoch is fixed.

$$\begin{pmatrix} X_{e-1}^k & X_0 & .. & X_{t-1} & X_t \\ Y_{e-1}^k & Y_0 & .. & Y_{t-1} & Y_t \\ G_{e-1}^k & \tau_0 & .. & \tau_{t-1} & \tau_t \end{pmatrix}$$

Figure 17. The State Matrix of the *k*-th Adversary at Epoch *e*

Each adversary is controlled by an independent DQN and the decision of each adversary is independent of the other. The state input to the DQN of adversary k at epoch e consists of the location and the corresponding grid $(X_{e-1}^k, Y_{e-1}^k, G_{e-1}^k)$ where fake users were placed and the trajectory of the UAV-BS $(X_0, Y_0, D_0)\ to\ (X_t, Y_t, D_t)$ along with the time spent at each location during epoch $e-1$ as shown in Figure 17.

The reward function of the adversaries is made to be proportional to the number of fake users covered. Hence, the reward at epoch e of k-th adversary is proposed to be:

$$r_e^k = -\sum_{n=1}^{F_{max}} F_n^k \tag{89}$$

where $F_n^k \in \{0,1\}$ denotes whether or not the n-th fake user generated by the k-th adversary was serviced by the UAV-BS or not.

4 RESULTS AND DISCUSSION

4.1 IRS-Aided 3D Beamforming for mmWave Communications

In this section, we provide numerical results that shows the importance of optimizing the IRS orientation. To highlight the variation in the received power with respect to the IRS orientation, IRS beamforming and transmitter/receiver beamforming are performed at the respective ends for all case studies.

First, consider a simple case study with an IRS placed near the receiver. The IRS consists of a total number of elements 250×250 whose element size in each direction is λ (i.e, $d_I^x = \lambda$ and $d_I^y = \lambda$). The IRS is placed at a position of r_I^r = 100-m, $\theta_{rx}^I = 60°, \varphi_{rx}^I = 140°$ and the transmitter is placed at a distance of 800-m from the receiver. Moreover, $\theta_{rx} = 120°, \varphi_{rx} = 325°$ and $\theta_{tx} = 20°, \varphi_{tx} = 70°$ are the position angles to the receiver and transmitter from the IRS, respectively. By using the derived analytical solution, the optimal IRS orientation angles were computed to be $\theta_n = 62°$ and $\varphi_n = 348°$. On the other hand, Figure 18 shows the variation of the received power from the Tx-IRS-Rx path with the IRS orientation angles as computed through numerical calculations. Figure 18 shows that indeed the received signal power is maximum when the IRS orientation matches the optimal IRS orientation angles obtained by the analytical solution.

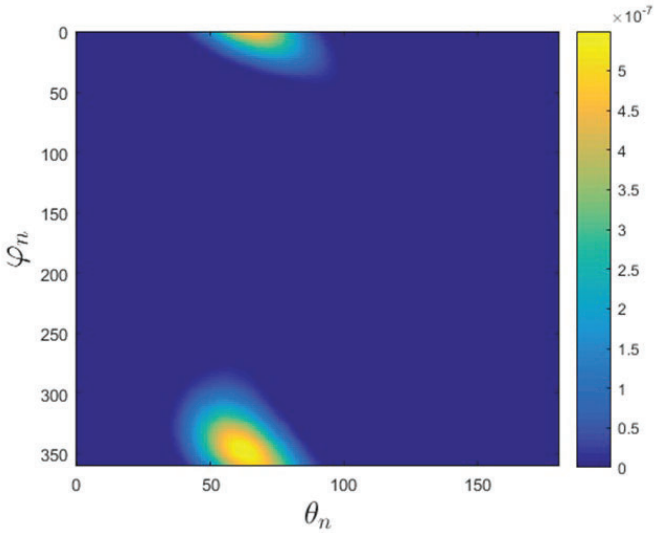


Figure 18. Variation of the Received Power in Watts from Tx-ITS-Rx Path with θ_n and φ_n

Figure 19 shows the received power variation as a function of the distance r_I^r between the receiver and the IRS when the IRS is oriented (a) optimally, (b) randomly, and (c) randomly in between the transmitter and receiver location vectors (i.e., $\mathrm{E}\{p_\Omega\}$). Figure 19 shows that the difference between the expected received power when the IRS orientation is completely random and when it is optimal is significantly large. For the given example, it is seen to be about 21.79 dB. Moreover, as can be seen from Figure 19, the expected received power difference between the optimal IRS orientation and $\mathrm{E}\{p_\Omega\}$ can also be significant. For the given example, it is about 13.82 dB.

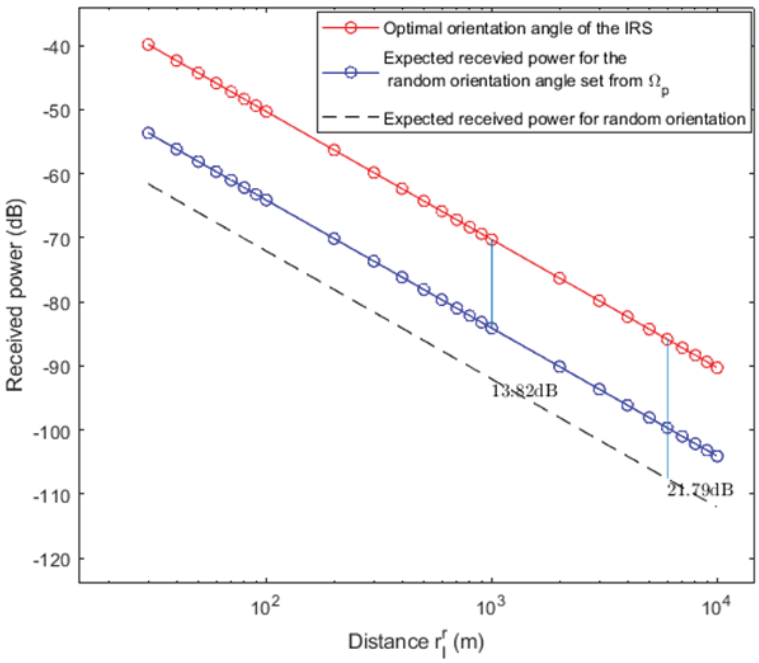


Figure 19. Variation of the Received Power from Tx-ITS-Rx Path for a Specific Location Tx and Rx With Respect to the r_I^r

In the next example, we consider an IRS placed near the ground receiver of a GEO downlink operating on frequency of 74 GHz. The IRS is placed at a position of r_I^r= 20-m, $\theta_{\mathrm{rx}}^{\mathrm{I}} = 60^{\mathrm{o}}, \varphi_{\mathrm{rx}}^{\mathrm{I}} = 145^{\mathrm{o}}$ whereas the GEO satellite is at an altitude of 36000-km. Moreover, $\{\theta_{\mathrm{rx}} = 120^{\mathrm{o}},\ \varphi_{\mathrm{rx}} = 325^{\mathrm{o}}\}$ and $\{\theta_{tx} = 20^{\mathrm{o}}, \varphi_{\mathrm{rx}} = 70^{\mathrm{o}}\}$ are the position angles to the ground receiver and satellite transmitter from the IRS, respectively.

As before, we assume an IRS with a total number of elements of 250 × 250 with each element size of $d_x = \lambda$ and $d_y = \lambda$. Figure 20 shows the variation of power received from the IRS as a function of the orientation of the IRS. As can be seen from Figure 20, the ground receiver received maximum power through the IRS when the orientation of the IRS is $\{\theta_{\mathrm{n}} = 62^{\mathrm{o}}, \varphi_{\mathrm{n}} = 348^{\mathrm{o}}\}$. We note that this is exactly the same as the analytical solution.

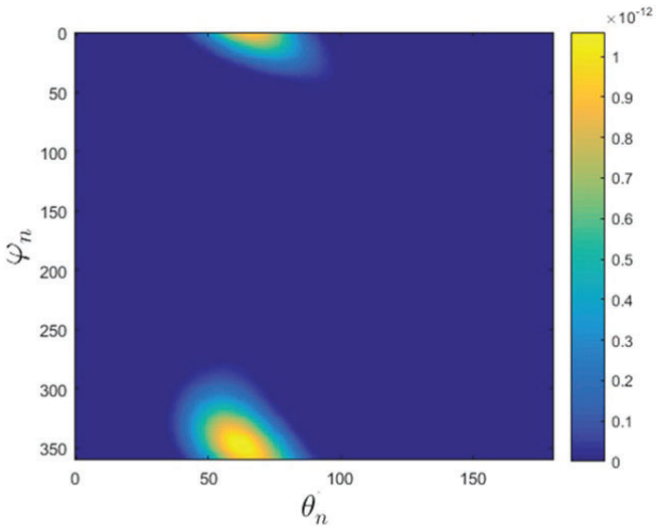


Figure 20. Variation of the Received Power as a Function of Orientation Angles, θ_n, φ_n

Figure 21 shows the optimal orientation angles of the IRS for the above example. It is seen that indeed the optimal orientation angles of the IRS lie between the corresponding angles of the locations of the ground receiver and the satellite transmitter.

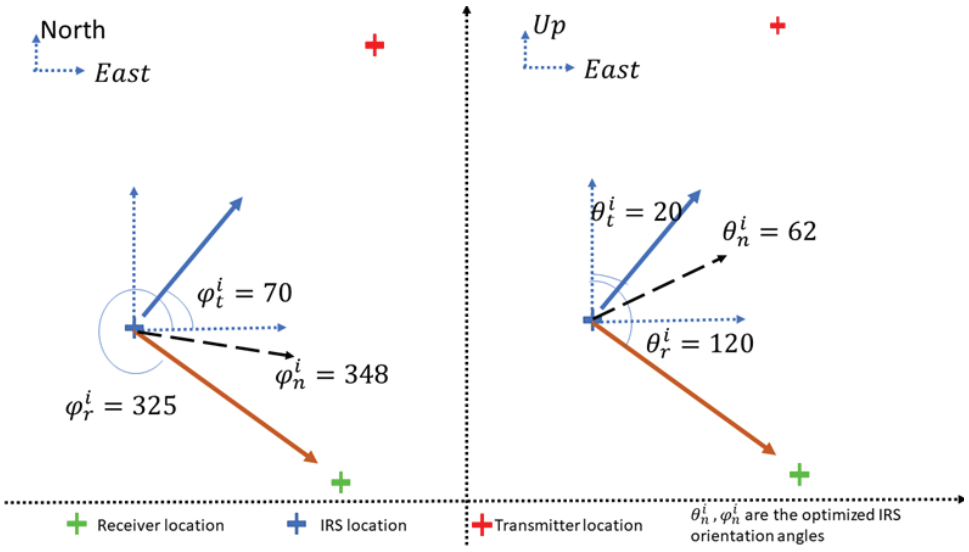


Figure 21. The Optimal IRS Orientation Angles

4.2 Coverage area Optimization

To highlight the variation in the achievable rate with respect to the IRS orientation, IRS beamforming and transmitter/receiver beamforming are performed at the respective ends. The BS location with respect to the IRS location is assumed to be r_I^t = 500-m, $\theta_I^{tx} = 30°$, $\varphi_I^{tx} = 50°$. For this case study, the low signal strength area A_ϵ is assumed to be defined by 25-m $\leq \rho \leq$ 50-m, $110° \leq \varphi_I^{rx} \leq 130°$, and -30-m $\leq z \leq$ -50-m (with respect to an ENU coordinate system at the IRS).

Figure 22 shows the average received power from the transmitter-IRS-receiver path for receivers located in the low signal strength area with respect to the IRS orientation angles. As can be seen from Figure 22, the maximum power is delivered to the selected low signal strength area when the IRS orientation angle is (θ_n=72°, φ_n =95°) which matches the analytical solution given in the previous section.

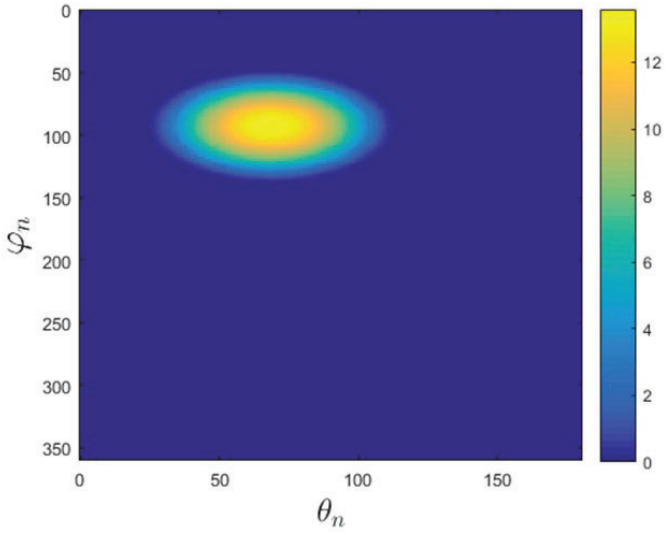


Figure 22. Average Received Power in Selected Area from Tx-ITS-Rx Path for a Specific Tx Location

Figure 23 shows the received power variation as a function of the average distance r_I^r between the receiver and the IRS when the IRS is oriented (a) optimally, (b) randomly, and (c) randomly in between the transmitter and receiver location vectors (i.e., $p_{\Omega_{A_\epsilon}}$).

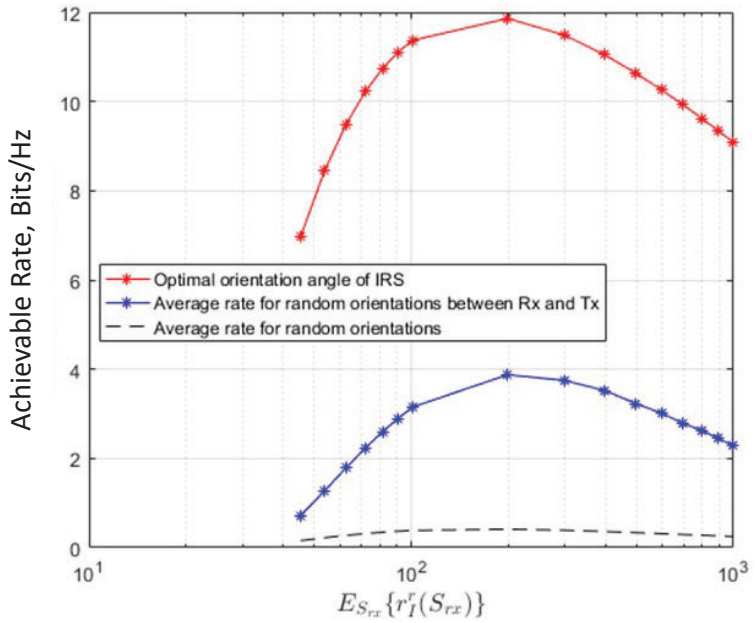


Figure 23. Average Received Power in Selected Area from Tx-ITS-Rx Path for a Specific Tx Location with Respect to the Average Distance r_I^r

As shown in Figure 22 and Figure 23, there is a 10.9 bps/Hz achievable data rate improvement with the optimally oriented IRS compared to the randomly oriented IRS. As before, the optimal IRS orientation angle is always obtained in between the low signal strength area and the BS transmitter. If we denote the set of IRS orientation angles that lie between the low signal strength area and the BS transmitter by Ω_ϵ, it is seen that there is 8.1 bps/Hz achievable rate improvement when the optimal IRS orientation is used as opposed to when the IRS orientation is set to be along an arbitrary direction in set Ω_ϵ.

Figure 23 also shows the average achievable rate with the average distance $\mathrm{E}_{\mathrm{S_{rx}}}\{r_I^r(S_{rx})\}$ to the weak signal area A_ϵ from the IRS. According to Figure 23, the maximum average achievable rate in A_ϵ is achieved when $\mathrm{E}_{\mathrm{S_{rx}}}\{r_I^r(S_{rx})\}$ is approximately 110-m for the above considered case study. Hence, when the position of the IRS is fixed, there exists an $\mathrm{E}_{\mathrm{S_{rx}}}\{r_I^r(S_{rx})\}$ that maximizes the average maximum achievable rate. As a result, the IRS position can also be optimized to maximize the average achievable rate at locations within A_ϵ.

4.3 Beamformer Weights Update Period Simulation

First, to investigate the accuracy of the Δt_{approx} derived using only the information available at time *t*, consider a UAV transmitter and a receiver where the receiving UAV is initially located at $\theta_{tx}^{rx}(t) = 0$ and is traveling in a tangential direction to the normal vector to the transmit antenna array. Based on the receiver location and velocity, the beamforming weight vector update timing can be calculated using Δt_{approx}. On the other hand, the time it takes for it to cross the HPBW

range of the transmitter UAV is the Δt^* computed by taking into account the actual trajectory of the receive UAV. We may define the error due to the approximation as $e_{\Delta t} = \Delta t^* - \Delta t_{approx}$. Figure 24 shows the variation of $e_{\Delta t}$ with respect to HPBW of the transmit antenna at different relative speeds of the receive UAV. As can be seen in Figure 24, the error due to Δt_{approx} increases as the HPBW of the transmit antenna increases. This can be attributed to the fact that the tangential velocity at time t only holds for time t to $t + \Delta t$ if the arc length is small in Figure 6. However, note that Figure 24 is based on the assumption that the RX UAV travels in a fixed direction during time t to $t + \Delta t$. In practice, this may not be true and the choice of tangential direction may impact the performance obtained using Δt_{approx} as we will show in later results.

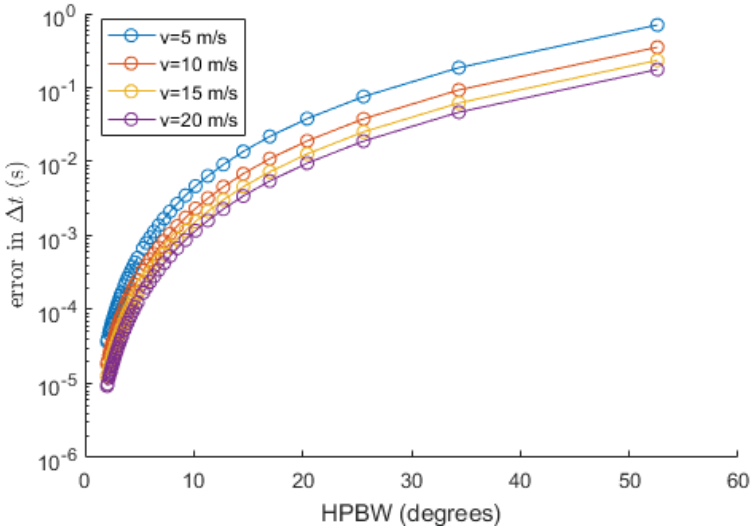


Figure 24. Variation of Error Between Exact and Approximate Δt with Respect to the HPBW of an Antenna

Next, we investigate the behavior of $\Delta t^*$with respect to the operating frequency and determine an effective method to calculate Δt_{approx} in a given scenario. Recall that $\underline{w}_t$ needs to be updated within Δt^* time to maintain the normalize beamforming gain $G_{B,tx} > G_{B,tx,min}$. We assumed that the antenna aperture is fixed at 100-cm^2, $G_{B,tx,min}$= 0.5, $\frac{d_{tx}^x}{\lambda} = 0.5$, $\frac{d_{tx}^x}{\lambda} = 0.5$, r = 30-m, and the relative speed of the receiving UAV in a tangential direction with respect to the transmitter is v = 10 ms^{-1} [33] This corresponds to a relative angular velocity of the receiver with respect to the transmitter $\omega_{rx} = 0.33$ rad/s.

Note that the maximum aperture of the transmit antenna can be approximated by the physical area of the antenna, $A_{e,tx} = A_{p,tx}$ where $A_{p,tx}$ denotes the physical area of the transmitter antenna [16]. Since $A_{p,tx}$ of the antenna can be expressed as $A_{p,tx} = (M_{tx}^x - 1)(M_{tx}^y - 1)d_{tx}^x d_{tx}^y$, for a fixed aperture antenna, the number of antenna elements increases as the frequency increases. Hence, for

a fixed aperture antenna, the radiation pattern beamwidth becomes narrower when operating frequency increases. Therefore Δt^* decreases with increasing frequency as shown in Figure 25.

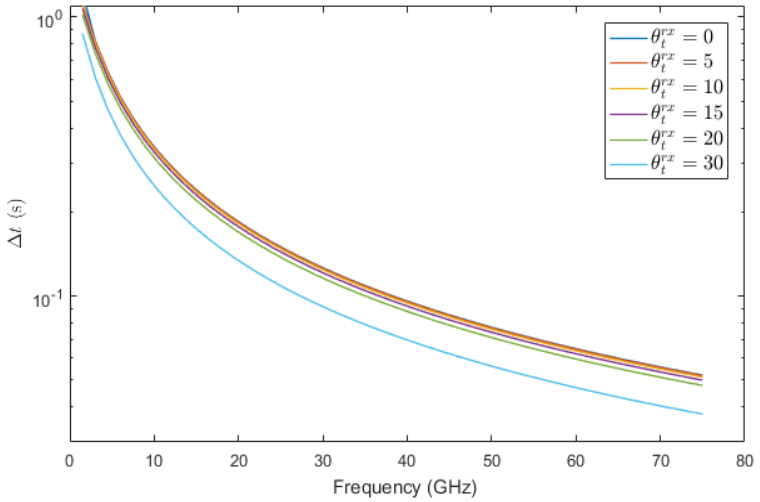


Figure 25. Beamforming Update Time Period Variation with Frequency for Different Receiver Locations with Fixed Transmitter Aperture 100-cm², ω = 0.33 rad/s, and *r* = 30-m

Figure 25 also shows the variation of Δt^* with the operating frequency when different initial elevation positions are assumed for the receive UAV. Recall from Figure 7b that the shape of the boundary line of the HPBW changes with the change in transmit beam direction. Therefore, Δt^* varies depending on the position of the receive UAV at time t according to the expression for Δt^*. Figure 25 illustrates that is smaller Δt^* when the transmitter orientation is aligned with receive UAV so that $\theta_{tx}^{rx}(t) = 0$, than when they are not (i.e., $\theta_{tx}^{rx}(t) \neq 0$).

Smaller Δt^* requires frequent transmission of overhead training bits. Let us consider the same scenario as that in Figure 25. Figure 26 shows the dependence of the ratio of the overhead bit rate to the bit rate on the operating frequency. Clearly, the relative overhead is lower at higher frequencies even with frequent weight updates.

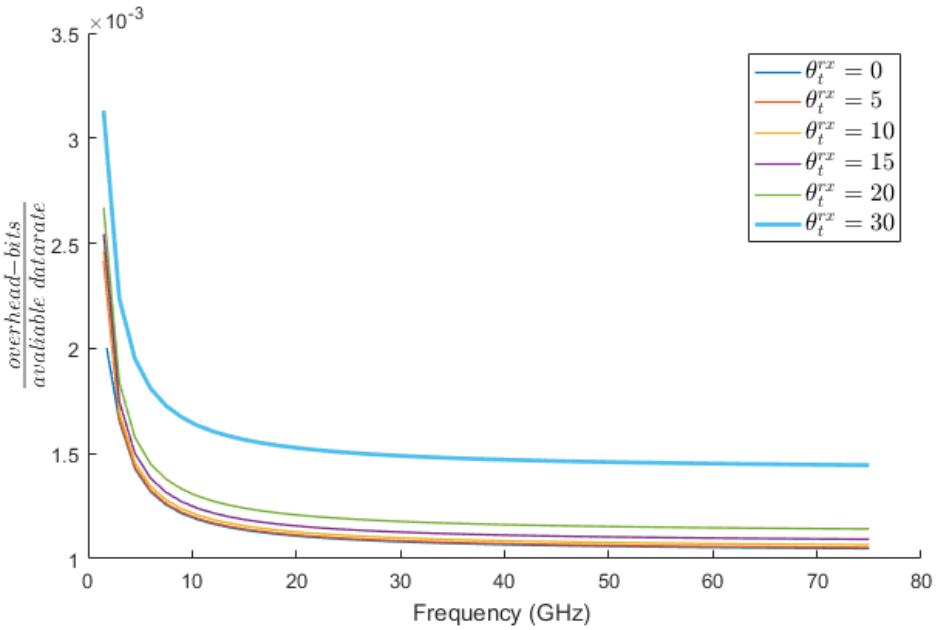


Figure 26. Ratio of Overhead Bit Rate and Available Bit Rate

4.3.1 2D UAV Network Simulation

To investigate above findings in an actual UAV communications system, next we consider two UAVs flying over a square area of side-length 1-km divided into 50 × 50 identical square cells [34]. Without loss of generality, we may assume that the UAV transmitter is fixed at the center of the square area, while the UAV receiver moves around it. At each step, it moves from one cell center to the nearest cell center. With this model, the movement of the receiver UAV is restricted to eight directions namely: North, North-East, East, South-East, South, South-West, West and North-West. Let b_j denote the j-th possible movement direction, for $1 \leq j \leq 8$, and $\mathcal{B} = \{b_j, 1 \leq j \leq 8\}$ denote the set of possible directions a UAV is allowed to move. It is assumed that the UAV receiver travels between the cells at the constant speed of $v = 10\ ms^{-1}$, the antenna aperture is $A_{e,t} = 100\text{-cm}^2$ and the operating frequency is 60 GHz.

Figure 27 shows the variation of the beamforming gain, $G_{\text{B,tx}}$, with respect to time as the receiver moves around. As shown in Figure 27, the Δt_{approx} results in beamforming weights getting updated about 1532 times within 240 seconds of simulation time. In those 1532 instances, the beamforming gain drops 10 times below the half power required according to $G_{\text{B,tx,min}}$= 0.5. On the other hand, according to the Δt^*, the beamforming weights only needed to be updated 153 times and it completely avoids beamforming gain dropping below half power level, as can be seen from Figure 27.

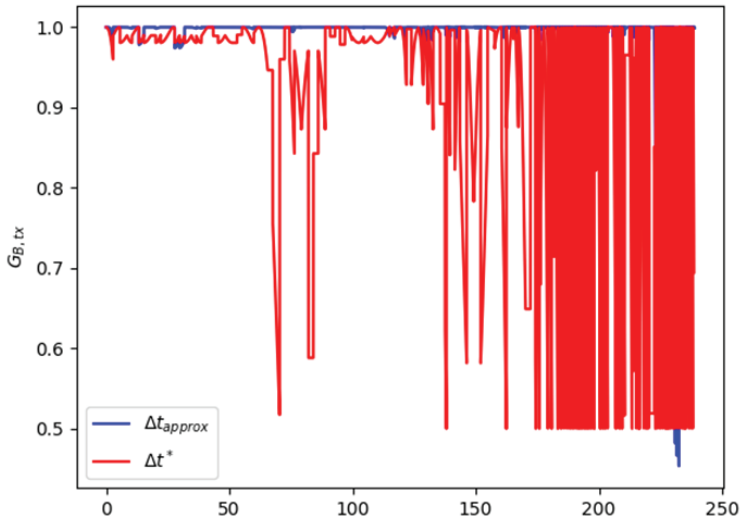


Figure 27. Variation of the Normalized Beamforming Gain with Time when Using Δt^* and Δt_{approx}

In our UAV system simulation, we assume that at each location the receiver selects a travel direction on its own. Hence, the transmit UAV only knows the current location of the UAV receiver but not which direction it may travel next. Clearly, the required Δt depends on the travel direction (Action) selected by the UAV receiver. Since the transmit UAV knows that the UAV receiver's movements are restricted to the possible 8 directions in the action space, a possible solution is to use the average Δt averaged over all possible 8 directions as the update time period.

Figure 28 shows the beamforming gain $G_{\mathrm{B,tx}}$ over time when $\Delta t = E\{\Delta t_{approx}(b)\}$ where $E\{\}$ denotes the expectation assuming all directions $b_j \in \mathcal{B}$ are equally likely. Using this method, the beamformer weights are updated 379. As can be seen from Figure 28, the beamforming gain drops below 0.5 for 148 seconds during the total 700 seconds of simulation time (i.e., 22% from the total simulation time).

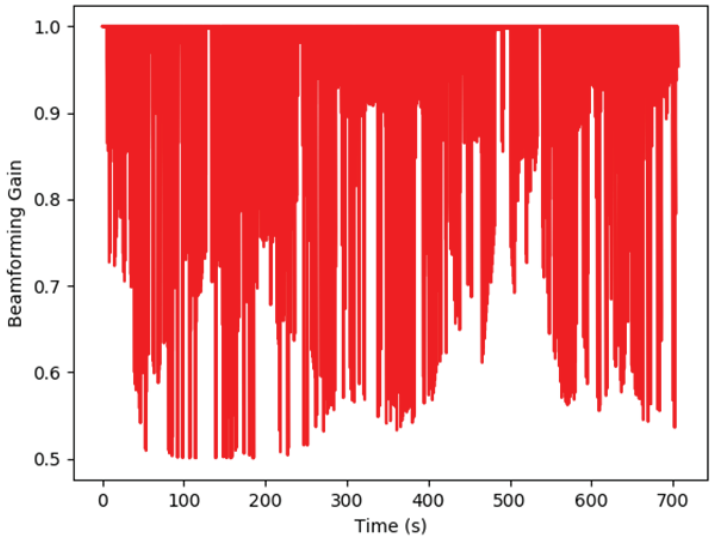


Figure 28. Variation of the Normalize Beamforming Gain with Time When Using Approximate Equation in the Simulation Model when $\Delta t = E\{\Delta t_{approx}(b)\}$

An alternative is to use $\Delta t = min_{b\in\mathcal{B}}\{\Delta t_{approx}(b)\}$ where minimum is over all possible 8 directions. Figure 29 shows the beamforming gain $G_{\mathrm{B,tx}}$ over time when $\Delta t = min_{b\in\mathcal{B}}\{\Delta t_{approx}(b)\}$ is used. During the total simulation time, the beamformer is now updated 974 times. However, as Figure 29 shows, the beamforming gain drops below 0.5 only for 0.02 seconds (i.e., 0.002% from the total simulation time), a significant improvement over average Δt_{approx} criterion.

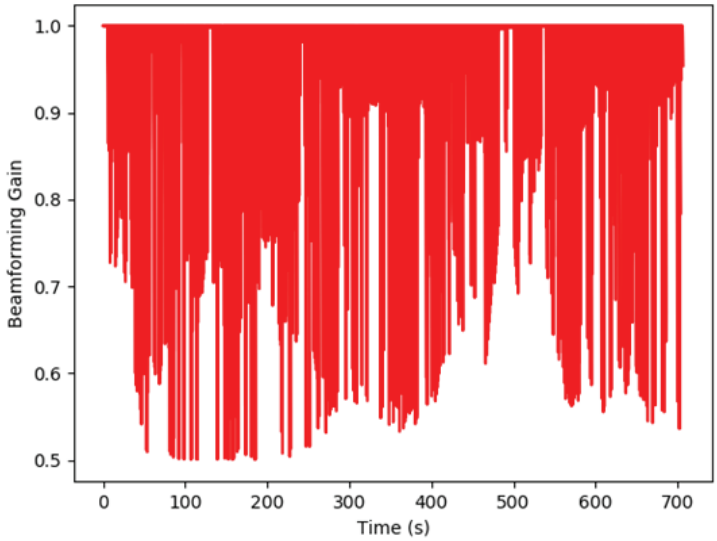


Figure 29. Variation of the Normalize Beamforming Gain with Time When Using Approximate Equation in the Simulation Model when $\Delta t = min_{b\in\mathcal{B}}\{\Delta t_{approx}(b)\}$

Figure 30 examines the distribution of the time durations during which beamforming gain drops below the required threshold of $G_{\mathrm{B,tx}} > 0.5$ when $\Delta t = E\{\Delta t_{approx}(b)\}$. As can be seen from Figure 30, the time duration of beamforming gain dropping below the required threshold has a mean of about 208 seconds (over the total simulation time of 700 seconds).

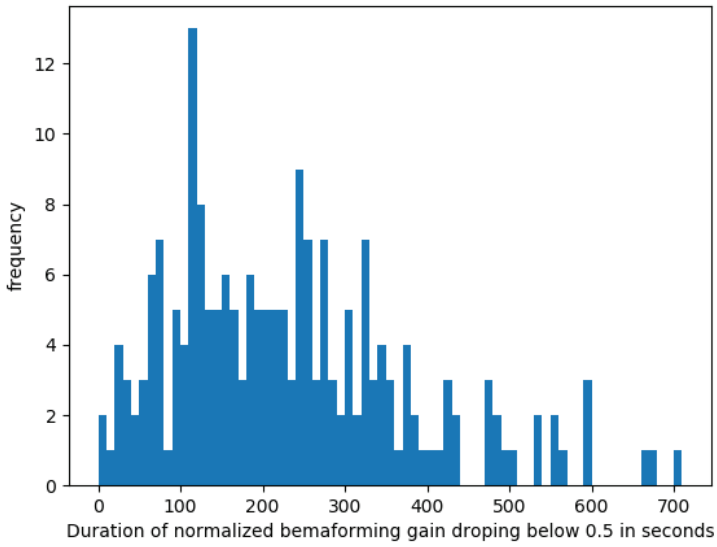


Figure 30. Histogram of Total Time Duration Where $G_{\mathrm{B,tx}}$ Drops Below 0.5 When $\Delta t = E\{\Delta t_{approx}(b)\}$

Figure 31 and Figure 32 show the comparison of histograms of the number of times beamformer gets updated during simulations for each of the Δt calculation methods. When $\Delta t = min_{b\in\mathcal{B}}\{\Delta t_{approx}(b)\}$, the beamformer weights are updated more frequently than when $\Delta t = E\{\Delta t_{approx}(b)\}$. However, it reduces the time duration where $G_{\mathrm{B,tx}} < 0.5$ to almost zero during the simulation time.

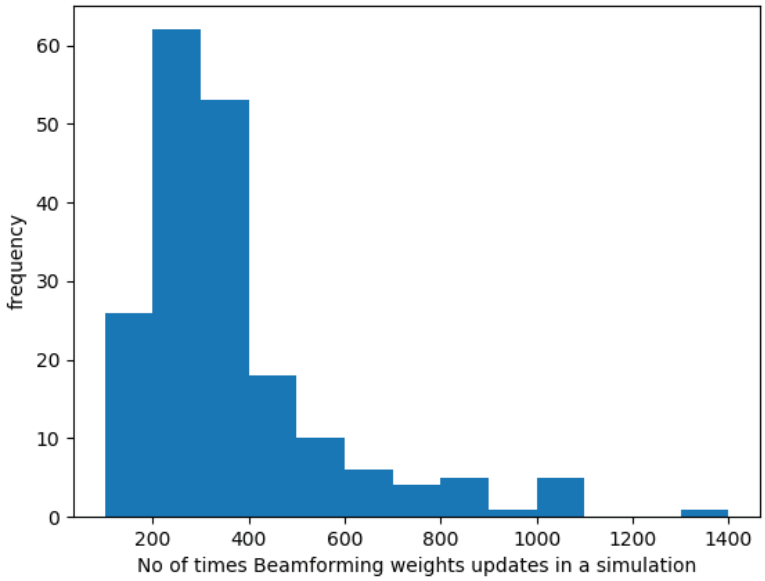


Figure 31. Histogram of Number of Beamformer Weights Update Times When $\boldsymbol{\Delta t = E\{\Delta t_{approx}(b)\}}$

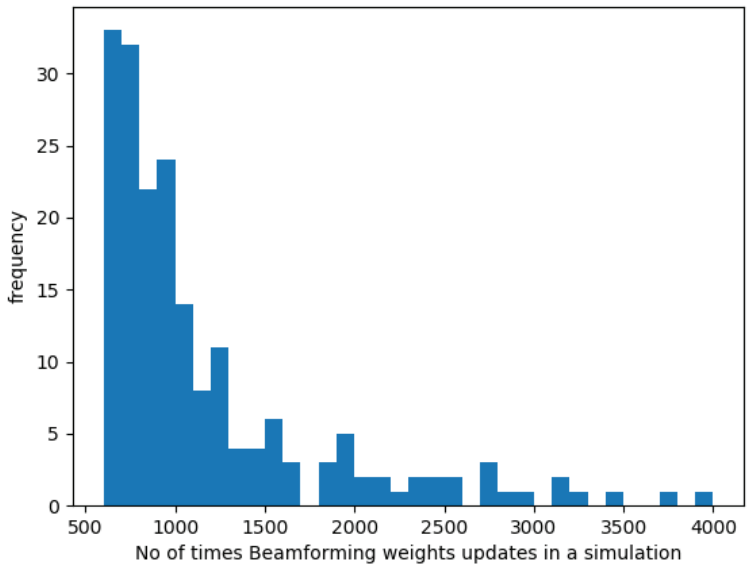


Figure 32. Histogram of Number of Beamformer Weights Update Times When $\boldsymbol{\Delta t = min_{b\in\mathcal{B}}\{\Delta t_{approx}(b)\}}$

4.3.2 3D UAV Network Simulation

To further investigate the implications of above findings in a U2U communication system, let us consider a 3D geographical area of 1-km × 1-km × 100-m divided into 50 × 50 × 10 cuboids of

equal size. Assume that the transmit UAV is located at the center of the cuboidal geographical area and the receiving UAV is moving from the center of a cuboid cell to that of its nearest cell. This model restricts the movement of the receive UAV to a maximum of 26 directions in 3D space. Let a_i denote the i-th possible movement direction, for $1 \leq i \leq 26$, and $\mathcal{A} = \{a_i, 1 \leq i \leq 26\}$ denote the set of possible directions a UAV is allowed to move. In simulations, we assume that the receive UAV travels at a constant speed of 10 ms^{-1} transmit UAV antenna has an aperture of $A_{p,tx}$ = 100-cm^2 and the operating frequency of 60 GHz. Since 2D network scenario considered previously shows that the number of times where the beamforming gain drops below $G_{\mathrm{B,tx}} < 0.5$ is minimized by updating Δt according to $\Delta t = min\{\Delta t(a_i)\}$, in the following we limit ourselves to this choice.

Figure 33 shows the variation of $G_{\mathrm{B,tx}}$ with time when using Δt^* and Δt_{approx} for updating $\underline{w_t}$. Note that, Δt^* is calculated using information on full trajectory of the receiving UAV. Hence, to obtain a fair comparison Δt_{approx} is also calculated using the information on present location and velocity of the receiving UAV. In Figure 33, the use of Δt_{approx} results in updating $\underline{w_t}$ approximately 532 times within the 240 seconds long simulation time. In those 532 instances, $G_{\mathrm{B,tx}}$ drops below the minimum threshold 9 times. With the use of optimal Δt^*, on the other hand, $\underline{w_t}$ is updated only 53 times during the simulation time and, as expected, never results in $G_{\mathrm{B,tx}}$ dropping below the minimum threshold. This confirms that having the knowledge of the entire flight path of the receive UAV reduces the frequency of $\underline{w_t}$ updates. However, from now on we will use Δt_{approx} to update $\underline{w_t}$ due to unavailability of receive UAV flight path in a realistic U2U communications network.

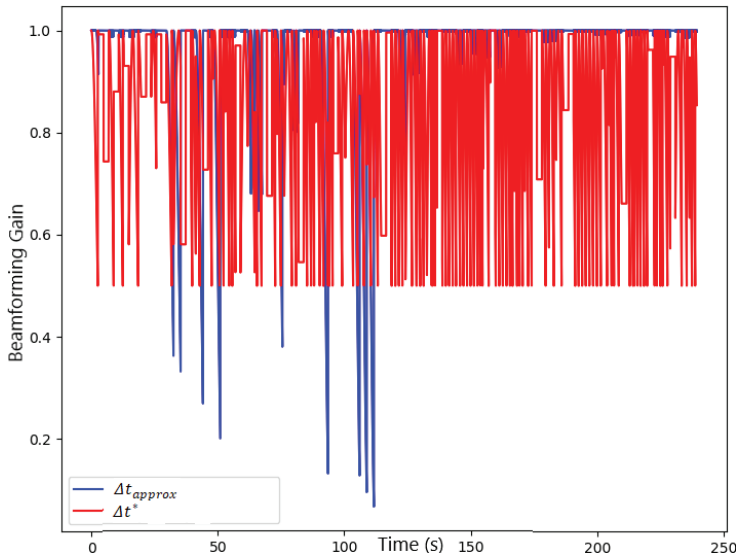


Figure 33. Variation of $G_{B,tx}$ with Time When Using Approximation Δt_{approx} and Optimal Δt^*

At any given time t, the receive UAV may choose to stay at the current location itself or move in one of the possible 26 directions. Clearly, the required update time Δt_{approx} will depend on this selection, but it is unknown a prior to the transmit UAV. In this case, we consider receive UAV travels 350 steps which is approximately 950 seconds in each simulation. Figure 34 shows the variation of $G_{B,tx}$ during the simulation time when $\Delta t = min_{a\in\mathcal{A}}\{\Delta t_{approx}(a)\}$ where $\Delta t_{approx}(a)$ is the Δt_{approx} computed assuming that the receive UAV selects direction $a \in \mathcal{A}$. During the simulation time, $\underline{w_t}$ is updated 1023 times with this approach for computing Δt but $G_{\mathrm{B,tx}}$ drops below $G_{\mathrm{B,tx,min}}$ only for 0.005 seconds which is 0.0005% in total simulation time. Alternatively, if $\Delta t = E\{\Delta t_{approx}(a)\}$ is used, where $E\{\}$ denotes the expectation assuming all directions $a_i \in \mathcal{A}$ are equally likely, it increases to 5.2% of the total simulation time as shown in Figure 35 but reduces the number of $\underline{w_t}$ update times.

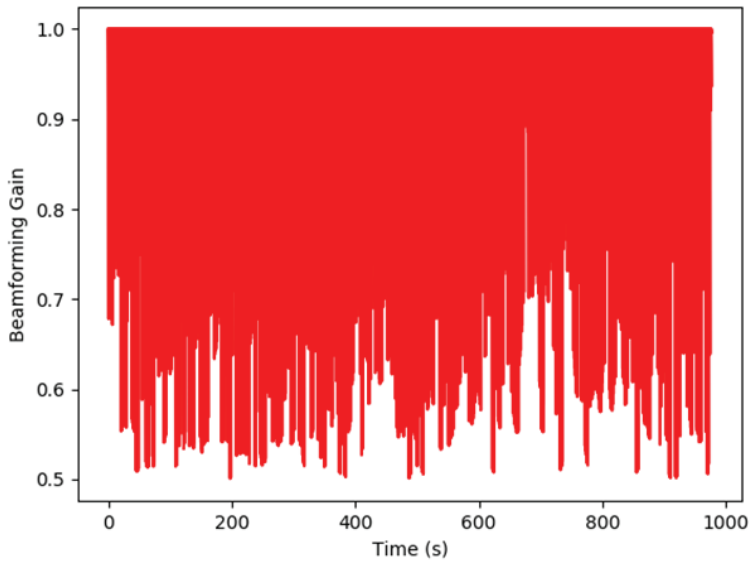


Figure 34. Variation of the Normalize Beamforming Gain $G_{B,tx}$ With Time When Using $\Delta t = \min_{a\in\mathcal{A}}\{\Delta t_{approx}(a)\}$ in 3D Simulation Model

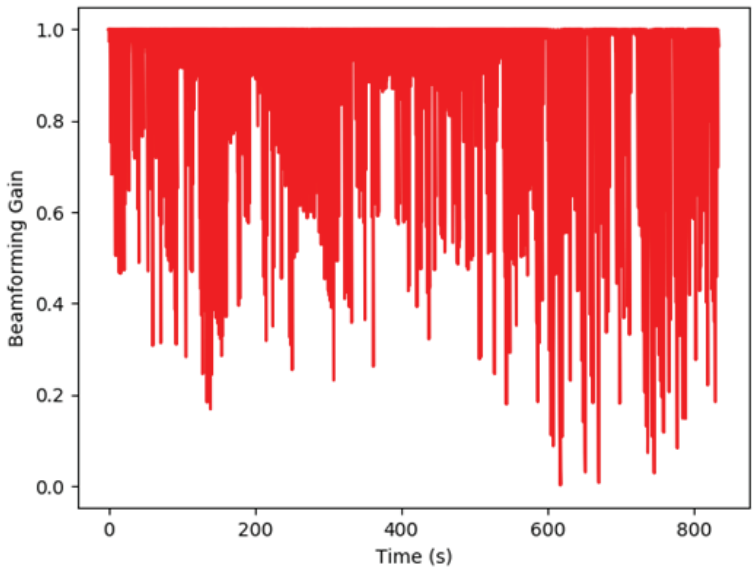


Figure 35. Variation of the Normalize Beamforming Gain $G_{B,tx}$ With Time When Using $\Delta t = E\{\Delta t_{approx}(a)\}$ in 3D Simulation Model

For histograms in Figure 36, Figure 37, and Figure 38, we used 200 simulations for each Δt_{approx} calculation method. Figure 36 shows the histogram of time duration when $G_{B,tx} < G_{B,tx,min}$ in simulations with $\Delta t = E\{\Delta t_{approx}(a)\}$. The average time duration of $G_{B,tx} < G_{B,tx,min}$ in simulations is 254.8 seconds when $\Delta t = E\{\Delta t_{approx}(a)\}$. It was also observed that the average time duration of $G_{B,tx} < G_{B,tx,min}$ is negligible (close to zero) if we were to use $\Delta t = min_{a\in\mathcal{A}}\{\Delta t_{approx}(a)\}$. However, this requires faster beamforming weighs which leads to increase of overhead bitrate.

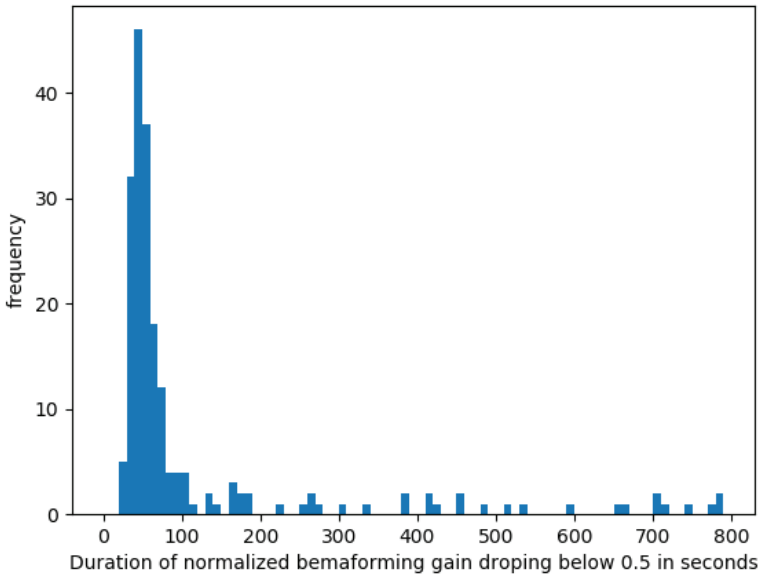


Figure 36. Histogram of Total Time Duration Where $G_{B,tx}$ Drops Below 0.5 When $\Delta t = E\{\Delta t_{approx}(a)\}$

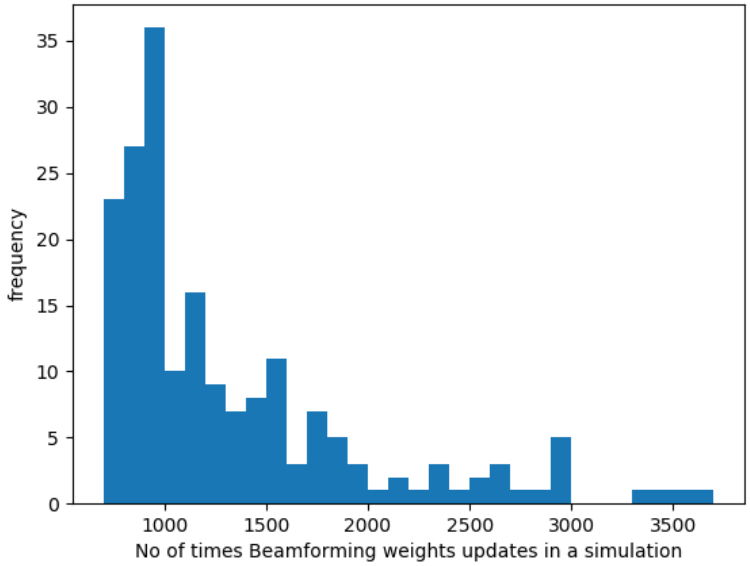


Figure 37. Histogram of Number of Beamformer Weights Update Times When $\Delta t = E\{\Delta t_{approx}(a)\}$

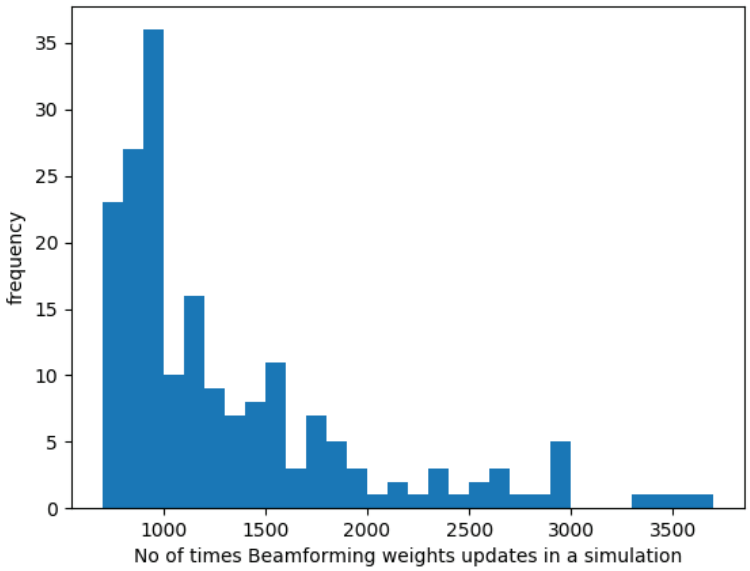


Figure 38. Histogram of Number of Beamformer Weights Update Times When $\Delta t = min_{a \in A} \{\Delta t_{approx}(a)\}$

To examine the increment of the overhead bitrate, Figure 37 and Figure 38 illustrate the frequency distribution of $\underline{w}_t$ update in simulations. It clearly shows that the average of $\underline{w}_t$ update increase 72.6% if we were to use $\Delta t = min\{\Delta t_{approx}(a)\}$. Hence, overhead bitrate increases 72.6% in comparison to using $\Delta t = E\{\Delta t_{approx}(a)\}$.

Comparing Figure 37 and Figure 38, we note that the number of times beamforming weights are updated is comparatively lower when $\Delta t = E\{\Delta t_{approx}(a)\}$ is used. However, this results in beamforming gain dropping below the required threshold for a significant amount of time. In contrast, when $\Delta t = min\{\Delta t_{approx}(a)\}$ is used, the number of updates can be much higher, but it improves the overall performance by avoiding violation of the minimum beamforming gain requirement.

In summary, when $\Delta t = E\{\Delta t_{approx}(a)\}$, $G_{B,tx}$ drops below the minimum threshold more frequently. However, the frequency of $\underline{w}_t$ updates is comparatively lower. In contrast, when $\Delta t = min\{\Delta t_{approx}(a)\}$ the frequency of $\underline{w}_t$ updates is comparatively higher but $G_{B,tx}$ rarely drops below the minimum threshold resulting a stable connection.

4.4 Optimal UAV Trajectory Planning

Figure 39 shows the performance of the two proposed methods compared to the random action selection strategy when the users are static and uniformly distributed over the area of interest. For comparison, a Q-table based reinforcement learning approach used in [24] is also shown in Figure 39. The number of distinct ground users covered was used as the metric to assess the performance of the proposed methods. From Figure 39, it can be seen that the proposed methods using DQN

and the reward-based greedy algorithm are able to cover significantly more numbers of users in contrast to random action strategy as well as the previous state of the art approach. Specifically, the proposed methods seen to cover roughly about twice as many users as the previous approach of [24].

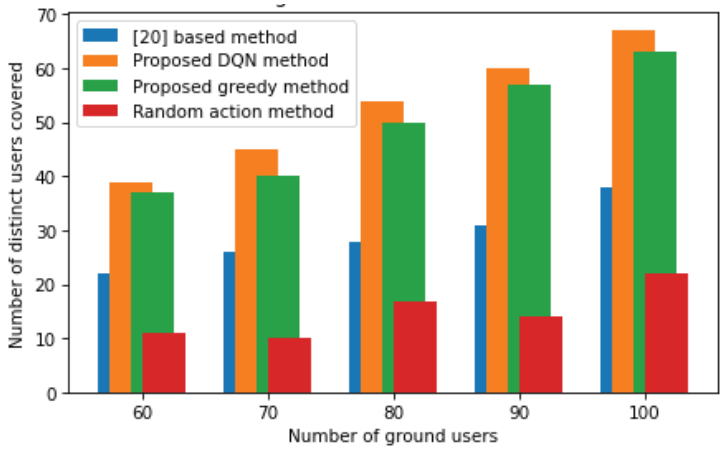
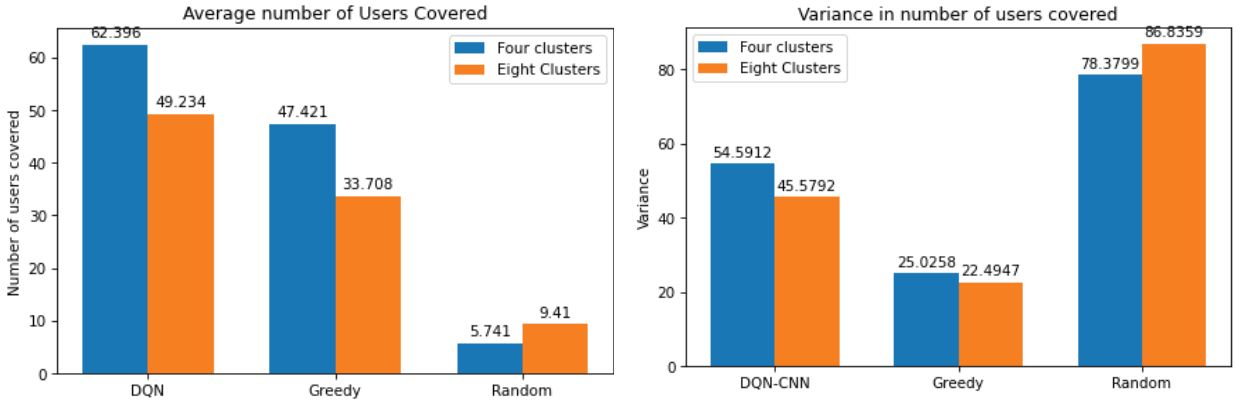


Figure 39. Number of Users Covered for Varying Number of Total Ground Users Uniformly Distributed Over the Area of Interest

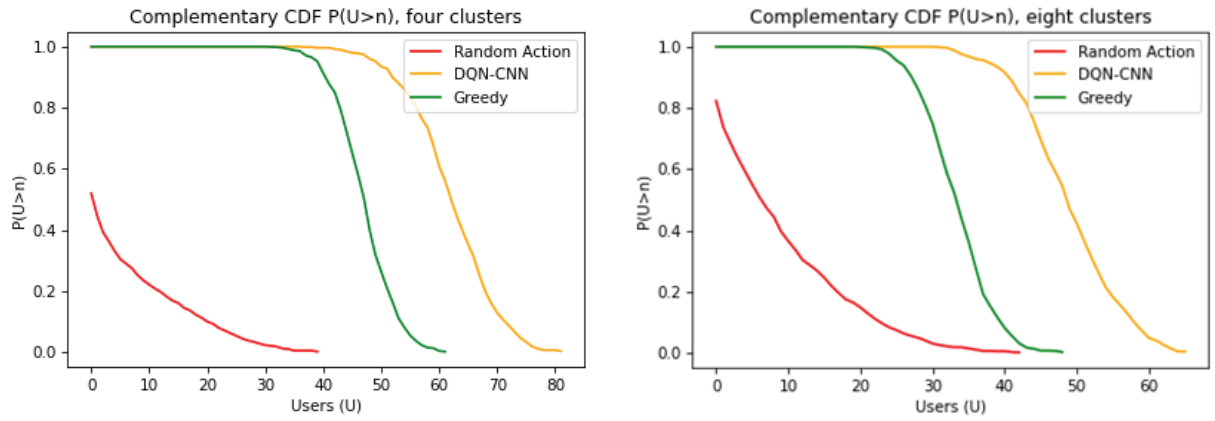
Figure 40(a) shows the average number of distinct users covered by the proposed DQN method (using a convolutional neural network, i.e., DQN-CNN), proposed reward based greedy method, and random action method -- when the distribution of users was modelled as a Gaussian mixture with four and eight clusters. It should be noted that in both the cases the total number of ground users was 100. Figure 40(b) shows the variance in number of users covered by the methods.



(a) (b)

Figure 40. Average and the Variance of Number of Users Covered Using Proposed Methods When Users are Distributed According to Gaussian Mixtures

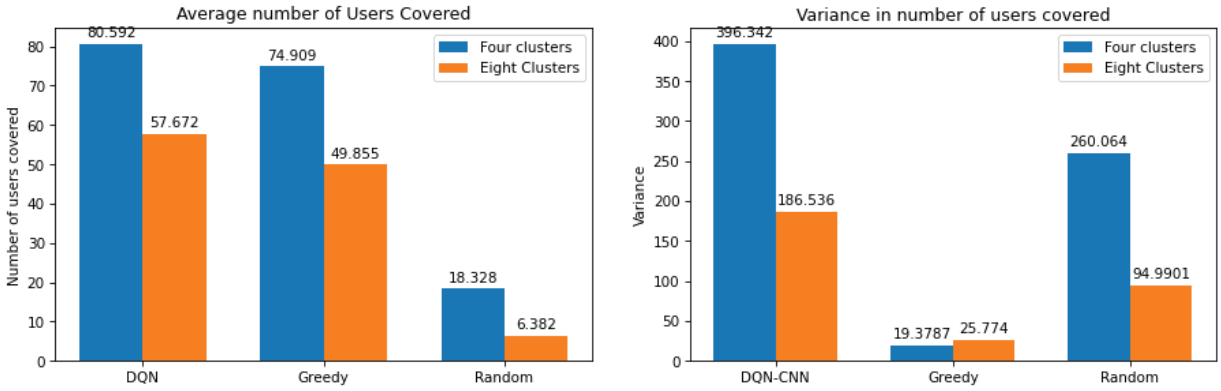
Figure 41(a) shows the probability of covering more than a given number of users by the UAV when users are distributed according to Gaussian mixtures. In other words, if U denotes the number of users covered, Figure 41(b) shows the Complimentary Cumulative Distribution Function (CCDF) of U defined as $P\ (U > n)$ (here, P denotes "probability").



(a) (b)

Figure 41. CCDF of Number of Users Covered When Users are Distributed According to Gaussian Mixtures

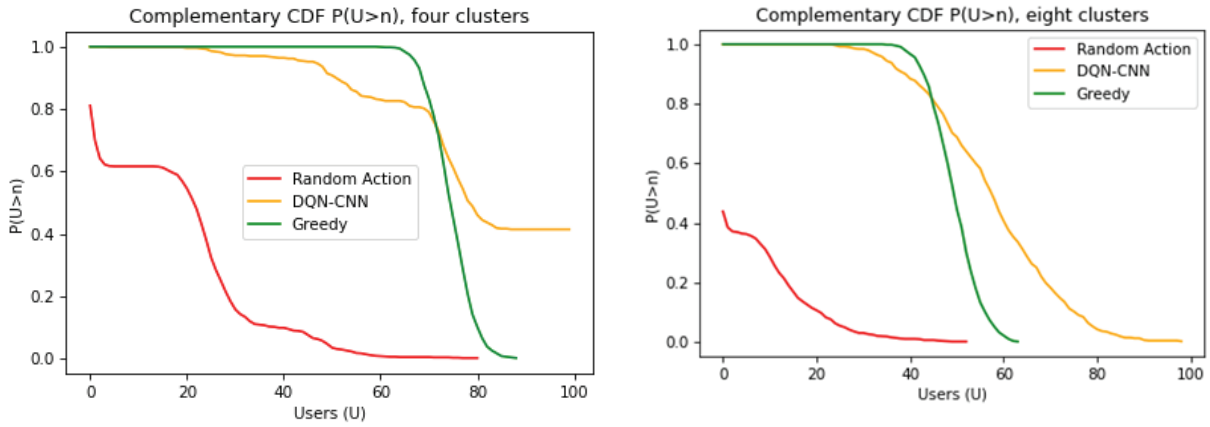
Figure 42 shows the corresponding results when the distribution of users was modelled as a Gaussian mixture with four and eight clusters but with more densely located users. As before, the total number of ground users was 100.



(a) (b)

Figure 42. Number of Users Covered and the Variance in Number of Users Covered with Densely Packed Users Distributed According to Gaussian Mixtures

Figure 43 shows that both the proposed DQN and the reward-based greedy methods are capable of covering more than 40 users with very high probability. Moreover, both proposed methods significantly outperforms the random action selection approach.



(a) (b)

Figure 43. CCDF of Number of Users Covered When Users are Densely Packed and Distributed According to Gaussian Mixtures

Next, we considered the case in which users are modeled as being mobile. In this case, the cluster means of the GMM were allowed to be time varying. Since the position of users at a given time instant is correlated with their previous locations, a Recurrent Neural Network (RNN) was used as the DQN. Figure 44 and Figure 45 show the average numbers of users covered and the CCDF of the numbers of users covered.

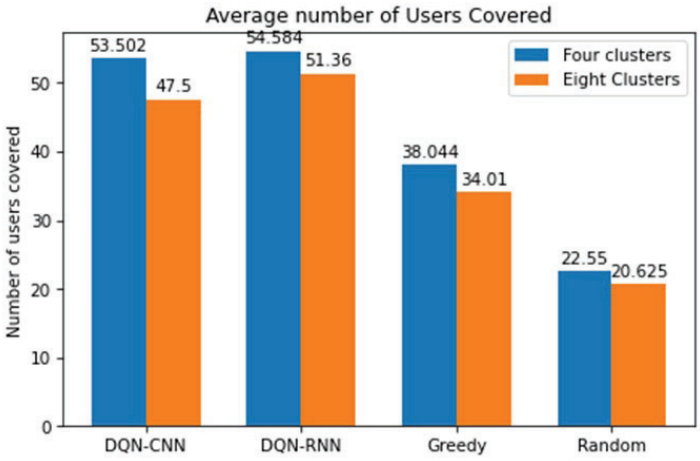
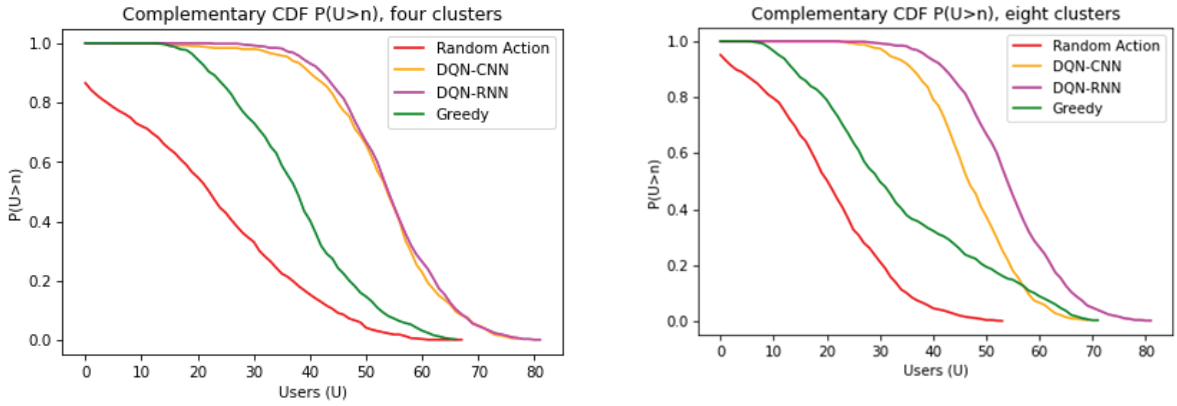


Figure 44. Number of Users Covered When Users are Mobile and Distributed According to Gaussian Mixtures



(a) (b)

Figure 45. CCDF of Number of Users Covered When Users are Mobile and Distributed According to Gaussian Mixtures

Figure 45 shows that the proposed DQN and rewards-based greedy methods are still able to cover more than 40 users with very high probability. Moreover, from Figure 45 we may see that the proposed DQN based method which uses RNN seems to outperform the DQN with a convolutional neural network (CNN) in terms of covering more ground users when the number of clusters is large.

4.5 Security and Robustness of Machine Learning Based Trajectory Planning Systems

We considered a square geographical area with 100 users spread across in four clusters according to a GMM model. Each cluster had equal number of users. In the fixed user position scenario, the locations of users remained the same throughout the epoch. In time varying position scenario, the locations of all ground users changed by a fixed magnitude and direction during the epoch. The performance of the UAV-BS in terms of the number of distinct users covered and the amount of user data relayed in the absence of adversaries was first investigated to provide a baseline to measure the impact of adversarial attacks. We considered four adversaries with spawning radius large enough to almost cover the square area (see Figure 16). The positions of the adversaries were assumed to be fixed regardless of whether the ground users were fixed or mobile and so were the number of fake users generated by each adversary.

Figure 46 shows the path taken by the trained UAV-BS in the absence and presence of adversaries during the learning phase of the DQN. It was observed that the DQN takes about 250 training epochs to learn an optimal flight path. The green circles denote the coverage area of the UAV-BS at each location of its trajectory. It can be observed by comparing Figure 46(a) and Figure 46(b) that the adversaries indeed have the potential to alter the trajectory of the UAV-BS drastically leaving large fractions of users without coverage.

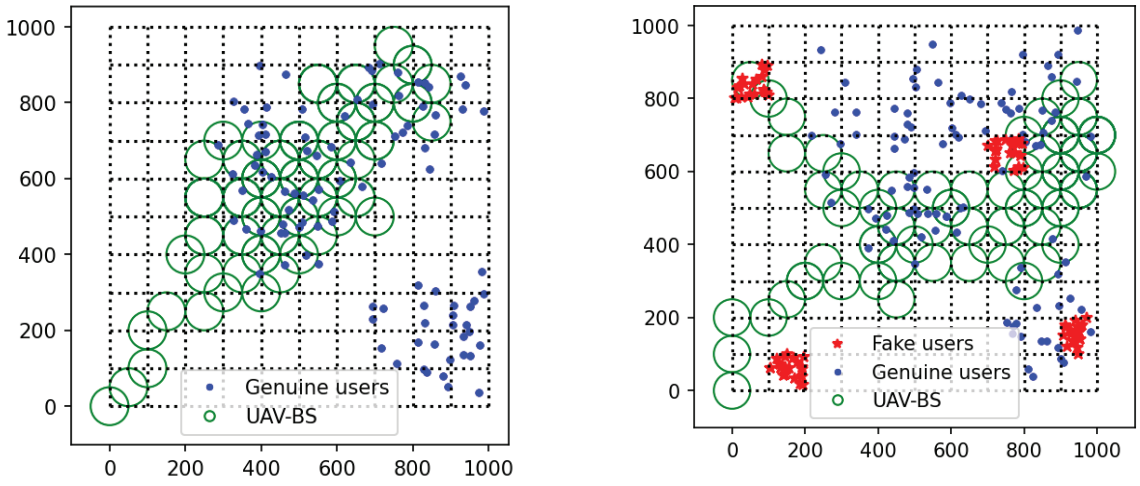


Figure 46. Trajectory of the UAV-BS in the (a) Absence and (b) Presence of Adversary

The UAV-BS trajectory itself does not provide the full details of the impact of adversaries. Figure 47(a) shows the time spent by the UAV-BS at the selected locations at each time step in the absence of spoofed users. On the other hand, when there are adversaries, the time the UAV-BS spent servicing fake users is illustrated in Figure 47(b). The pink circles indicate the minimum amount of time the UAV-BS spends at any location which is the waiting time, T_w. The yellow and red circles correspond to longer durations. From Figure 47(b) it is clear that under the Sybil attack, the UAV-BS wastes much of its time in serving counterfeit users leaving genuine users stranded.

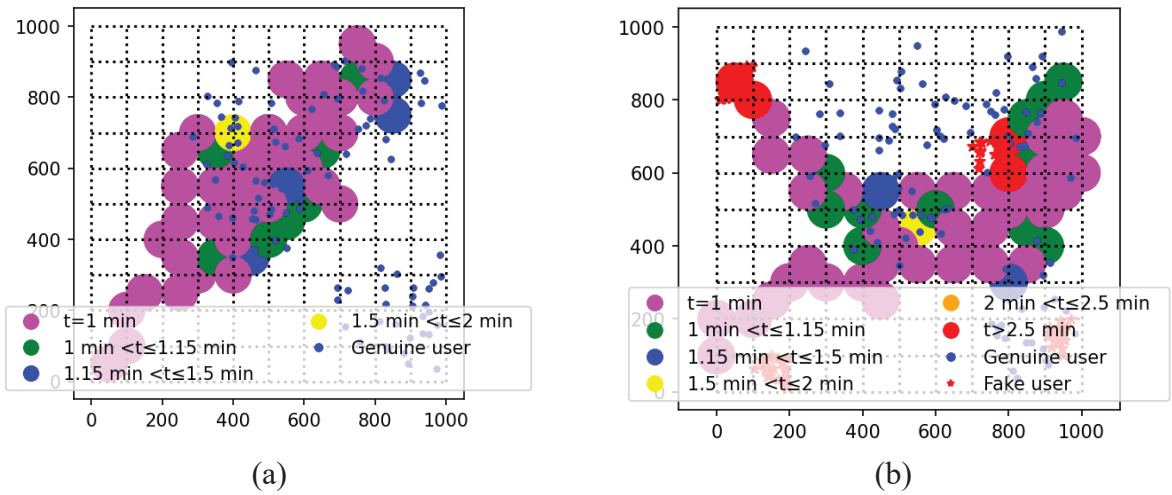


Figure 47. Time Spent at Each Location in Presence and Absence of Adversaries

Table 1 and Table 2 show the performance of UAV-BS in the two scenarios mentioned above. It can be seen that there is a definitive degradation in system performance due to the presence of adversaries. Indeed, the presence of adversaries could result in up to 20% reduction in number of distinct users serviced and up to 24% reduction in the data relayed in the case of stationary user distribution. Similarly, in the case of mobile users, up to 22% reduction in number of distinct users serviced and up to 10% reduction in data relayed is possible.

Table 1. Performance of the UAV-BS With Stationary Ground Users in the Presence and Absence of Adversaries

Experiment No.	Without Distinct Users	Adversaries Data relayed (GB)	With Distinct Users	Adversaries Data Relayed
1	59	9.46	53	8.16
2	51	8.24	41	6.23
3	38	6.13	33	4.76
4	61	9.72	58	8.13
5	57	9.26	46	6.45
Average	**53.2**	**8.56**	**46.2**	**6.75**

Table 2. Performance of the UAV-BS With Non-stationary Ground Users in the Presence and Absence of Adversaries

Experiment No.	Without Distinct Users	Adversaries Data relayed (GB)	With Distinct Users	Adversaries Data relayed (GB)
1	44	5.92	35	4.31
2	48	6.39	43	5.83
3	54	7.47	47	6.52
4	55	7.45	43	6.61
5	55	7.39	45	6.64
Average	**51.2**	**6.92**	**42.6**	**5.98**

5 CONCLUSIONS

The mmWave spectrum is currently being investigated as the possible solution to meeting the increasing demand for wireless services both in military as well as civilian applications. The large bandwidths possible with mmWave operation, however, come at the price of high attenuation seen by smaller wavelengths. It is not immediately clear, whether the communications theoretic foundations that have been developed up to now for microwave frequency bands, in particular sub-6 GHz spectrum, are still optimal and always relevant in taking advantage of these enormous bandwidths. In the specific case of satellite communications, in particular, it is of interest to investigate whether the imbalance in power and bandwidth constraints can better be exploited by developing new techniques. This project has started the investigation of two of the promising new concepts in wireless communications, namely the Intelligent Reflecting Surfaces (IRS) and 3D beamforming, in the context of mmWave satellite communications.

An analytical solution was obtained for the optimal IRS orientation to maximize the signal strength at a receiver location for a given fixed transmitter location. In addition, an optimization problem was formulated to maximize the average received rate at receiver locations distributed across a low signal strength area. Results from case studies showed that there can be a significant improvement in received signal strength when the IRS is orientated optimally. As a result, by manipulating the mechanical zenith and azimuth angle of the IRS, the coverage efficiency of the network can be significantly improved.

The investigation of 3D beamforming for mobile platforms showed that beamforming weights may need to be updated much more frequently when higher frequencies are used. To reduce computational complexity, simplified expressions were also derived for computing beamformer weight update periods. When using higher frequencies, however, it is possible to achieve much higher bit rates since the available bandwidths can be much larger. As a result, the overhead bit rate can still decrease relative to the overall bitrate of the system even with much more frequent weight updates. This shows that the mmwave frequencies may provide overall better efficiencies with 3D beamforming.

Application scenarios of 2D and 3D unmanned aerial vehicle (UAV) networks were considered to examine how to determine the transmit beamformer weight vector update timing (Δt) for maintaining a link between two UAVs. The optimal Δt^* required the knowledge of receive UAV flight path at the transmit UAV. As an alternative, an approximate solution for Δt that requires only the present estimate of receiver location was also proposed. It was shown that the proposed approximation for Δt has about 98% accuracy compared to the optimal Δt^*. It was observed that the overhead bit rate penalty decreases as operating frequency increases showing that even with more frequent weight updates, mm-wave frequencies can provide better performance in terms of data rate. Proposed method for update timing computation was simulated in a U2U communication system and results showed that beamforming gain rarely drops below the desired minimum threshold resulting in a stable connection.

Two algorithms, a reward based greedy algorithm and a deep reinforcement learning (DRL) algorithm implemented with a Deep Q-Network (DQN), were proposed for determining an optimal

trajectory for a UAV base station to maximize the coverage of distinct ground users. The performance of the two algorithms on static and non -stationary user distributions showed that the proposed methods are able to provide coverage to significantly more ground users with very high probability compared to the state-of-the-art alternatives as well as random action selection strategy.

The performance of a DRL based approach for learning the UAV Base Station (UAV-BS) coverage trajectory that maximizes the distinct ground users and the total data throughput was investigated for stationary and mobile user distributions in the presence of adversarial Sybil attacks. The attacks were assumed to be in the form of spawned fake users whose locations were also learnt through deep learning. It was observed that such attacks are capable of altering the trajectory of the UAV-BS drastically resulting in performance degradation. Furthermore, the UAV-BS expended significant resources in covering fake users when under a Sybil attack. The work investigated the impact of Sybil attacks during the training stage of DQN-implemented reinforcement learning (RL) algorithm.

REFERENCES

[1] Zada, M., Shah, I. A., and Yoo, H., "Integration of Sub-6-GHz and mm-Wave Bands With a Large Frequency Ratio for Future 5G MIMO Applications," in *IEEE Access*, Vol. 9, pp. 11241-11251, 2021, doi: 10.1109/ACCESS.2021.3051066.

[2] Cotuk, H., Tavli, B., Bicakci, K., and Akgun, M. B., "The Impact of Bandwidth Constraints on the Energy Consumption of Wireless Sensor Networks," *IEEE Wireless Communications and Networking Conference*, 2014 WCNC, 10.1109/WCNC.2014.6952870.

[3] Uwaechia, A. N. and Mahyuddin, N. M., "A Comprehensive Survey on Millimeter Wave Communications for Fifth-Generation Wireless Networks: Feasibility and Challenges," in *IEEE Access*, Vol. 8, pp. 62367-62414, 2020, doi: 10.1109/ACCESS.2020.2984204.

[4] Giordani, M. and Zorzi, M., "Satellite Communication at Millimeter Waves: A Key Enabler of the 6G Era," *2020 International Conference on Computing, Networking and Communications (ICNC)*, pp. 383-388, 2020, doi: 10.1109/ICNC47757.2020.9049651.

[5] Kodheli O., "Satellite Communications in the New Space Era: A Survey and Future Challenges," in *IEEE Communications Surveys & Tutorials*, Vol. 23, No. 1, pp. 70-109, First Quarter 2021, doi: 10.1109/COMST.2020.3028247.

[6] Saeed, O. Y. A., Saeed, A. A. A., Gaid, A. S. A., Aoun, A. M. H., and Sallam, A. A., "Multiband Microstrip Patch Antenna Operating at Five Distinct 5G mm-Wave Bands," *2021 International Conference of Technology, Science and Administration (ICTSA)*, pp. 1-5, 2021, doi: 10.1109/ICTSA52017.2021.9406521.

[7] Yang, Y., Zheng, B., Zhang, S., and Zhang, R., *Intelligent Reflecting Surface Meets OFDM: Protocol Design and Rate Maximization*, 2019.

[8] Zhao, M., Wu, Q., Zhao, M., and Zhang, R., *Exploiting Amplitude Control in Intelligent Reflecting Surface Aided Wireless Communication with Imperfect CSI*, 2020.

[9] Tang, W., Chen, M., Chen, X., Dai, J., Han, Y., Di Renzo, M., Zeng, Y., Jin, S., Cheng, Q., and Cui, T., *Wireless Communications with Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement*, 2019.

[10] Wang, Y., Lu, H., and Sun, H., "Channel Estimation in IRS-Enhanced mmWave System with Super-Resolution Network," in *IEEE Communications Letters*, doi: 10.1109/LCOMM.2021.3079322.

[11] Shen, H., Xu, W., Gong, S., Zhao, C., and Ng, D., "Beamforming Optimization for IRS-Aided Communications With Transceiver Hardware Impairments," in *IEEE Transactions on Communications*, Vol. 69, No. 2, pp. 1214-1227, Feb. 2021, doi: 10.1109/TCOMM.2020.3033575.

[12] Zhao, J., *A Survey of Intelligent Reflecting Surfaces (IRSs): Towards 6G Wireless Communication Networks with Massive MIMO 2.0*, 2019.

[13] Wu, Q. and Zhang, R., "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," in *IEEE Transactions on Wireless Communications*, Vol. 18, No. 11, pp. 5394-5409, Nov. 2019, doi: 10.1109/TWC.2019.2936025.

[14] Wu, Q., Zhang, S., Zheng, B., You, C., and Zhang, R., *Intelligent Reflecting Surface Aided Wireless Communications: A Tutorial*, 2020.

[15] Huang, K. and Wang, H., "Passive Beamforming for IRS Aided Wireless Networks," *IEEE Wireless Communications Letters*, Vol. 9, No. 12, pp. 2035-2039, Dec. 2020, doi: 10.1109/LWC.2020.3011596.

[16] Balanis, C., *Antenna Theory: Analysis and Design*. 3rd ed. Hoboken, NJ: John Wiley, pp 285-384, 2005.

[17] Xu, S., Liu, J., Cao, Y., Li, J., and Zhang, Y., "Intelligent Reflecting Surface Enabled Secure Cooperative Transmission for Satellite-Terrestrial Integrated Networks," *IEEE Transactions on Vehicular Technology*, Vol. 70, No. 2, pp. 2007-2011, Feb. 2021, doi: 10.1109/TVT.2021.3055965.

[18] Cai, W., Li, H., Li, M., and Liu, Q., "Practical Modeling and Beamforming for Intelligent Reflecting Surface Aided Wideband Systems," *IEEE Communications Letters*, Vol. 24, No. 7, pp. 1568-1571, July 2020, doi: 10.1109/LCOMM.2020.2987322.

[19] Özdogan, Ö., Björnson, E., and Larsson, E., "Intelligent Reflecting Surfaces: Physics, Propagation, and Pathloss Modeling," *IEEE Wireless Communications Letters*, Vol. 9, No. 5, pp. 581-585, May 2020, doi: 10.1109/LWC.2019.2960779.

[20] Stanczak, J., Kovacs, I., Koziol, D., Wigard, J., Amorim, R., and Nguyen, H., "Mobility Challenges for Unmanned Aerial Vehicles Connected to Cellular LTE Networks," *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, pp. 1-5, 2018, doi: 10.1109/VTCSpring.2018.8417736.

[21] Abeywickrama, H., He, Y., Dutkiewicz, E., Jayawickrama, B., and Mueck, M., "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, doi:10.1109/OJVT.2020.2971594.

[22] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M., "Playing Atari with Deep Reinforcement Learning," arXiv preprint arXiv:1312.5602, 2013.

[23] Andrychowicz, O., Baker, B., Chociej, M., Józefowicz, R., McGrew, B., Pachocki, J., Petron, A., Plappert, M., Powell, G., Ray, A., Schneider, J., Sidor, S., Tobin, J., Welinder, P., Weng, L., and Zaremba, W., "Learning dexterous in-hand manipulation," *International Journal of Robotics Research*, 2020, Vol, 39, No., 1, pp. 3–20, https://doi.org/10.1177/0278364919887447.

[24] Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., Quillen, D., Holly, E., Kalakrishnan, M., Vanhoucke, V., and Levine, S., "Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation," *Proceedings of The 2nd Conference on Robot Learning*, *Proceedings of Machine Learning Research*, 2018, Vol. 87, pp. 651-673, https://proceedings.mlr.press/v87/kalashnikov18a.html.

[25] Abeywickrama, H., He, Y., Dutkiewicz, E., Jayawickrama, B., and Mueck, M., "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, doi:10.1109/OJVT.2020.2971594.

[26] Jain, R., Chiu, D., and Hawe, W., "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems," *DEC Research Report TR-301*, 1984.

[27] Kurumbanshi, S., Rathkanthiwar, S., and Patil, S., "Packet Arrival Analysis Using Pareto and Exponential Distribution in Wireless Adhoc Networks," *2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*, Jalandhar, India, pp. 338–343, 2018.

[28] Xia, W., Polese, M., Mezzavilla, M., Loianno, G., Rangan, S., and Zorzi, M., "Millimeter Wave Remote UAV Control and Communications for Public Safety Scenarios," *16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Boston, MA, USA, pp. 1–7, June 2019

[29] Goldsmith, A., *Wireless Communications*, Cambridge University Press, 2005.

[30] Wang, X., Ma, R., and Xu, Y., "The Role of Data Cap in Optimal Two-Part Network Pricing," *IEEE/ACM Transactions on Networking*, Vol. 25, No. 6, pp. 3602–3615, 2017.

[31] Abeywickrama, H., Jayawickrama, B., He, Y., and Dutkiewicz, E., "Comprehensive Energy Consumption Model for Unmanned Aerial Vehicles, Based on Empirical Studies of Battery Performance," *IEEE Access*, Vol. 6, pp. 383–394, 2018.

[32] Bali, T., "The Generalized Extreme Value Distribution," *Economics Letters*, Vol. 79, No. 3, pp. 423–427, 2003.

[33] Simonsson A., "Beamforming Gain Measured on a 5G Testbed," *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1-5, 2017, doi: 10.1109/VTCSpring.2017.8108648.

[34] Bhandarkar, A. and Jayaweera, S., "Optimal Trajectory Learning for UAV-Mounted Mobile Base Stations using RL and Greedy Algorithms," *2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2021, pp. 13-18, doi: 10.1109/WiMob52687.2021.9606384.

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

3D	Three Dimensional
5G	Fifth Generation
AAM	Advanced Aerial Mobility
AOA	Angle of Arrival
AOD	Angle of Departure
AI	Artificial Intelligence
B5G	Beyond Fifth Generation
CCDF	Complimentary Cumulative Distribution Function
DoS	Denial of Service
DQN	Deep Q-Network
DQN-CNN	DQN Convolutional Neural Network
DQN-RNN	DQN Recurrent Neural Network
DRL	Deep Reinforcement Learning
EHF	Extremely High Frequency
ENU	East North Up
FD	Full Dimensional
FIFO	First In First Out
GEO	Geosynchronous Earth Orbit
GEV	Generalized Extreme Value
GHz	Giga Hertz (10^9 Hz)
GMM	Gaussian Mixture Model
HPBW	Half Power Beam Width
IRS	Intelligent Reflecting Surfaces
ISP	Internet Service Provider
LOS	Line of Sight
LTP	Local Tangent Plane
MAC	Medium Access Control

MILSATCOM	Military Satellite Communications
ML	Machine Learning
mmWave	Millimeter Wave
MRC	Maximal Ratio Combining
NIRS	Number of Intelligent Reflecting Surfaces
NN	Neural Networks
PDF	Probability Density Function
PHY	Physical Layer
RL	Reinforcement Learning
RNN	Recurrent Neural Network
Rx	Receive
SHF	Super High Frequency
SNR	Signal to Noise Ratio
Tx	Transmit
U2U	UAV to UAV
UAV	Unmanned Aerial Vehicle
UAV-BS	UAV Base Station
UHF	Ultra High Frequency
VHF	Very High Frequency
WiGig	Wireless Gigabit

DISTRIBUTION LIST

DTIC/OCP
8725 John J. Kingman Rd, Suite 0944
Ft Belvoir, VA 22060-6218 1 cy

AFRL/RVIL
Kirtland AFB, NM 87117-5776 1 cy

Official Record Copy
AFRL/RVB/Dr. Steven A. Lane 1 cy

This page is intentionally left blank.