NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

IMPROVING UXS NETWORK AVAILABILITY WITH ASYMMETRIC POLARIZED MIMO

by

Robert N. Severinghaus

June 2013

Dissertation Co-Supervisors: Murali Tummala
John McEachen

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704–0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE June 2013	3. REPORT TYPE AND DATES COVERED Dissertation

4. TITLE AND SUBTITLE: IMPROVING UXS NETWORK AVAILABILITY WITH ASYMMETRIC POLARIZED MIMO	5. FUNDING NUMBERS
6. AUTHOR(S) Robert N. Severinghaus	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number _____N/A_____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

To improve the efficiency of communications among unmanned systems, the research focused on the novel use of asymmetric polarized MIMO and network availability. The dissertation objective was to maintain the highest network availability for a mobile ad hoc network with heterogeneous communication capabilities. Using a hybrid dual-polarized Rayleigh fading channel model, asymmetric antenna configurations were studied in simulation for bit error rate and capacity. For a 1×1 configuration, polarization reciprocity was used to exploit the polarized channel knowledge, thereby maximizing received uplink power. The optimum gains to maximize uplink capacity were also derived for varying channel cross-polarization values. Larger configurations of 2×1 and 2×2 were investigated, including overlays of orthogonal space-time block coding, which improved diversity performance in the polarized channels. Extending these link results to realistic scenarios with unmanned systems, a reference point group mobility model including large-scale propagation was proposed to compute the network availability. Another scenario detailed robot exploration of unknown environments, which included large-scale path loss models. While deploying the network, the factors of exploration strategies, signal thresholds and routing were shown to impact the availability metric. Lastly, four extensible formation models were analyzed for their influence on network availability.

14. SUBJECT TERMS MIMO, polarization, unmanned systems, channel modeling	15. NUMBER OF PAGES 153
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540–01–280–5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

IMPROVING UXS NETWORK AVAILABILITY WITH ASYMMETRIC POLARIZED MIMO

Robert N. Severinghaus
Civilian, United States Navy
B.S., University of Kansas, 1989
M.S., Naval Postgraduate School, 1996

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 2013

Author: _____
Robert N. Severinghaus

Approved by:

_____ Murali Tummala Professor of Electrical and Computer Engineering Dissertation Committee Chair and Dissertation Co-Advisor	_____ John McEachen Professor of Electrical and Computer Engineering Dissertation Co-Advisor
_____ Xiaoping Yun Distinguished Professor of Electrical and Computer Engineering	_____ David Jenn Professor of Electrical and Computer Engineering
_____ Sherif Michael Professor, Space Systems Academic Group	

Approved by: _____
R. C. Robertson, Chair, Dept. of Electrical and Computer Engineering

Approved by: _____
Douglas Moses, Vice Provost for Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

To improve the efficiency of communications among unmanned systems, the research focused on the novel use of asymmetric polarized MIMO and network availability. The dissertation objective was to maintain the highest network availability for a mobile ad hoc network with heterogeneous communication capabilities. Using a hybrid dual-polarized Rayleigh fading channel model, asymmetric antenna configurations were studied in simulation for bit error rate and capacity. For a 1×1 configuration, polarization reciprocity was used to exploit the polarized channel knowledge, thereby maximizing received uplink power. The optimum gains to maximize uplink capacity were also derived for varying channel cross-polarization values. Larger configurations of 2×1 and 2×2 were investigated, including overlays of orthogonal space-time block coding, which improved diversity performance in the polarized channels. Extending these link results to realistic scenarios with unmanned systems, a reference point group mobility model including large-scale propagation was proposed to compute the network availability. Another scenario detailed robot exploration of unknown environments, which included large-scale path loss models. While deploying the network, the factors of exploration strategies, signal thresholds and routing were shown to impact the availability metric. Lastly, four extensible formation models were analyzed for their influence on network availability.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

A/D	analog to digital
AODV	ad-hoc on-demand distance vector
ARQ	automatic repeat request
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase shift keying
COST	Co-Operative for Scientific and Technical research
CR	cognitive radio
CSIT	channel state information at the transmitter
CSMA-CA	carrier sense multiple access collision avoidance
CTS	clear to send
dB	decibel
DFS	depth first search
DOA	direction of arrival
DoD	Department of Defense
EGC	equal gain combining
EP	equal power
GCS	ground control station
IEEE	Institute of Electrical and Electronic Engineers
JTRS	Joint Tactical Radio System
LMS	least mean square
LOS	line of sight
MANET	mobile ad-hoc network
MAC	medium access control
MCS	modulation coding scheme
MIMO	multiple-input multiple-output
MKS	meters, kilograms, seconds
MRC	maximal ratio combining
OFDMA	orthogonal frequency division multiple access
OSTBC	orthogonal space time block coding

PDF	probability density function
PHY	physical layer
PR	polarization reciprocity
QoS	quality of service
QPSK	quadrature phase shift keying
RF	radio frequency
RPGM	reference point group mobility
RTS	request to send
RV	random variable
SAE	Society of Automotive Engineers
SDR	software defined radio
SNR	signal to noise ratio
STC	space time coding
TCP/IP	transmission control protocol / Internet protocol
TDMA	time division multiple access
TEM	transverse electromagnetic
UAV	unmanned aerial vehicle
UGV	unmanned ground vehicle
U.S.	United States
UXS	unmanned system
WF	water filling
WLAN	wireless local area network
WSN	wireless sensor network
XML	extensible markup language
XPR	cross polarization ratio

EXECUTIVE SUMMARY

The use of unmanned systems (UXS) has increased greatly due to their success in combat operations. These systems relay their sensed information to their own ground control stations (GCS) and are not designed for interoperability with dissimilar UXS. This paradigm is changing because new capabilities are possible for teams of interoperable UXS. Specific to the communication challenge are the radio technologies, communication protocols, and the wide disparity of antenna configurations and power capabilities of transmitters. Adding to this are the unique formations and movements of UXS, which have a strong effect on network availability. This led to the dissertation objective: maintaining the highest network availability for a mobile ad hoc network (MANET) with heterogeneous capabilities.

Network availability, defined as the probability that there is a path between any two nodes in the network, is a critical metric of network performance. The computation of network availability grows exponentially with the number of edges in the graph representation. In this work, an efficient algorithm was developed to compute the availability, and it was used with many topologies throughout the research. The research showed how to translate the real-world probabilistic limits of the communication channel to graph theory representations that could use this availability algorithm. Because each link makes an uneven contribution to availability, the research narrowed to examine one of the latest technologies to improve these individual links.

Having evolved from smart antenna concepts in the last two decades, the tremendous capacity benefits of multiple-input multiple-output (MIMO) are already incorporated in the latest wireless local area network standards. The dissertation research extends MIMO techniques to polarized channels with asymmetric configurations. Typically viewed as a negative effect on signals, the depolarization that can occur in channels is exploited instead. Small UXS can utilize dual polarization antennas due to space constraints. The resulting asymmetric configuration, with single polarization antennas at one end and dual polarization at the other end of the channel, can take

advantage of the channel depolarization as a diversity path and capture orthogonal polarization power on downlink.

Extensive polarized channel models exist that split the effects of spatial correlation and polarization, but their normalization is insufficient for asymmetric configurations. In this effort, a hybrid, dual-polarized, Rayleigh fading channel model was developed. This hybrid model was then used in simulations to test a number of antenna configurations for bit error rate (BER) and capacity. For the simplest, a $1\times1\text{p}$ configuration, the novel use of polarization reciprocity (PR) could maximize the received uplink power to a single polarization antenna. The optimum uplink gain multiplier was also derived, showing improvements in ergodic capacity at a range of cross-polarization ratios, especially at high signal to noise ratio. These improvements were found using the same uplink power as a reference single polarization antenna. Improvements to diversity gain were found by the use of orthogonal space-time block coding (OSTBC). Two larger configurations ($2\times1\text{p}$ and $2\times2\text{p}$) were also tested, showing improved BER and capacity that were insensitive to channel depolarization. Notably, the $2\times1\text{p}$ configuration could be overlaid with OSTBC in two ways to provide BER performance that improved with or was insensitive to channel depolarization. Any of these configurations could be used to improve a single link, and the research next expanded to the network-level effects.

Many UXS analyses assume ideal communication or perfectly random motion of the network nodes. In contrast, this research proposed realistic scenarios for UXS and analyzed the network availability. The availability metric was first used with the reference point group mobility (RPGM) model that represented the clusterhead control topology used in ground UXS. It was computed by three methods (including shadowing) and considered the two data rates associated with control and sensor data typical of UXS. A fast approximation of the availability was also shown in the work. A second scenario detailed a robot exploration of five unknown indoor and outdoor environments, simulated with appropriate large-scale path loss models. A single exploring robot was faced with the task of movement beyond its communication range, so a fixed relay node deployment method was developed. The method was robust in the environments and deployed a tree topology that maintained adequate availability during the entire exploration. Comparisons

of the two exploration methods of depth first search (DFS) and frontier search showed that they placed relay nodes differently in the environments, even in non-convex areas. Quantitative results of these simulations slightly favored the frontier exploration strategy. The variable of signal threshold was shown to simulate different capability radios or high noise environments, with higher thresholds being more advantageous. Aggressive network routing was also investigated and showed a tradeoff between connectivity and network availability. In one case, there were 80 routing switches and an increase of 9.1% in the metric. The linear chain of active relay nodes in this network was a simple topology that was then extended for the third UXS scenario. This scenario studied the geometric formations that occur in MANETs as they move in concert to utilize sensors, provide overlapping protection, or transit an area. Four extensible formation models were analyzed for their influence on network availability. The results provided a concise answer to individual link availability required for network availability thresholds.

With individual communication links supported by asymmetric polarized MIMO and shown in realistic UXS scenarios, the objective of maintaining the highest network availability for a mobile ad hoc network with heterogeneous capabilities was met.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my God who made all of this work possible through His abiding love.

I would like to thank my wife, Christina, for her consistent support and the sacrifice of many hours together so I could complete this work.

And I would like to express my sincere appreciation to Prof. Murali Tummala for his mentoring and guidance that made the challenging dissertation process very rewarding. My committee members: Prof. Murali Tummala, Prof. John McEachen, Prof. David Jenn, Prof. Xiaoping Yun, and Prof. Sherif Michael have all provided significant insights and improvements to this dissertation, for which I am grateful.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Unmanned systems are coming into greater use in such wide-ranging commercial and scientific applications as driverless cars, household vacuums, urban search and rescue, and mine exploration. The United States (U.S.) military has benefitted over the past decade as similar individual systems have matured and become successful in military missions. However, these military systems exist independently and have yet to achieve the synergy from exchanging data among dissimilar systems.

A. MILITARY UNMANNED SYSTEMS

The U.S. military is entering the era of robots at war [1]. Unmanned systems (UXS) have an increasing role in the U.S. Department of Defense (DoD) due to their success in combat operations [2]. Typically used on a military mission that utilizes its sensors, these unmanned systems have transformed the way in which war is fought. From defusing a roadside bomb to dropping a precision-guided weapon on high value targets, these unmanned systems have become an integral part of the U.S. military. They can take on missions that are described as dull, dirty, or dangerous.

One type of UXS is an unmanned aircraft system (UAS), which is composed of a number of unmanned aerial vehicles (UAV), its datalinks, and a ground control station (GCS). Built for different missions, individual UAVs range in wingspan from a few inches to over 130 feet for a Global Hawk UAV [3]. In general, UAS have two main advantages over manned aircraft: they are considered more cost effective, and they minimize the risk to a pilot's life. As a result, there are over 7,500 UAVs in the DoD inventory today, and the spending on UAS was at the level of $3.3 billion in FY 2010 [3]. One of the most common UAS is the Army's Raven. Weighing four pounds, a Raven can fly on battery power for about one hour. Allowing its controller to "look over the hill," it relays real time video within line of sight (LOS) of a GCS [4].

Another type of UXS is an unmanned ground vehicle (UGV). UGVs are mostly wheeled or tracked and have cameras for real-time video and sometimes manipulators, such as multiple degrees-of-freedom arms. Approximately 8,000 UGVs of various types

have seen action in Operation Enduring Freedom and Operation Iraqi Freedom [2]. These deployed UGVs have been used in missions including suspected object identification, route clearance, and to locate and defuse improvised explosive devices (IEDs). During these counter-IED missions, Army, Navy, and Marine Corps explosive ordnance teams detected and defeated over 11,000 IEDs using UGVs [2]. One example of an UGV is the SUGV, a 30 pound UGV that moves on tracks. It is battery powered and can operate for 6 hours within range of its GCS [5].

Although not typically classified as an UXS, unattended ground sensors (UGS) are stationary sensors that report information to a control station. A system diagram would be very similar to other UXS, with multiple sensors and a control station with a radio receiver. Their use for sensitive surveillance is not generally documented in open literature, but they can use seismic, magnetic or acoustic sensors to detect movement, and then relay that data via radio frequency [6]. These low power military versions are mimicked by commercially available wireless sensor networks (WSN). An example is the National Instruments WSN-3202, which is a little larger than a chalkboard eraser. It is battery powered for up to 3 years and has a radio range of 300 meters to send analog data [7].

At the very least, it can be seen that these UXS are wide-ranging in their characteristics. There is a tremendous size difference ranging from a matchbox-size sensor to large vehicles and resulting power requirements. There are fixed UGS and mobile UGVs. There may be fields of hundreds of UGS, but only one large UAV operating in an area. However, they all can provide some persistent surveillance, lower operating costs, and reduced risk for soldiers. Another remarkable similarity is that these systems each have their individual control systems, sometimes referred to as "stovepiped" control systems. Having been acquired as a "system," which only communicated data to a control station, few of these systems are interoperable. Nevertheless, this interoperability is a key enabler for realizing new capabilities for unmanned systems. Teams of unmanned systems are possible with this interoperability. What is envisioned is a future with highly networked unmanned systems, as in Figure 1.

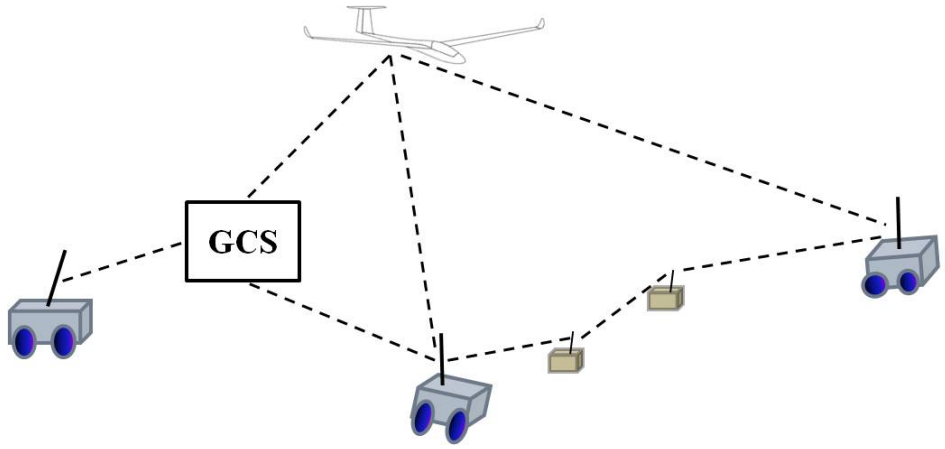


Figure 1. An Unmanned Systems network composed of UAS, UGVs, and UGSs.

B. UXS COMMUNICATIONS

There are a number of efforts to address the lack of interoperability among UXS. Since 2009, the DoD has been pursuing a common GCS with open architecture [8]. This would allow any GCS of this standard to operate a number of different UXS. A second advantage of a common GCS is reduced time for integration with other DoD command and control systems compared to proprietary GCS. But even if all UXS were controlled centrally, the network connections among these units would be undefined.

The future vision is not just communication of the UXS back to its control station, but communications to other unmanned (and manned) systems. Since 2011, Army network modernization efforts have labored to support a command capability by a robust transport layer, networking among many systems. These include command posts, air platforms, ground vehicles, dismounted soldiers, and unmanned sensors sending data, voice, imagery and video [9]. It is this sharing of data among the unique systems that is the focus of interoperability.

Another effort has focused on the narrower goal of UXS-UXS communication. The Joint Architecture for Unmanned Systems (JAUS) is a messaging architecture based on extensible markup language (XML). With JAUS working group members in the military, industry and academia, it was originally chartered in 1998, and it is now an

established set of standards maintained by the Society of Automotive Engineers (SAE) [10]. The JAUS encourages a strict partitioning of subsystems into software components with addressing similar to Internet protocol (IP) subnet addresses [11]. JAUS does not specify the transport mechanism, just the sets of messages. This enables a common reference "language" carried in message packets among all UXS and their components.

In calling for the next steps in improving UXS communications, the 2011 DoD UXS roadmap calls for the DoD to significantly improve efficiencies in communication transmission, transmitter and receiver designs, bandwidth usage and power usage [2]. The current efforts all have in common some standardization to achieve interoperability whether by interfaces, transport layer or message formats. But all of these utilize wireless communications for their effective interoperability. Although the message format is well defined, this dissertation will investigate what then happens to the signals in the wireless medium (physical layer) and how this affects the network of these dissimilar UXS. Considering that individual links are just a component of the network, a network-wide measure is also needed to measure the effect of link changes.

C. DISSERTATION OBJECTIVE

The improvement in the efficiency of communication among UXS could be measured in many ways. The specific measure we will use is network availability, which can encompass a number of factors. The objective of this dissertation is to investigate and develop schemes for *maintaining the highest network availability for a mobile ad hoc network (MANET) with heterogeneous communications capabilities.* The term MANET is the common scholarly term for a number of moving nodes forming a temporary network without fixed infrastructure and forms a common case for UXS. And, as discussed, the numerous different types of systems have dissimilar communication capabilities. It is this heterogeneous characteristic that makes networking even more difficult.

To improve availability the research will first investigate individual links in the wireless channel. The underlying channel property of polarization has been previously examined for potentially doubling spectral efficiency, but this does not apply when one

end of the link has only single polarization antennas. We propose the novel use of polarization reciprocity to maximize uplink power for asymmetric antenna schemes. These configurations are tested in simulation for both error rates and spectral efficiency.

Then considering the broader context of mobile nodes, some realistic situations for UXS are investigated. While most analyses use generic motion models, the unique formations and movements of UXS on military missions are examined. The effect on network availability when in realistic wireless propagation scenarios is computed. The results show what movements of UXS should and should not be conducted for maintaining network availability. When the mission requires movement that would certainly disconnect the network, we propose deployment methods for additional fixed nodes to maintain network availability.

D. RELATED WORK

Foschini and Gans' 1998 landmark paper [12] elicited interest in multiple element arrays, namely MIMO, for greatly increased channel capacity. Nabar's 2002 paper [13] was the first to use MIMO techniques on dual polarized antennas and derived estimates for error performance in different channel conditions. His conclusion was that polarization diversity generally incurred a performance loss for transmit diversity techniques (Alamouti coding). Coldrey [14] pointed out Nabar's error in power modeling of the channel and compared single polarization with dual polarization, noting the best performance of dual polarization was in Rician channels at high signal to noise ratio (SNR). Oestges [15] used an extensive channel model and concluded that dual polarization schemes are preferred when single polarization spatial correlations are high. All of these models used symmetric transmitters and receivers. But what is needed for analysis of the unique mechanism of polarization reciprocity is a hybrid polarized channel model. Our contribution will be this hybrid model and its use in asymmetric configurations.

The UXS that use the wireless communication channel are frequently in motion, complicating network performance analysis. Many simulations use mobility models that are variants of the random waypoint model [16]. This type of model is analytically

tractable and has been studied for effects on: link changes [17], clustering [18], network capacity [19], link lifetime [20], and link stability [21]. A de facto standard, this mobility model describes random, uncorrelated movements for a large network of nodes. The difficulty of using the random mobility model is that the mobility characteristics are very different from real-life scenarios. A realistic mobility model is important because different mobility models have an influence on the performance of the routing protocols [22], [23], [24]. To remedy this, more recent research has matched a large network mobility model to empirical data [25]. However, to meet the needs of the research in this dissertation, we provide analytical results for availability of a communication channel overlayed on a realistic mobility model.

Research related to UXS networks has also benefitted from the rich and diverse field of robotics (typically UGV). Specific to exploration, single robot strategies without heuristics have difficulty surpassing some simple algorithms, such as depth first search [26], [27]. Another major strategy proposed by Yamauchi [28] is frontier-based exploration, which moves the robot to the boundary between open space and uncharted territory. Many of these exploration algorithms can take the UGV beyond the range of its GCS and, therefore, require some form of communication relay. Typically, the WSN relay node placement is a discrete optimization problem that maximizes availability. There is substantial literature striving to achieve smaller polynomial time approximation algorithms for this availability problem [29], [30]. However, our proposal operates in the unknown environment in which optimization approaches are not applicable to deployment. Some prototype systems to solve this problem have been developed for robots [31] and first responders [32]. Their purpose is to reach maximum linear range by deploying nodes for relay communication. Our approach shows the effect of robot exploration strategies, signal thresholds and routing on the system-level availability metric.

E. ORGANIZATION OF THE DISSERTATION

The dissertation first investigates MANET availability and details a unique approach for individual links. The system-level effects are then considered in three realistic scenarios.

An introduction to network availability is discussed in Chapter II. Because of the probabilistic nature of the channel (shadowing, fading and polarization), each link in the network can be represented as a weighted edge in graph form. This MANET reduced to a graph is the input to the proposed algorithm that computes the availability. Using graph theory, three basic structures provide a framework for looking at various solution approaches to improve availability, as detailed in the following chapter.

There are a number of valid approaches to improving network availability, and these are discussed in Chapter III. An overview of other, fundamental solution approaches to the dissertation objective is presented. The various advantages and disadvantages of these fundamental solutions provide reasoning for the proposed approach of asymmetric polarized MIMO.

Using asymmetric polarized MIMO between two nodes is considered in Chapter IV. The unique characteristics of the polarized channel and its analysis led to the development of a hybrid channel model. This model is used to determine the performance of a number of asymmetric and reference antenna configurations. The concept of polarization reciprocity and its unique capability for maximizing uplink power is explored. Whereas this chapter examines the link between two nodes, the next chapter extends the availability concept to multiple nodes.

The entire MANET is considered in Chapter V. The network availability metric is first studied in relation to the movement of individual nodes. This is done within the context of a group mobility model with realistic shadowing effects, and the specific geometric analysis is explained in the Appendix. Next, the metric is used for a single exploring UXS in different environments, including realistic large-scale propagation

effects. A third scenario considers formations within a MANET. These three studies all use the graph theory relationships of Chapter II and could include the asymmetric polarized MIMO of Chapter IV for individual communication links.

II. BACKGROUND

To maintain high availability for the MANET, we must determine the source of wireless transmission errors. This prompts us to discuss propagation effects in the channel. Then these effects of the single link must be measured by the system-wide metric of availability.

We first consider a simplified model of the MANET. We can generically discuss each UXS in the subset of the total network as being composed of a layered communication protocol architecture as in Figure 2.

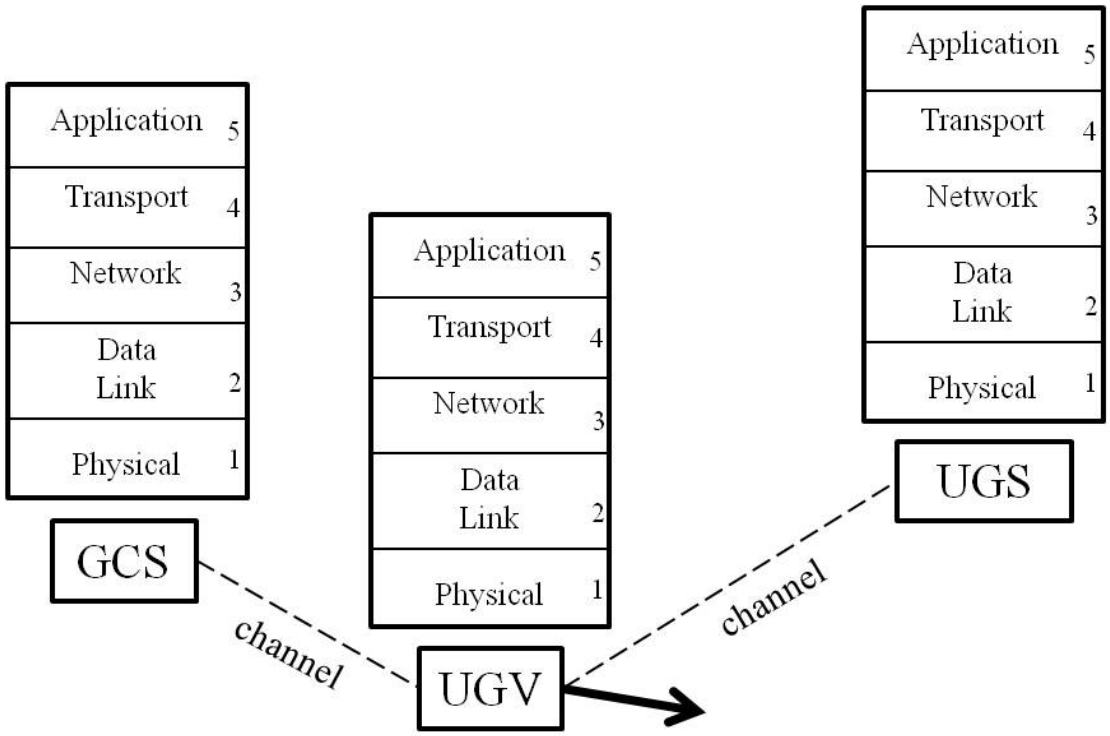


Figure 2. A MANET of three nodes shown with TCP/IP model layers.

The message packets among the network nodes are generated at the top layer of the protocol stack, proceeded down through the layers and eventually are transmitted through the wireless channel as symbols. With sufficient received signal strength, the transmitted 1 or 0 is received and properly decoded with a bit error rate (BER) dictated by a BER curve [33]. This BER is a function of signal strength (if noise is considered

constant in SNR). With deficient signal strength, the BER will be higher than a desired threshold. *It is this insufficient signal strength that is the source of bit errors*. Admittedly, the movement of the nodes in a MANET can change the received signal strength, changing the BER.

If there is a bit error, these are not always fatal to message packets. The impact of these bit errors can be minimized by a mechanism(s) within one of the layers, such as error correction coding or automatic repeat request (ARQ). We will also refer to some instances throughout the dissertation in which the solution to insufficient signal strength on one link is to find new routing mechanisms (layer 3). Another solution may be at layer 4 with a request for retransmission of the data. Acknowledging these favorable mechanisms, we consider these upper layer solutions as "last resort" and choose to focus on the channel and layers 1–2 interfaces.

A. WIRELESS SIGNAL PROPAGATION

We will next look at some of the channel effects that modify the transmitted signal while in the channel. These are a source of bit errors and are *a limiting factor of network availability*.

1. Large-scale Path Loss: Shadowing

The power received P_R in free space propagation by an antenna at a distance d from the transmitter is given by the Friis equation [34]

$$P_R(d) = \frac{P_T G_T G_R \lambda^2}{(4\pi)^2 d^2}. \quad (1)$$

The transmitted power is P_T, and the gains G_T and G_R are for the transmitter and receiver antennas respectively, and λ is the wavelength. The formula only applies to direct path or line of sight (LOS) between transmitter and receiver, which does not always occur in wireless communications. The system of units for this dissertation is MKS.

To account for the non-uniform obstructions and surroundings around a transmitter, an empirical path loss model with shadowing is commonly used. The model utilizes a log normal variable to represent the uneven signal strength received at the edge

of the range of coverage. The model that can be used for determining the received signal power for a chosen distance, is given by [35]

$$P_R(d)_{[dB]} = P_{T[dB]} - L(d_o)_{[dB]} - 10n_p \log\left(\frac{d}{d_o}\right) + X_{[dB]}. \quad (2)$$

In (2), $L(d_o)$ is the path loss at a reference distance in LOS, the value n_p is a path loss exponent particular to an environment, and X is a zero mean log normal RV. The random variable's distribution describes shadowing caused by objects in the wireless channel. The log normal nature is well-supported by experimental work such as [36], but if there is very little shadowing, a Rayleigh RV is more appropriate [37]. To give some perspective, Cox [38] measured $n_p = 4.5$ and the standard deviation of X as $\sigma = 10.1$ dB in a residential area.

2. Small-scale Path Loss: Fading

Another observed effect for channels is at a small scale, as the receiver moves only a short distance, such as a few wavelengths. It is caused by the multiple paths travelled by a broadcast signal in the medium and results in random amplitudes and phases at the receiver. But these combined fluctuations may add destructively or constructively at the receiver resulting in *fading* [35].

While there are propagation models that incorporate factors, such as time delays of each path and doppler spreading, we assume for this research that the channel exhibits slow, flat fading. These narrowband channels have an amplitude distribution that is Rayleigh [39].

3. Polarization

An often disregarded but an important mechanism of the channel is how it responds to different polarizations. The transmitted electromagnetic (EM) field from an antenna radiates outward and the radiated electric and magnetic field vectors ($\vec{E}, \vec{H}$) are treated as spherical waves in the far field. For a sufficiently distant observer, the radius of curvature of the spherical waves is so large that the phase fronts of the radiated fields are

approximately planar. These plane waves can be described by magnitude, phase, frequency and polarization. The polarization of any plane wave is described by the figure that the tip of the instantaneous electric field vector traces out with time at a fixed observation point.

This plane wave has an electric field ($\vec{E}$) and magnetic field ($\vec{H}$) that are perpendicular to each other and the direction of propagation; this is referred to as a transverse electromagnetic (TEM) wave, as shown in Figure 3.

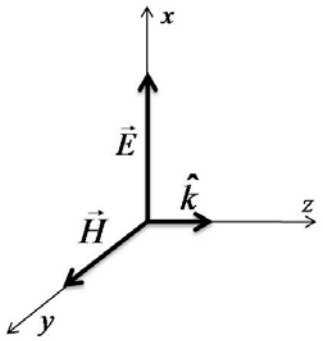


Figure 3. A TEM wave propagating in the $\hat{k}$ direction.

The instantaneous electric field intensity can be described by

$$\vec{E}(z,t) = E_o \cos(\omega t - kz + \phi)\ \hat{x} \tag{3}$$

and the instantaneous magnetic field intensity can be described by

$$\vec{H}(z,t) = \frac{E_o}{\eta} \cos(\omega t - kz + \phi)\ \hat{y} \tag{4}$$

where E_o is the magnitude, ω is the frequency, and $k = 2\pi / \lambda$. The values of ϕ are the same for $\vec{E}(z,t)$ and $\vec{H}(z,t)$, indicating they are in phase, which is typical for lossless material transmission.

Polarization, by convention, is only referenced to the electric field direction. (The magnetic field is always orthogonal to the electric field in isotropic media but generally ignored in polarization descriptions.) More generally, the electric field is not perfectly aligned with a chosen reference frame, nor will it keep the same orientation with time.

The polarization ellipse of Figure 4 describes all possible polarizations completely, with linear polarization as a degenerate case.

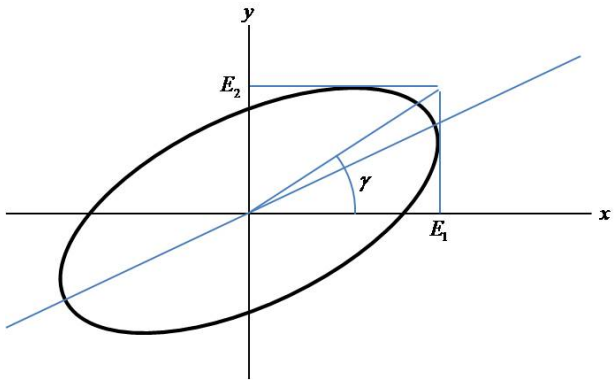


Figure 4. The polarization ellipse is described by three values. After [40].

The instantaneous electric field for the polarization ellipse can be written as

$$\vec{E}(z,t) = E_1\cos(\omega t - k\,z)\,\hat{x} + E_2\cos(\omega t - k\,z + \delta)\,\hat{y} \tag{5}$$

where δ is the phase angle lead of the *y*-component ahead of the *x*-component. The tip of the instantaneous electric field vector traces out the ellipse with time. At $z = 0$, (5) becomes

$$\vec{E}(t) = E_1\cos(\omega t)\,\hat{x} + E_2\cos(\omega t + \delta)\,\hat{y}\,. \tag{6}$$

If the components are in phase, δ =0, and the wave is linearly polarized at an angle determined by E_1 and E_2. If E_1=0, then the wave is vertically polarized (V-polarized). Similarly, if E_2=0, the wave is horizontally polarized (H-polarized). If the components are out of phase, various elliptical polarizations occur. Circular polarizations will result if $|E_1| = |E_2|$ and $\delta = \pm 90^\circ$. The different polarizations states can be fully described by δ and the angle

$$\gamma = \tan^{-1}\left(\frac{E_1}{E_2}\right). \tag{7}$$

In summary, the three wireless propagation effects of shadowing, fading and polarization exist in the channel and have a probabilistic effect on the transmitted

signal. These effects are for each link, but their impact on the entire network must be determined.

B. SOLUTION METHODS FOR SIMPLE NETWORK AVAILABILITY

A communications link between two radios has a unique probability associated with it, influenced by the propagation mechanisms described in the previous section. We can then structure a number of these links together to create a network. These individual links that now compose the network each have an effect on network-wide measures. Graph theory is a convenient tool that has been used to represent communications networks [41]. The representation of the physical radio in graphs allows the use of methods developed in graph theory to compute valuable network-wide measures, such as reliability [42], and solve problems, such as flows [43] and routing [44].

Basic graph models can be constructed and characterized with just a few definitions [45]:

- A graph $G = (V, E)$ consists of two sets V and E.
- The elements of V are called vertices (or nodes).
- The elements of E are called edges. Each edge has a set of one or two vertices associated to it, which are its endpoints.
- The degree of a vertex is the number of edges incident on the vertex.
- A weighted graph is a graph in which each edge is assigned a number, called its edge weight.

Assigning each radio as a vertex (node) and each communication link as an edge weight $p_a, (0 \le p_a \le 1)$ creates a weighted graph that represents a network. Assuming independence of edge weights, one can determine the probability that nodes A and B are connected theory [46], as in these simple networks of Figure 5. The expressions for availability between A and B with series edges and parallel edges, respectively, are

$$P_A = \prod_{i=1}^{j} p_{a_i}, \tag{8}$$

$$P_A = 1 - \prod_{i=1}^{j} (1 - p_{a_i}). \tag{9}$$

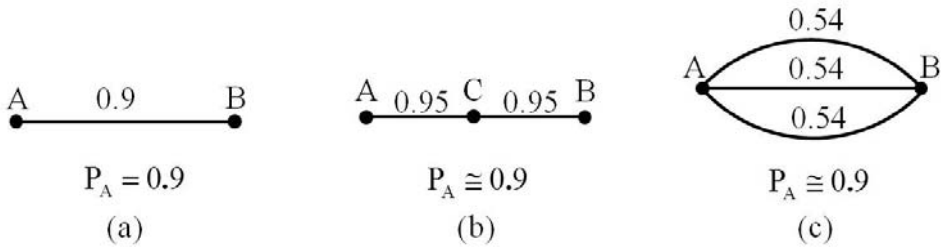


Figure 5. Graphs showing (a) single edge (b) series edges and (c) parallel edges.

With the goal of increasing network availability between A and B, there are three types of solutions. The first is demonstrated in Figure 5(a); the network availability can only be increased by improving the link availability. The second is by adding another node as in Fig. 5(b), the two resulting links need to have a product higher than the previous single link to improve network availability. "Weak" links are part of the calculation as in (8) and have a detrimental effect. The third type of solution creates diverse paths between the nodes. This diversity permits low availabilities of individual paths, but with the same net result on network availability. We will see in the next chapter that all the solutions to improving network's availability fall into one of these three categories.

There are certainly more complexities to modeling a network than the ideas just described. Additional nodes, such as node C in Figure 5(b) are placed optimally to create equal probability in the resulting edges, but this may not be the case in practice. Next, the edge weights can also be time-varying, so each of these graphs is only a snapshot of the network. Also, the edge weights are assumed to be bi-directional. If there was some asymmetry, (e.g., the transmitter power from one radio is much larger than the other), one must specify explicitly the direction of transmission that is being described. The alternative, using directed graphs with arrows on each link, is equally valid but more cumbersome in notation, so it is not used in this research. Acknowledging these complexities of the edge weights, we next show how they can be combined to provide a network-wide metric for networks with more than just two nodes.

C. NETWORK AVAILABILITY COMPUTATION

The network availability P_A is defined as the *probability that there exists a path between any two nodes in the network*, whether direct or through a number of connected nodes. Relating this to graph theory, of the many reliability metrics for a graph *G*, this has been referred to as the all-terminal reliability, $Rel_A(G)$ [47]. We concentrate on the all-terminal reliability because each node is an important communication and sensing entity. Each node is assumed to be capable of acting as a communications relay. This metric is used rather than the reliability between two nodes at the ends of a network, $Rel_2(G)$, to relay multi-hop messages [47].

There are a number of assumptions with any network metric, so here we explicitly state these assumptions. Each node is assumed to be capable of relaying messages. In a wireless network, as compared to a wired network, the link failures are more common, due to the unreliability of the channel. Therefore, for this analysis, we assume that the nodes are immune to failure and concentrate on the link failures. Link failures are also assumed to occur independently.

Next, we present an algorithm that can compute the network availability given any number of nodes and edge weights. Again representing each radio as a vertex and the communications link between radios as an edge, a mesh network of nodes is represented as an undirected graph, $G=(V,E)$. Each of the $|E|$ edges of G has a unique probability of availability p_a. An edge-cut in a graph G is a set of edges D such that $G-D$ has more components than G. The edge-connectivity of a connected graph G, denoted $\delta_e(G)$, is the minimum number of edges whose removal can disconnect G. Thus, if G is a connected graph, the edge connectivity $\delta_e(G)$ is the size of a smallest edge-cut. [45]

Complete state enumeration requires generating all $2^{|E|}$ states of G. For certain edges that are failed, a subgraph G' is formed. Each corresponding state has a probability of availability of

$$p_{state}=\prod_{i\in G'} p_{a_i} \prod_{i\notin G'} (1-p_{a_i}). \tag{10}$$

The network probability of availability, P_A, is the summation of all p_{state} that represent a connected subgraph G'. For a number of edge failures, n, this is given by

$$P_A = \sum_{n=0}^{|E|} \binom{|E|}{n} c_{n,e}\, p_{state} \tag{11}$$

where $c_{n,e} = 1$ or 0 determined by the state's G' being connected or not connected, respectively.

Classically, p_a for all edges is assumed identical to p [47], not unique as above. In this case, if N_i is the number of connected subgraphs G', (10) and (11) can be simplified to:

$$P_A = \sum_{i=0}^{|E|} N_i\, p^i\, (1-p)^{|E|-i}. \tag{12}$$

For computation of P_A, a connected, undirected, planar graph G is fully described by an "edge list" of: from vertex v, to vertex w and the p_a of the edge joining v and w. The algorithm of Figure 6 can process the "edge list" in any order and only requires having $0 < p_a \leq 1$.

```
1:  pa(0 fails) ← ∏ p_a
2:  for #fails 1: |E| - |V| -1 do
3:    compute edge failure combinations
4:    for each edge failure combination do
5:      G' ← (edge list – failed edges)
6:      connected← false
7:      changes←true
8:      L ←{v1}
9:      while (not connected) and (changes=true) do
10:       changes←false
11:       for all v,w in G' do
12:         if vi or wi in L then
13:           L ←append(vi,wi)
14:           changes←true
15:           (vi,wi) mark for deletion from G'
16:         end if
17:       end for
18:       if L contains all vi in G'
19:         connected←true
20:         pa(#fails) ← pa(#fails) + state probability
21:       end if
22:       if (connected=false) and (changes=true) do
23:         G' ←remove all vi,wi marked for deletion
24:       end if
25:     end while
26:   end for
27:  end for
28:  P_A←∑ pa(#fails)
```

Figure 6. The algorithm pseudocode for computing the network availability P_A. From [48].

The algorithm progresses through the number of edge failures from zero to $|E|-|V|-1$, which is the limit beyond which the graph is certainly disconnected. This is equivalent to constraining the upper limit of the summation in (11). For each state, a graph G' is formed consisting of the "edge list" minus the failed edges. The "edge list" is used to compile a new list of vertices connected to v_1, namely L. A modified depth first search (DFS) through G' occurs until the G' is established as connected or not connected. If connected, the polynomial for that specific failure condition of edges is

computed and added in to the case of (#fails) edges. The algorithm returns the network availability P_A , which is a sum of the (#fails) polynomials.

This details of wireless channel propagation were examined in this chapter. These factors are what cause the probabilistic nature of the wireless channel. The effects of each communication link will limit the network availability, and the total effect is found by the exact computation of the network availability. Equipped with this metric, we next examine what fundamental solutions can improve network availability.

THIS PAGE INTENTIONALLY LEFT BLANK

III. PROBLEM STATEMENT AND SOLUTION APPROACH

The MANET formed from the various UXS is first translated to a graph, and we discuss the effects of the MANET on the graph representation. At the abstract level of a graph, we examine the feasible solutions to improve network availability. Lastly, the selected approach is presented.

A. UXS NETWORK GRAPH

To study the complex problem of these networked communications, the graph theory context can abstract away from the UXS network diagram of Figure 1 to a graph [45]. If we replace each radio from Figure 1 with a lettered vertex and each wireless communication path with an edge, the created graph is in Figure 7. The edge weights signify a probabilistic measure associated with the corresponding communication link. Edges can have different weights due to the different node capabilities and the different channel conditions between the connected pairs of nodes.

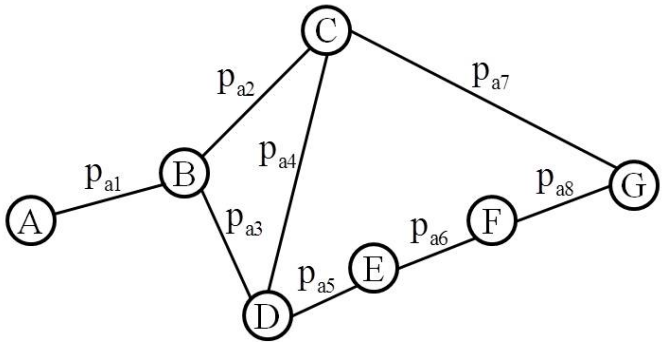


Figure 7. The UXS network diagram represented as a graph.

Because the graph can be a combination of fixed and mobile vertices, it can change edge weights with time. The edges may also reconfigure, creating a new topology. Nevertheless, conditions that reduce p_a, such as the movement of two nodes away from each other and/or a more difficult channel, will reduce P_A. The magnitude of this reduction in P_A in proportion to the reduction in p_a is not directly obvious except for simple networks, such as in Section II. C. The effects of each reduced p_a on P_A are not

equal, depending on their position and connections in the network, and require an algorithm, such as Figure 6, to determine their effect on P_A. As a result, it is desired to keep individual edge weights as high as possible. As a result, the network-wide metric of P_A may fall below some threshold that is needed for mission accomplishment.

In the context of these observations, we repeat the dissertation objective as: *maintaining the highest network availability for a mobile ad hoc network (MANET) with heterogeneous communications capabilities.* We have used the widely acknowledged term MANET, but there may be fixed nodes in the network, such as WSN. Our treatment considers having fixed nodes as a special case of the more general MANET.

B. SOLUTION APPROACHES

Because networking includes a combination of elements across the TCP/IP protocol stack, the solutions to maintaining availability also range across these layers. Hence there is not "one solution" to maintain high network availability. The fundamental solutions outlined here are a sampling, not an exhaustive list. Among these fundamental solutions, we observe that each can have an influence among the protocol stack layers (vertically) and across the deployed network (horizontally). The vertical influence of these solutions could be used to increase performance in cross-layer designs [49]. To briefly outline these fundamental solutions, we define them and then discuss their advantages and disadvantages in the following sections.

1. Fundamental Solutions

a. Increased Number of Nodes

One way to increase link availability is to insert more wireless nodes to create a multi-hop link. The new communication path made by the additional nodes may replace the existing link by relay as in Figure 5(b), when two nodes are moving away from each other. The new path might also supplement the existing link for redundancy. In both cases, the goal of the placement of these intermediary nodes is improved availability.

The advantage of more nodes is that in aggregate, they can bridge between or extend existing networks. Two networks that become disconnected due to depleted nodes can be bridged by fixed node placement, restoring or improving availability [48]. For a moving node that goes into an inaccessible area, deploying fixed nodes can extend a multi-hop relay in a linear topology to maintain availability [50].

Beyond the obvious increased costs to the system owner and the necessity of a means of deployment, additional nodes have vertical and horizontal impacts. The medium is a shared resource, so additional nodes needing access are controlled by the MAC layer. For example, the 802.11 MAC layer primarily employs the carrier-sense multiple access collision avoidance (CSMA-CA) scheme with reliability enhancement through the use of Request to Send (RTS) and Clear to Send (CTS) control frames [51]. More nodes can create the hidden terminal problem, which occurs when two nodes (that cannot communicate with each other) attempt to simultaneously transmit to a common node. The use of RTS/CTS not only solves the hidden terminal problem, but also reduces the collision period when the traffic is heavy or the number of nodes is large [52].

One might also think that the total throughput of the network would increase with more nodes, but this is not the case due to increased competition for a shared finite resource. The network throughput is defined as the fraction of time the channel is used to successfully transmit payload bits. For the 802.11 network described above, this throughput is steady, almost insensitive to the number of nodes in the network (with a suitably chosen contention window size) [53]. From a theoretical derivation [54], if there are n identical nodes placed optimally in a disk of unit area, the throughput capacity for each node is of order $\Theta\left(\frac{\mathrm{B}}{\sqrt{n}}\right)$. This assumes each node is capable of transmitting B bits/second over a fixed range. These results show steady or decreasing network throughput with an increasing number of nodes.

Neighbor discovery (layer 2) and routing (layer 3) certainly take more time with an increasing number of nodes. Searching for and maintaining routes among MANETs requires a scalability to this larger number of nodes. In table-driven protocols,

this routing information must be exchanged with each update of network size and/or topology change, leading to a network traffic burden. Reactive protocols such as ad hoc on-demand distance vector (AODV) routing, [55] might be used, but this has a longer delay in routing than a table-driven protocol.

Lastly, the addition of more nodes may require topology adjustment; for instance, when a node has many neighbors close together. To reduce the overhead of maintaining too many neighbors, the ideas of mobile topology management have been proposed. Such algorithms seek to change the topology to achieve lower node degree and low hop stretch in which the changes are based on factors, such as distance, clustering, degrees of angular separation, number of antenna sectors, and two-hop neighborhoods [56]. In summary, while the increase in the number of nodes seems a simple way to increase availability, there are numerous effects across the protocol stack and no increase in network throughput.

b. Increased Transmit Power

Instead of more nodes, simply increasing the transmit power of each node has a *linear* effect on the power received as seen in the Friis equation (1), repeated here for convenience,

$$P_R(d) = \frac{P_T G_T G_R \lambda^2}{(4\pi)^2 d^2} \tag{13}$$

This could be a simple solution within the designed power limits of a transmitter. In a MANET, with two nodes moving away from each other, increasing transmit power can certainly be used to offset increasing distance d to keep the received power the same, but it must increase as the *square* of the distance. Assuming a fixed distance, the gain in channel capacity C with received signal power P_R is not linear but *logarithmic* in the presence of additive white Gaussian noise (AWGN) as given by [57]

$$C = W \log_2\left(1 + \frac{P_R}{N}\right) \tag{14}$$

where we have used W as the fixed channel bandwidth and N as the average noise power. Together, these two equations point out that the doubling of transmit power can improve but does not double the data rate at a fixed distance.

The transmit power changes must also consider the power resources of the node. Fixed nodes, such as cellular base stations, have connection to a power grid, essentially an unlimited power source. But small fixed nodes, such as those that comprise a wireless sensor network (WSN), have small batteries. The commercially available MICA2DOT node only transmits between 0 and 10 dBm, but using a battery-aware approach to selecting transmit power level increased useful life by 52% [58]. Mobile nodes, such as small military UGVs that are only battery powered, fall between these two extremes. These small UGVs use more energy on movement than communications but are still limited by battery life [59].

Increased transmit power can also increase interference. By way of analogy, consider a crowded gathering of people who are conversing. If one pair starts talking louder, the increased noise interferes with other conversations. Similarly, transmit power increases must also consider their interference effect on the rest of the network. This effect has been noted in dense cellular systems that are interference limited. The increase in all users' transmit power has no net benefit due to the mutual interference [60]. This problem exists in MANETs also. To solve this, one cross-layer design (layer 1 and 2) first centrally schedules nodes' transmissions to reduce interference, then uses a distributed algorithm for transmit power control [61]. Research has also yielded distributed layer 3 algorithms that control the transmit power of network nodes, balancing network connectivity and interference [62].

The transmit power increase can also be detected by others outside the network. Specifically, the intelligent adjustment of transmit power can address a military wireless network security concern: the low probability of intercept (LPI) by reducing the detectability of the network links by others [63]. In total, the seemingly simple increase of transmit power has positive effects on received power and capacity but potentially negative effects on battery depletion and interference with other nodes in the network that

must be managed. Having assumed omni-directional antennas to this point, we next look at directional antennas and their effects.

c. Directional Antennas

Increasing the gain of the transmitter and/or receiver antenna will increase received power, as shown in the Friis equation (13). Alternatively, as the distance increases, increasing the antenna gains can keep the received signal above a certain threshold. Another benefit is that directional antennas concentrate energy in a certain direction, thereby reducing the interference they cause and the interference they receive. Together, these effects at the physical layer can increase the spatial concurrency of transmissions [54].

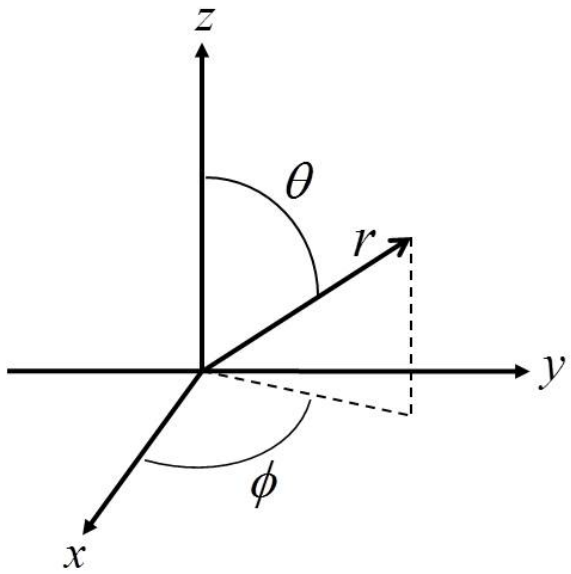


Figure 8. A vector in the spherical coordinate system.

The well-studied half-wave dipole is a basic antenna. A z-directed dipole has the familiar "doughnut" omnidirectional radiation pattern that is only a function of the zenith angle θ ; it radiates equally in all azimuth ϕ directions, and the normalized field pattern is [40]

$$g(\theta) = \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \tag{15}$$

The gain of the half-wave dipole (assuming a perfect radiation efficiency of 1) is 1.64 [40]. This gain is with reference to a hypothetical isotropic radiating antenna, which would radiate equally in all directions. Also, the stated antenna gain is the maximum value, which for the half-wave dipole is met when $\theta = \frac{\pi}{2}$; in other directions, the gain is lower.

The benefit of increased gain from directional antennas is proportional to physical size, so it may be difficult to use directional antennas on small nodes. From [34], the gain G is related to the physical area A_p of the antenna by the efficiency of the aperture ($0 < \varepsilon_{ap} < 1$) as given by

$$G = \frac{4\pi\varepsilon_{ap}}{\lambda^2} A_p . \tag{16}$$

For example, uniformly illuminated aperture antennas, such as horns and reflectors, have $\varepsilon_{ap} \approx 1$, so gains are produced by increasing the physical area of the antenna [40]. The end result is that antenna gains can be limited by the area available for the antenna on the node.

One network problem with implementing directional antennas is the deafness of some receivers as they miss a transmission when they are listening in another direction. Another problem is slow neighbor discovery with all directional antennas. The resulting MAC layer solutions include scheduling [64], using directional RTS/CTS [65], and using control channel busy tones [66]. There are many variations and combinations of these approaches, and relevant surveys of directional antenna MAC protocols are available in [67] and [68]. In summary, the directional antenna can improve received power if there is available physical space, and it has network-wide impacts.

d. Software Defined Radio

Potentially, the connectivity of a network could be improved through the digital processing of received signals. First termed in the 1990s [69], the phrase software-defined radio (SDR) had the vision to change all previously hardware (analog) functions of a radio to software (digital) functions, as shown in Figure 9.

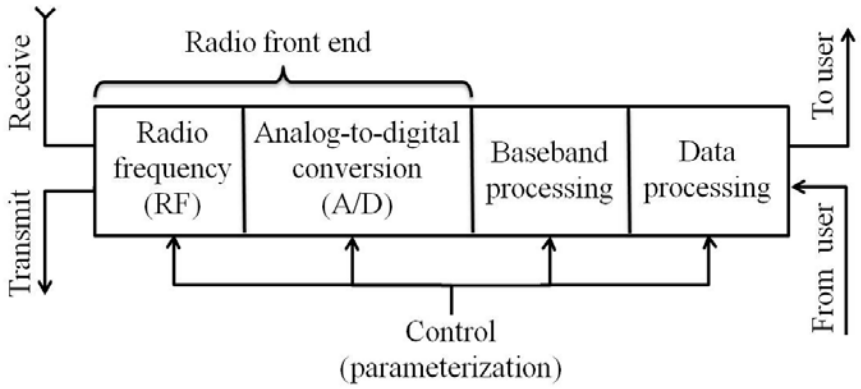


Figure 9. The major components of a SDR transceiver. From [70].

This was idealistically seen as digitizing at the antenna with wideband A/D converters. The practical limits of such conversion and processing requirements resulted in a more practical definition as the reconfigurability at any level of the radio protocol stack with some signal processing done in the RF circuitry [71]. Even with this technology limitation, the great improvement in SDR is the flexibility and interoperability from operating across multiple bands and multiple standards [70]. It is precisely these qualities that the DoD wanted when it created the Joint Tactical Radio System (JTRS) program. One of the JTRS goals was to create a family of interoperable radios with portability of software among many hardware platforms [72]. Besides the military, the reconfigurability by software download has been used in commercial cellular base stations, supporting multiple radio standards on the same platform [73].

The primary advantage of SDR is the use of constantly improving digital performance (power usage and clock speeds) of evolutionary devices, such as CPUs and FPGAs. One such radio prototype operates from 0.4 to 6 GHz with three different waveforms (802.11A, 802.11b, and W-CDMA) [74]. A military radio solution used a VHF/UHF SDR with an external frequency translation and power amplifier to reach the lower frequency HF band [75]. This type of solution still maximizes a SDR to not replicate additional hardware. The RF units of SDRs may switch among multiple antennas or use broadband antennas with bandpass circuits. Although impedance matching between the transmitter and antenna attempts to maximize transmitted power,

the gain of the antenna may change as the frequency changes due to the antenna radiation efficiency.

Improvements of digital technology may not keep up with the higher data rate demands, new waveforms and smaller radios. Questioning if the current scaling approach of processing is sufficient, some researchers propose fundamental changes to the DSP architecture allowing for the needs of mobile computing such as minimal power usage [76]. Regardless, the overwhelming advantage of a digital, software-defined radio is the flexibility of reconfiguring to become more than one radio with just one radio.

e. Cognitive Radio (CR)

The cognitive radio (CR) is built on the basis of a SDR, but in addition it is "*aware*" of its environment. This means that it can *adapt its operating parameters* (e.g., transmit power, frequency and modulation) in real time in response to the incoming RF signal [77]. Of necessity, to adjust the transmit power, the CR needs the feedback from the receiver. This idea, called channel state information at the transmitter (CSIT) is covered in more detail in Chapter IV.

Cognitive radio may also adjust operating bandwidth or frequency to provide network connectivity. Having enough spectrum is an important consideration for maintaining network availability, yet average spectrum usage might be low. One study measured only 13% spectrum usage in an urban, high communication area [78]. One way CR can increase spectrum usage is by the coexistence of opportunistic users with incumbent users, even in a licensed band. To do this, the CR must sample the spectrum to become aware of used and unused (white) spectrum. An example is the IEEE 802.22 standard, which opportunistically uses licensed television spectrum to provide broadband access to rural areas [79]. The sensing is enabled by many receivers that monitor the spectrum and centrally report results to geographically dispersed base stations. The distributed receivers sense the spectrum both in-band and out-of-band to detect any incumbent users, thereby preventing interference [80].

The change of modulation in response to the environment is another CR characteristic. The widely used 802.11a standard has eight modulation and coding

schemes (MCS) at the PHY level [51]. The listed sensitivity values in Table 1 are the minimum signal levels as measured at the antenna connector of the receiver, and these minimum signal levels ensure a certain maximum packet error rate [51].

Table 1. IEEE 802.11a PHY showing data rate and signal levels. From [51]

MCS	Modulation	Code Rate	Rate (Mbit/sec)	Min. Sensitivity (dBm)
7	64-QAM	3/4	54	-65
6	64-QAM	2/3	48	-66
5	16-QAM	3/4	36	-70
4	16-QAM	1/2	24	-74
3	QPSK	3/4	18	-77
2	QPSK	1/2	12	-79
1	BPSK	3/4	9	-81
0	BPSK	1/2	6	-82

The standard only provides the MCS choices, so an algorithm must decide which MCS to use for the next transmission. One highly cited algorithm uses a combination of the data payload length, the number of retransmission attempts and an estimate of the receiver's SNR value (environment estimate) to determine the next PHY transmission mode [81]. Even more sophisticated, a change in modulation and bandwidth can be accommodated by a managed MAC scheme, such as orthogonal frequency division (OFDMA), as in the 802.22 standard [79].

These advantages of CR come at the cost of complexity. A difficult design problem for the CR is the multi-gigahertz wide bandwidth RF processing coupled with the reliable detection of primary users [82]. To reliably perform this sensing, there must be coordination, such as quiet times for in-band sensing, when the network of opportunistic users are not transmitting [79]. This can cause a possible reduction of the Quality of Service (QoS). Another challenge with this coordination is also that as the channel conditions change, the network scheme must change, potentially leading to capacity reduction due to small coherency time [82]. Overall, the advantages of CR are worth the complexity as evidenced by their incorporation in published standards.

f. Arrays

The modern application of multiple-input multiple-output (MIMO) schemes for communications has its genesis in the mature fields of antenna arrays and direction finding algorithms. We briefly review these ideas and then present MIMO in a broader context.

Several antennas operating together defines an array. The advantage is the ability to produce a highly directional radiation pattern using a collection of small antennas rather than a single large antenna. The radiation pattern of the array depends on the number and type of antennas used, the spacing in the array, and the relative phase and magnitude of the excitation currents [83].

A well-studied and still relevant example is the uniformly excited, linear, equal spaced array. Referring to Figure 10, we define the variable ψ as

$$\psi = k\,\mathrm{d_a}\cos(\theta)+\nu \tag{17}$$

where $\mathrm{d_a}$ is the distance between the elements, and θ is with reference to the line of N identical elements (the array axis). The value ν is the current phase by which the n+1th element leads the nth element. This array can then be shown to have a normalized array factor of [40]

$$\mathrm{f}(\psi) = \frac{\sin(\mathrm{N}\psi/2)}{\mathrm{N}\sin(\psi/2)}. \tag{18}$$

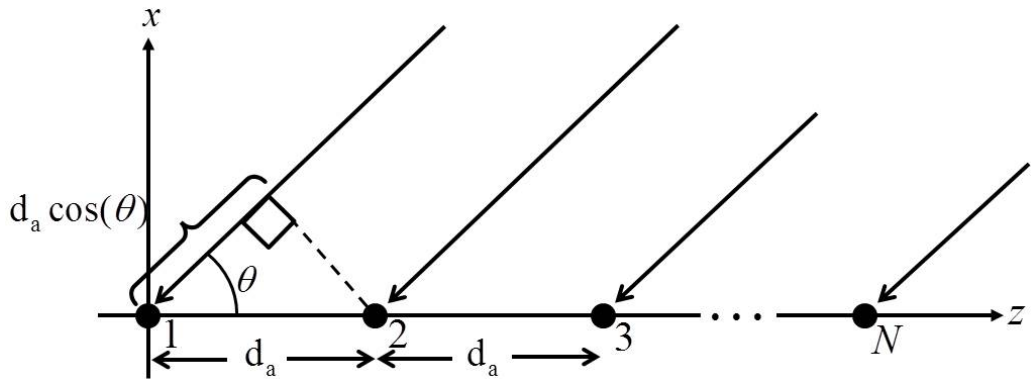


Figure 10. An equally spaced linear array of N elements. After [40].

What is observed from (18) is that as N increases, the main lobe becomes narrower, and this becomes an increase in gain. Another observation of (17) and (18) is that the value of the phase shift ν can be changed and have an effect on the array factor. The maximum of the array factor occurs at $\psi = 0$, so changes in ν create changes in the direction (θ) of the main lobe. This ability to move the radiation pattern in space is called a *phased array*.

The total radiation pattern in space of the phased array is found by the principle of pattern multiplication. The complete, normalized pattern of the antenna array is the product of the element pattern (assumed identical for all N) and the array factor, [40]

$$\mathrm{F}(\theta,\phi)=g(\theta,\phi)\ \mathrm{f}(\theta,\phi)\,, \tag{19}$$

where the normalized array factor (18) is converted to a function of θ and ϕ by the non-linear transformation of (17).

What are the drawbacks to using arrays? An array requires a network of feeds for the number of antennas. Also, each feed must manipulate the phase and amplitude of the signals sent to the individual antennas, which is possible with digital signal processing. Beyond this added complexity, there is another drawback of antennas operating near each other, namely mutual coupling. Mutual coupling is the interaction between antenna elements that changes the current magnitude, phase and distribution on each element [40]. This mutual coupling becomes significant as inter-element spacing is decreased. For example, for a fixed length of an array with mutual coupling, the array output SINR can decrease when more elements are added [84]. The general guideline for minimizing mutual coupling for similarly polarized antennas is to keep them at least a half wavelength apart [85]. The (presumed negative) change in performance of real world arrays due to mutual coupling can be approximated by substituting an average active-element pattern for the element pattern in (19) as [40]:

$$\mathrm{F}(\theta,\phi)=g_{ae}(\theta,\phi)\ \mathrm{f}(\theta,\phi)\,. \tag{20}$$

An array's usefulness can be increased with additional signal processing. With the added computational processing, the arrays just discussed can become a "smart antenna" or "adaptive array" [85], as illustrated in Figure 11, which shows a representative array pattern. The adaptation occurs as the antenna determines the direction of desired and undesired signals and shapes its pattern accordingly.

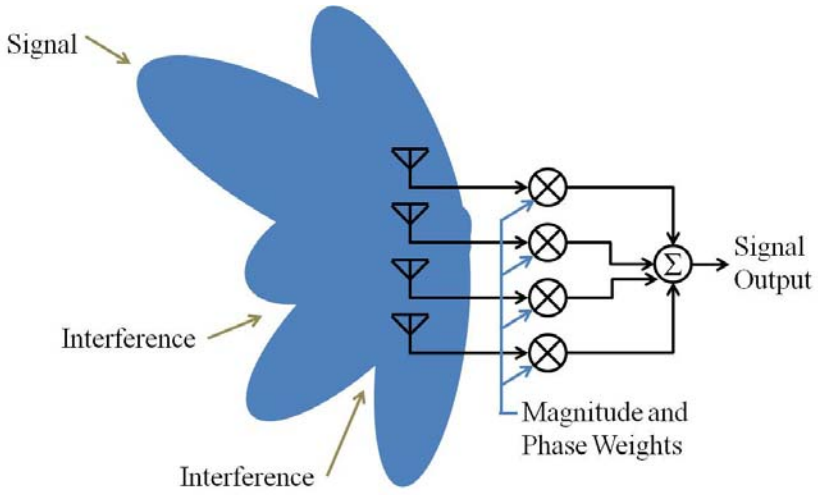


Figure 11. A smart antenna varies its magnitude and phase weights to shape its pattern. After [85].

First the antenna array receives all the signals from all directions, and a direction of arrival (DOA) algorithm determines the direction of all impinged signals based on the time delays or phase shifts implicitly supplied by the antenna array. An algorithm (such as the ESPRIT algorithm [86]) stacks the received data vectors and computes a basis for the signal subspace. Then a least mean squares solution (LMS) of two matrix equations is computed (size proportional to the number of array elements), and the azimuth and elevation are found from the eigenvalues of these matrix solutions. The output of the DOA algorithm (the θ and ϕ of the desired signals and interfering signals) is fed to an adaptive algorithm to compute the magnitude and phase weights, thereby placing array maximums toward desired directions and nulls toward interfering signals. The LMS adaptive algorithm has adequate convergence and simple complexity [85].

We have now seen that arrays, such as the linear phased array, can steer their patterns, and the more capable adaptive arrays can reshape their patterns to also minimize interfering signals.

g. MIMO

In contrast to the one input or output signal used in arrays, MIMO is a wireless communications system with multiple antennas on the transmitter and receiver that can greatly increase channel capacity. Remarkably, it does not change the resources of transmit power and channel bandwidth, but increases the computational effort instead [39]. The capacity benefit has been so dramatic that it is incorporated into the latest WLAN standard [87].

The MIMO channel is made up of the many subchannels between the transmitter and receiver antennas, as in Figure 12. This is represented by an N_r by N_t matrix

$$H=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & \ddots & & \\ \vdots & & & \vdots \\ h_{N_r1} & & \cdots & h_{N_rN_t} \end{bmatrix}, \tag{21}$$

where N_r is the number of receivers, N_t is the number of transmitters, and each subchannel is a complex number h_{mn} between the mth receive and the nth transmit antenna pair.

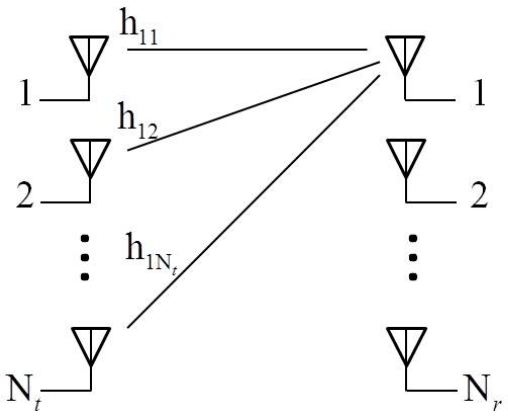


Figure 12. The MIMO channel diagram showing example subchannels that make up one row of the MIMO channel matrix. After [88].

Consequently, the classic Shannon capacity equation [57] in scalar form is extended to a matrix form as given by [12], [89]

$$C=E\{\log_2 (\det(I_{N_r}+\frac{\rho}{N_t}HH^{\dagger}))\} \text{ bits/s/Hz}, \quad (22)$$

where $E\{\bullet\}$ is the expectation operator and I_{N_r} is the N_r by N_r identity matrix. The average SNR at the receiver input is ρ [12]. The complex values that represent each subchannel (h) are constantly changing, making H a matrix of random variables. The resulting capacity would also be a random variable, but an average of the capacities achieved with each channel realization is summarized by the ergodic capacity, as in (22). This is the rate at which information can be transmitted assuming no CSIT [90]. The capacity could be increased further through an optimization if there is CSIT such that the transmitter adjusts signals to use stronger subchannels wisely, (e.g., a water-filling algorithm), which we discuss in Chapter IV.

There are four generally stated benefits to MIMO [88, 91]:

- Array gain. The coherent combining of signals at the receiver. This is similar to the adaptive array, which performs spatial processing at the receiver. There can also be spatial pre-coding at the transmitter assuming CSIT. So, the array gain is at the transmitter or receiver (or both) and requires knowledge of the channel.

- Interference reduction. Also similar to the adaptive array, array weights can be chosen to place nulls in the direction of undesired signals, improving the signal-to-interference noise ratio (SINR).
- Spatial multiplexing gain. The multiplexing gain is achieved by putting different data streams onto many parallel subchannels, which increases channel capacity (bits/sec). This offers a capacity gain on the order $min(N_t, N_r)$. This capacity gain is highly dependent on the spatial correlation.
- Spatial diversity gain. Sometimes called diversity gain, it is the improvement in reliability because of multiple signal paths. The multiple, hopefully independent paths mitigate channel fading. The spatial diversity order is equal to the number of independently fading links, which ideally is $N_t N_r$. The effect of increased diversity gain is seen in reduced BERs.

Not all of these advantages are achievable at all times, and there is a flexible tradeoff between spatial multiplexing gain and spatial diversity gain. As the number of antennas increases, the multiplexing gain remains linear with the number of antennas, but the diversity gain gives diminishing returns [92]. This research concentrates on array gain and spatial diversity gain, of which the latter is helped by some unique coding.

2. Overview of Proposed solution

The fundamental solution approaches just discussed cover a wide range of ways to maintain the highest network availability for a MANET with heterogeneous capabilities. From Figure 13, it can be seen that these fundamental solutions can be fit to a taxonomy based on the three graph theory concepts of Chapter II.

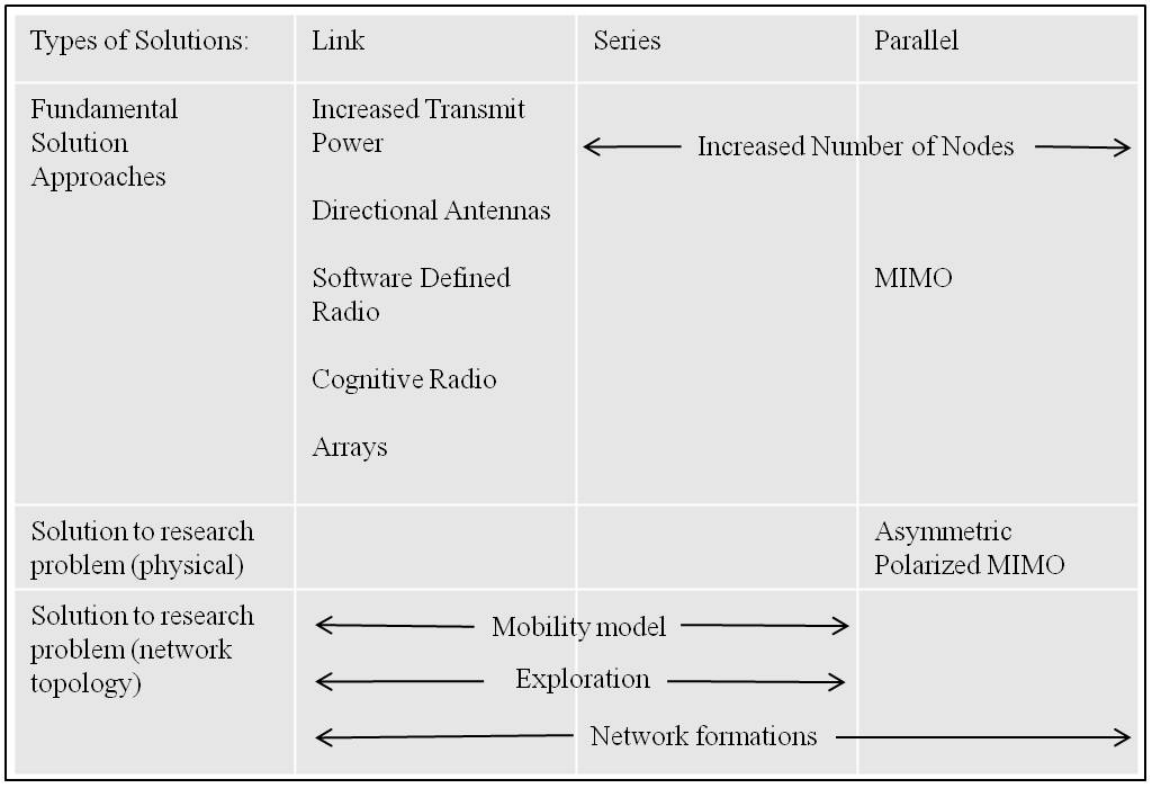


Figure 13. A taxonomy of fundamental solution approaches and the proposed solution.

A solution to the research problem is proposed first at the physical layer and then in the network topology. This proceeds from looking at what can be done to improve one link through asymmetric polarized MIMO (improve p_a) to the system level considerations of network topology (improve P_A).

The proposed solution method by asymmetric polarized MIMO uses a "parallel" category solution, at the physical layer. Beyond the MIMO capabilities already described, it has two unique characteristics. The asymmetric capability of communications systems was seen as typical of the UXS described in Chapter I, such as an UGV communicating with a WSN node. Polarization provides both the opportunity for a diversity in paths and making compact antenna arrays, which are potential advantages for small UXS. The metrics of the proposed solution are bit error rates (BER) and system capacity. We detail this solution method in Chapter IV.

The solution method of changing the network topology is considered in Chapter V. The network topology solution method takes on three different forms depending on the state of the MANET: a default state of all fixed nodes, one moving node, or all moving nodes. The link focus is maintained by considering realistic RF propagation mechanisms in every case. The use of asymmetric polarized MIMO is not shown but could be a solution for any link between nodes in these topologies. These system level changes in topology certainly have an effect across the entire TCP/IP stack. But the metric of the performance of these system level solutions will be abstracted to the network availability computation, as shown in Chapter II.

With the UXS MANET represented as a graph, we can more easily understand how changes in a single link between nodes affects network availability. All of the fundamental solution approaches, whether focused on link, series, or parallel, have the goal of improving link availability. Such solutions also can have negative effects, such as increased power causing increased interference. The unique solution of polarized MIMO is detailed in the next chapter utilizing the parallel transmission paths between nodes to improve the individual link availability.

IV. POLARIZED MIMO

In this chapter, we examine spatial diversity as a means to combat fading and what polarization can lend in terms of diversity. We then develop of various dual polarized antenna configurations and study their performance.

A. DIVERSITY

As outlined in Chapter II, there are large scale and small scale path loss effects in a channel. As a result of these effects and the movement of the transmitter and/or receiver, the combined fluctuations of the signal may add destructively at the receiver resulting in *fading*.

Diversity has the goal of communication link improvement and comes in various forms. The main idea of spatial diversity is to have multiple paths between the transmitter and receiver so that if one is subjected to deep fading, another may be chosen. With multiple paths, both the instantaneous and average SNRs at the receiver may be improved, often by as much as 20 dB to 30 dB [35]. This has led to very useful multi-antenna receiver implementations, such as selection diversity, maximal ratio combining (MRC), and equal gain combining (EGC) [39]. This focus on the multi-antenna receiver can be improved by transmitter strategies.

Spatial diversity gain is achieved by transmitting the same signal on multiple antennas that then have independently fading paths to the receiver. The spatial diversity gain is the number of independent paths with a maximum of $N_t N_r$, which is the product of the number of transmitting antennas and the number of receiving antennas. Interference reduction is possible using the multiple receive antennas to spatially differentiate co-channel users.

And yet another way to improve diversity is through coding, which is a form of spatial diversity gain. Alamouti space time coding (STC) [93] achieves the maximum spatial diversity gain without any transmitter channel knowledge. Although it transmits modified complex symbols from two antennas over two symbol periods in blocks, it is

considered a spatial diversity technique. With two transmit antennas, it has a code rate of one. Upon reception at the multi-antenna receiver, channel estimates are used in linear decoding. The decoding algorithm creates orthogonal subchannels for the two transmitted symbols. To be very specific, this is Alamouti orthogonal space time block coding (OSTBC), and it is widely used because of the simplicity of the linear receiver structure.

B. THE POLARIZED CHANNEL

While these diversity methods provide an improvement in received SNR, another consideration is the use of the multiple paths via polarization. In contrast to the methods just discussed, the diversity methods involving polarization are highly dependent on the channel.

1. Polarized Channel Characteristics

The transmitted signal has numerous changes due to the interactions of V-polarized and H-polarized signals with objects in the channel. Each reflection and refraction interaction can be described by the Fresnel equations, which indicate that different polarizations behave differently at interfaces. A macroscopic view of the numerous interactions (including diffraction) yields some probabilistic analysis of the channel. We begin by grasping a simple case.

One of the pioneers in polarization diversity is Vaughan [94] who considered a mobile V-polarized transmitting antenna and measured the V-polarized and H-polarized received signals at a base station. In the urban environment, for instance, he found that the mean power level of the V-polarized signal was 7 dB higher than the H-polarized signal. Regardless, because of the independent, NLOS paths (envelope correlation = -0.003), the polarization diversity could be used with MRC to get a 3 dB improvement in the mean received signal level. This example of polarization diversity was used to overcome deep fading, but it is dependent upon characteristics of the wireless channel.

Many empirical studies have followed, and tabulated results for the channels in some different environments and frequencies are summarized in [15]. Some insightful generalizations can be made from the data. Compared to the urban environment, the

suburban environment had a 12-dB power level difference between V-polarized and H-polarized [94]. With fewer interactions in the suburban environment, less of the V-polarized transmitted signal is scattered to H-polarized. This cross-polar scattering also increases with increasing path loss [95] (a longer path has more interactions). Certain environments, such as cities with tall buildings, form urban canyons that favor vertical polarization. Sometimes called "street guidance," a V-polarized transmitted signal is preserved, especially compared to propagation over rooftops, even with far reflections [96]. Another generally observed result is that the correlation between signals received (or transmitted) from V- and H-polarized antennas is very small [95].

Although V- and H-polarizations are well preserved in LOS and some unique scenarios, they are less common in many situations, such as NLOS for mobile nodes. We analyze only NLOS situations throughout this work such that the Ricean K-factor is zero, leading to Rayleigh fading. It is precisely this scattering that can yield gains in diversity.

2. Polarized Channel Model

We first describe specific notation for the antenna configurations at each end of the channel, then a unique hybrid channel model that meets the needs of the analysis. Figure 14 shows an asymmetric antenna configuration.

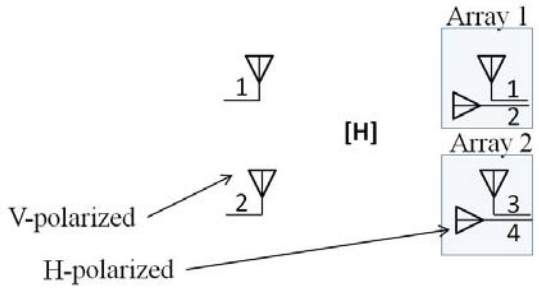


Figure 14. Asymmetric antennas configuration $(2 \times 2p)$ in a dual polarized channel, where each array is composed of vertical and horizontal polarized antennas.

The configurations of study are asymmetric, having one side of the communication link with a different capability antenna system (orthogonal dual polarized). The analysis is confined to the case of the same signal transmitted from one

"array" of dual polarized orthogonal antennas, differing only in the transmitted power of individual antennas. However, different signals may be transmitted on each array, as in coding. For this asymmetry, a simple notation is (number of downlink transmitter arrays) $\times$ (number of uplink transmitter arrays), where a dual polarization array is denoted by the suffix 'p'. We consistently refer to downlink as left to right in all diagrams and uplink as right to left. The example $2\times 2\mathrm{p}$ system of Figure 14 can then be compared to a familiar 2×2 (all vertical antenna) system, having identical signal inputs.

Along with the model description, there are some assumptions and limitations. First, it is recognized that there are two RF chains for each dual polarized array, increasing the cost and complexity of a transceiver. Another negative effect is that each receive antenna in the array is simulated with a noise source n with $\mathrm{E}\{nn^*\}=1$. We also strive to maintain valid comparisons with unipolar systems by limiting each array to the equivalent transmit power of a single antenna. The modeled antennas are assumed to have perfect cross-polar isolation, being insensitive to orthogonal polarizations. Imperfect cross-polar isolation is treated in [15]. The gain of each antenna is unity in order to focus upon the channel characteristics; higher gain antennas will obviously improve performance.

The model needed for the analysis of asymmetric configurations requires unique qualities. One polarized channel model [14] incorporates an important normalization motivated by power conservation, namely that the channel gains from a single vertical transmitter should sum to unity. Another polarized channel model [97] splits the effects of spatial correlation and polarization into matrices that are Hadamard multiplied but with the normalization to the vertical channel. For determining capacity of an all vertical configuration, this would show no change with changing cross-polarization ratio due to this normalization. We synergize these two models to one that effectively isolates the spatial correlation and polarization correlation and can be used to fully compare the VV and VH configurations against a benchmark under changing cross-polarization ratio.

This hybrid model also carries assumptions. One assumption of the model is that the Rayleigh fading channel has spatial fading correlation that is independent of the

polarization, provided the direction spectra are similar for all antennas [97]. Further, the model assumes flat fading, so the complex Rayleigh fading channel values remain constant for a number of symbol periods. This hybrid model infers that the channel treats all vertically polarized transmitted signals in an identical manner regardless of transmit or receive locations. Then using the dual polarized Rayleigh fading channel model of [15], the downlink channel is the N_r by N_t MIMO matrix

$$H = H_S \otimes X_P \tag{23}$$

where H_S is the uni-polarized spatially correlated Rayleigh channel, $\otimes$ represents the Kronecker product, and X_P includes the correlation and power imbalance from channel polarization.

The spatial correlation matrix R_s is constructed assuming all uni-polarized antennas, hence the elements of each array are co-located providing a symmetry that simplifies the R_S expression. The spatial correlation channel effect is described by

$$\mathrm{vec}(H_S) = R_s^{1/2}\mathrm{vec}(H_w) \tag{24}$$

where H_w is a matrix of complex i.i.d. Gaussian RVs of zero mean and equal variance. The vec( ) operator vectorizes a matrix by stacking its columns. The spatial correlation matrix is

$$R_s = \begin{bmatrix} 1 & r_1^* & t_1^* & c^* \\ r_1 & 1 & x^* & t_2^* \\ t_1 & x & 1 & r_2^* \\ c & t_2 & r_2 & 1 \end{bmatrix}, \tag{25}$$

where t and r are the familiar transmit and receive antenna spatial correlations. The variables c and x are, respectively, the co-channel (h_{11}, h_{22}) and cross-channel (h_{21}, h_{12}) correlations. For notation, a^* indicates the complex conjugate of a. Correlation matrix formation has the potential to increase the squared Frobenius norm $\|H\|_F^2$ (total power) in the channel, so the total channel power is normalized to the product of the number of uplink and downlink arrays in the studied configuration.

A 2 by 2 matrix X_P accounts for polarization effects as

$$\mathrm{vec}(X_P) = R_P^{1/2}\mathrm{vec}(X_w), \tag{26}$$

where X_w is a 2 by 2 matrix of independent complex exponentials $e^{j\theta_k}$ with θ_k having distribution $U[0, 2\pi)$. The polarization correlation matrix R_P models the power imbalance and correlation of the polarized channel as if they were co-located cross-polarized antennas. To describe this, subscripts denote the polarizations at each end of the four subchannels in the dual-polarized downlink channel matrix consisting of four impulse responses

$$H_{pol} = \begin{bmatrix} h_{vv} & h_{vh} \\ h_{hv} & h_{hh} \end{bmatrix}, \tag{27}$$

where $p_{ij} = \mathrm{E}\{|h_{ij}|^2\}$ is the power change in polarizations due to the channel. An important parameter is

$$\chi_V = \frac{p_{vv}}{p_{hv}} = \frac{\mathrm{E}\{h_{vv}h_{vv}^*\}}{\mathrm{E}\{h_{hv}h_{hv}^*\}} = \frac{1-\zeta}{\zeta}, \tag{28}$$

which describes the channel's response to power that is transmitted vertically compared to what is "leaked" to the horizontal polarization. Most literature equates χ to the cross-polarization ratio or XPR. The fraction of the transmitted power that is coupled from vertical to horizontal is represented by ζ [14], which we used extensively in this research. In a similar manner, (with the assumption that the co-polar ratio is 0 dB) the horizontal downlink power is expressed as

$$\chi_H = \frac{p_{hh}}{p_{vh}} = \frac{E\{h_{hh}h_{hh}^*\}}{E\{h_{vh}h_{vh}^*\}} = \frac{1-\eta}{\eta}. \tag{29}$$

These values compose the vital parts of the 4 by 4 matrix, given by

$$R_P = \mathrm{E}\{\mathrm{vec}(H)\mathrm{vec}(H^H)\} = \begin{bmatrix} 1-\zeta & r_{1p}^* & t_{1p}^* & c_p^* \\ r_{1p} & \zeta & x_p^* & t_{2p}^* \\ t_{1p} & x_p & \eta & r_{2p}^* \\ c_p & t_{2p} & r_{2p} & 1-\eta \end{bmatrix}, \quad (30)$$

where $r_{1p}, r_{2p}, t_{1p}, t_{2p}$ are the cross polar correlations, c_p the co-polar correlation, and x_p the anti-polar correlation and are labeled to show their similarity to the spatial correlations. The values of each are dependent on the communications environment. Some simplifying assumptions for our analysis are that off-diagonal elements of R_P are zero, similar to [15]. Despite the hybrid model's accommodation of any combination of V- and H-polarization transmitters, we chose V-polarization for downlink transmitters to simplify comparisons and limit complexity of the analysis.

C. MIMO CONFIGURATIONS FOR POLARIZED CHANNELS

We will analyze the asymmetric configurations by examining three special case configurations and comparing them to only V-polarized configurations. We should expect that a dual polarization receiver will capture more signal power than a single polarization receiver under dual polarized channel conditions.

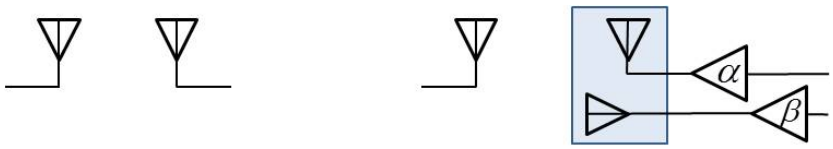
To determine the power advantage, the analyses will build from a general, scalar signal model

$$y = \sqrt{E_s}Hs + n, \quad (31)$$

where E_s is the transmit energy per symbol, s is the complex symbol from the QPSK constellation and n is the noise with a zero mean, circularly symmetric, complex Gaussian random process.

1. 1×1 and 1×1p Downlink

The most basic configurations to compare are the 1×1 and $1\times1\mathrm{p}$ configurations, as in Figure 15.



(a) (b)

Figure 15. Two configurations in dual polarized channels: the (a) 1×1 and (b) $1\times1p$ with uplink gain multipliers.

The first configuration is a single polarization downlink receiver, and the second configuration is dual polarization. Without the knowledge of the Rayleigh fading phase, the single channel receiver's performance approaches a bit error $P_E = 0.5$. For a fair comparison of the single channel receiver to a dual channel MRC receiver, we assume the single channel receiver has "ideal" coherent detection.

The symbol error rate is obtained by Monte Carlo simulation with the results in Figure 16. Note that there is unequal power in the downlink channels, except at $\zeta = 0.5$. When ζ is small, performance of the two receiver configurations is essentially the same, as expected. As ζ increases, the power leaked to the horizontal channel is available to the dual polarization configuration and not captured by the 1×1 configuration, creating the difference in P_e. The second antenna provides a receive diversity of order two that improves performance in the Rayleigh fading environment.

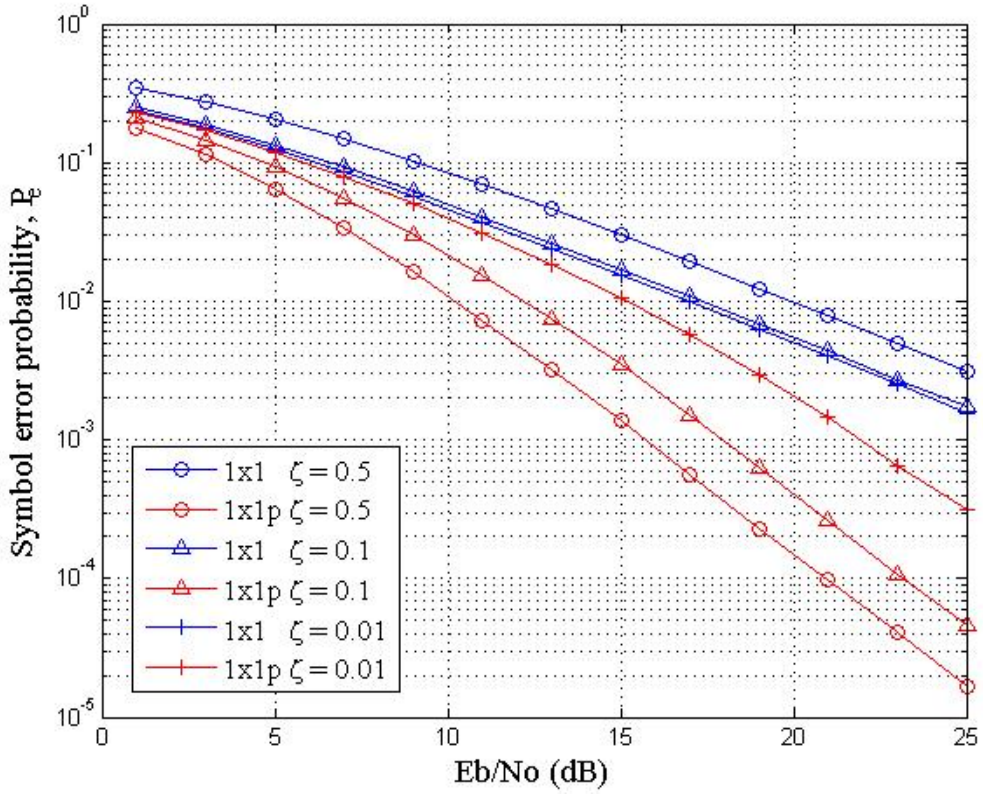


Figure 16. The effect of ζ on the 1×1 and $1\times1\text{p}$ downlinks with QPSK in Rayleigh fading.

Seeing these results from a single transmitter, we next consider if there is a multiple transmitter advantage. Actually, the same symbol error probability performance will result if there are many V-polarized transmit antennas. This comes from the transmit antenna diversity argument of [91] extended to this special case of polarized channels. Assume a flat fading channel with two V-polarized transmit antennas, each sending the same signal, but each with half the transmit power. If ζ is the same for the channels, then the received signal at the V-polarized receiver is

$$y_v = \sqrt{\tfrac{E_s}{2}}\sqrt{1-\zeta}\, w_{vv1}\, s + \sqrt{\tfrac{E_s}{2}}\sqrt{1-\zeta}\, w_{vv2}\, s + n. \qquad (32)$$

Knowing that $w_{vvx}, x \in 1,2$ is a complex Gaussian random variable and the sum of two complex Gaussian random variables is also a random variable, $\frac{1}{\sqrt{2}}(w_{vv1} + w_{vv2})$, yields a simpler equation,

$$y_v = \sqrt{E_s}\sqrt{1-\zeta}\, w_{vv}\, s + n = \sqrt{E_s}\, h_{vv}\, s + n, \tag{33}$$

which is of the same form as (31). This implies that there is no additional diversity gain from the same polarization, multi-antenna transmitter, under the constraint of equal transmit power. A similar argument can be applied to the H-polarization of the receiver. So, although generated from the *single* transmitter antenna, the results of Figure 16 could equivalently be labeled as the results from a single polarization, *multi*-antenna transmitter.

2. 1×1 and 1×1p Uplink and Polarization Reciprocity

The asymmetric configuration has the capability of transmitting maximum signal energy to the uplink receiver with changing ζ values. A scheme for maximizing power transmission using waveguides in cross-polarized channels is detailed in [98]. This 'variable polarization' system has two receive antennas of orthogonal polarizations and can extract maximum incident power by adjusting phase. The system then uses the same polarization power ratios to transmit back to a single polarization receiver. This is similar to what is used in the asymmetric configuration, but is in the LOS context; we will examine this idea in the NLOS, fading context.

To take advantage of this maximum signal transfer back to the original transmitter in a polarized channel, the channel is assumed unchanging for a period of time, the coherence time. For slow fading, this coherence time is much greater than the symbol period. The downlink receiver must measure the relative power of the incoming signal on two antennas (polarizations), then use this channel knowledge for transmitting during the coherence time. This temporarily valid knowledge is the reciprocal of the downlink polarized channel matrix, hence the term *polarization reciprocity* (PR).

To provide maximum signal energy to the uplink receiver, consider Figure 15. The downlink signal experiences independent Rayleigh fading in both polarizations in the channel. The receiver in Figure 15(b) will have different received powers at the V- and H-polarized antennas due to the channel ζ. This *channel knowledge can be exploited in the uplink channel via reciprocity* if the uplink transmission has the same ratio as the downlink received powers in the quasi-static channel.

To split the transmitted uplink power between polarizations, we define the uplink gain multipliers proportional to the downlink signals as the following scalar values:

$$\alpha = \frac{|y_{vv}|}{\|y\|} \ , \ \beta = \frac{|y_{hv}|}{\|y\|} \ . \tag{34}$$

So that they have the equivalent uplink power as the single transmitter, α and β must meet the constraint:

$$\alpha^2 + \beta^2 = 1. \tag{35}$$

The values of α and β modulate the uplink signal on the two polarizations.

For the uplink performance, if we make the same assumptions as the downlink of "ideal" coherent detection, then the dual uplink transmitter keeps maximum power to the receiver under all conditions of ζ. These Monte Carlo simulation results are shown in Figure 17.

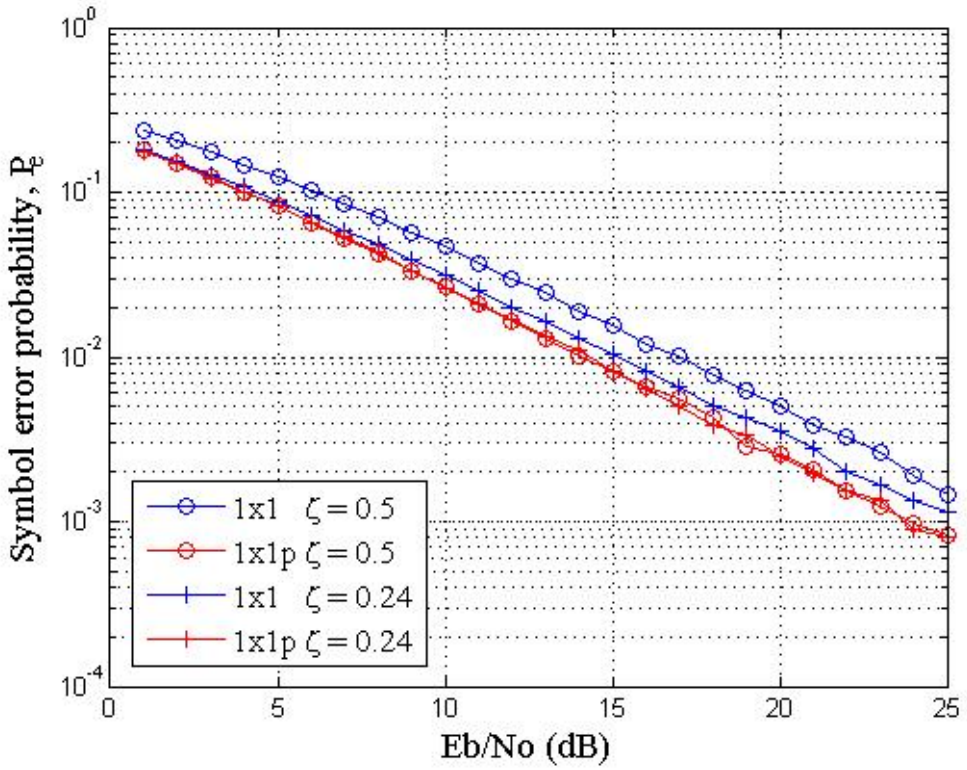


Figure 17. The effect of ζ on the 1×1 and $1\times1p$ uplink with QPSK in uncorrelated Rayleigh fading.

As expected, at the half-power channel ($\zeta = 0.5$) the single uplink transmitter requires 3 dB more E_b / N_o for the same performance.

We now examine if coding can improve the uplink performance. The uplink of Figure 15(b) could use Alamouti coding with two transmitters and one receiver [93]. This uses equal power (EP) transmission for each polarization. This performance as a function of ζ is shown in Figure 18.

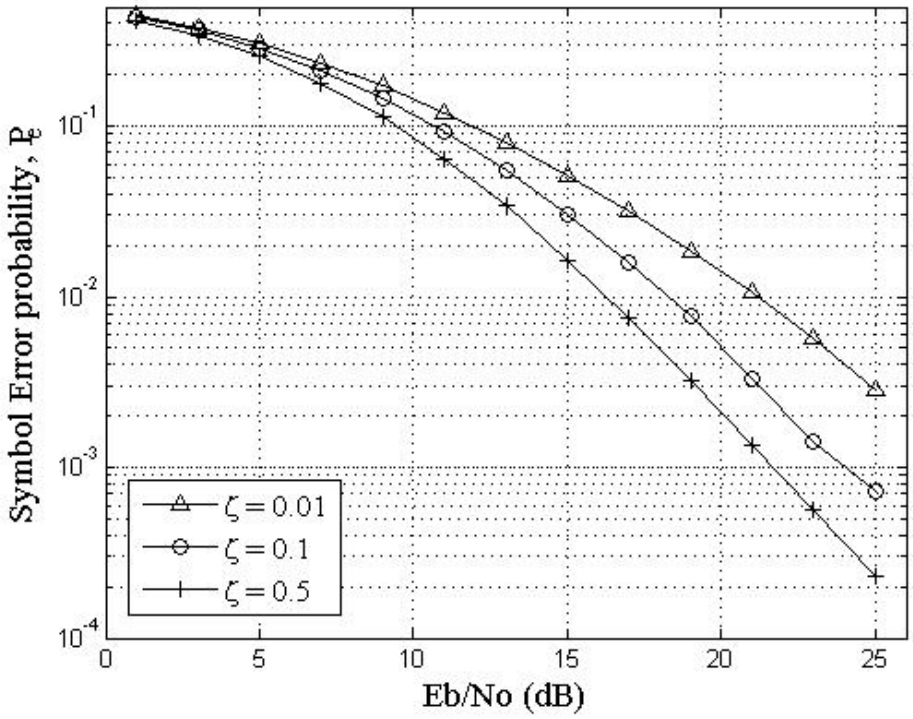


Figure 18. The effect of ζ on Alamouti coded 1×1p EP uplink with QPSK modulation and uncorrelated Rayleigh fading.

If the transmission scheme is changed from equal transmit power in each polarization to a variable power, we use Equations (34) and (35) to modulate the uplink signal power.

When this is joined with the Alamouti coding, the uplink received signals during two adjacent symbol times are

$$y_1 = \alpha h_1 s_1 + \beta h_2 s_2 + n_1, \tag{36}$$

$$y_2 = -\alpha h_1 s_2^* + \beta h_2 s_1^* + n_2. \tag{37}$$

After decoding and detection, the signal estimates $\tilde{s}$ are:

$$\tilde{s}_1 = (\alpha |h_1|^2 + \beta |h_2|^2) s_1 + (\beta h_1^* h_2 - \alpha h_1^* h_2) s_2 + h_1^* n_1 + h_2 n_2^* \tag{38}$$

$$\tilde{s}_2 = (\alpha|h_1|^2 + \beta|h_2|^2)s_2 + (\alpha h_1 h_2^* - \beta h_1 h_2^*)s_1 + h_2^* n_1 - h_1 n_2^* . \quad (39)$$

However, the second terms of (38) and (39) do not elegantly cancel out as the basic Alamouti code does with perfect channel estimation. Hence, we suspect worse performance, and this is in fact seen in Figure 19 when compared to Figure 18.

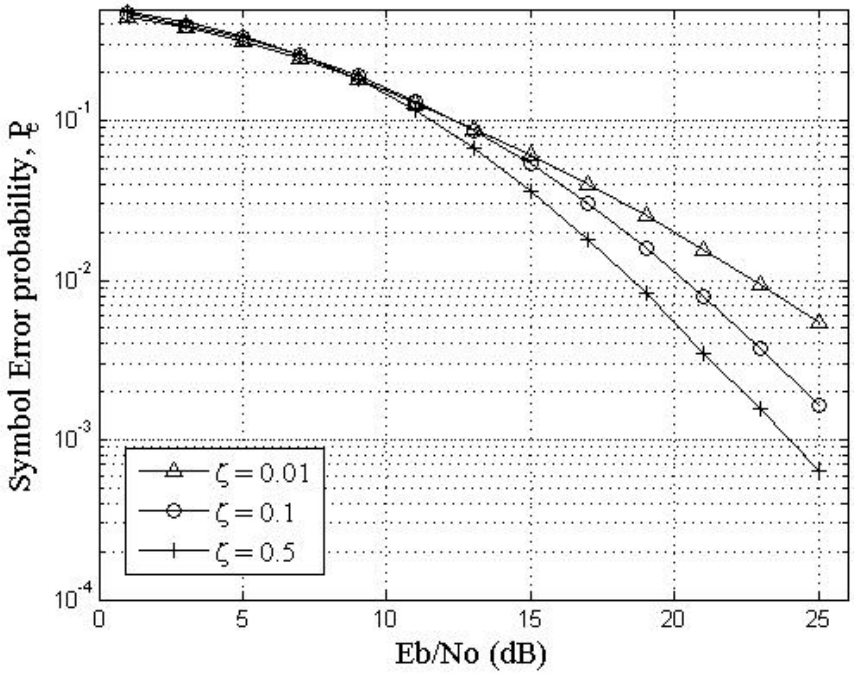


Figure 19. The effect of ζ on Alamouti coded 1×1p PR uplink with QPSK modulation and uncorrelated Rayleigh fading.

There are three main results from this uplink performance analysis. First, the use of dual channel PR is shown by simulation to provide lower P_e at the same E_b / N_o compared to the single transmitter. This effect is obviously more pronounced for higher values of ζ. Next, comparing Figure 17 and Figure 18, for high ζ and higher E_b / N_o (> 20 dB), the use of EP transmission with Alamouti coding outperforms uncoded single and dual transmitters. Lastly, comparing Figure 18 and Figure 19, the use of the PR scheme does not enhance the Alamouti coding performance. The gain multipliers α and β unbalance the cancellations in the Alamouti detection equations.

3. 1×1 and 1×1p Capacity

To compare uplink capacities of these configurations, we repeat the well-known MIMO channel capacity formula [12] in an uncorrelated channel where ρ is the average SNR at each receiver branch,

$$C = \log_2 \det(I_{N_r} + \frac{\rho}{N_t} HH^H). \tag{40}$$

Here N_r and N_t are the number of receive and transmit antennas, I_{N_r} is the N_r by N_r identity matrix, and H is the N_r by N_t channel matrix, and the conjugate transpose of any matrix A is expressed as A^H .

Because a receiver cannot separate a channel gain from a transmitter gain effect, the transmit uplink gain multipliers are included in the capacity equations. Referring to Figure 15(b), the capacity for the scalar vertical and horizontal channels can be expressed as:

$$C_V = \log_2(1 + \rho\alpha^2(\sqrt{1-\zeta})h_w(\sqrt{1-\zeta})h_w^*) \tag{41}$$

$$C_H = \log_2(1 + \rho\beta^2(\sqrt{\zeta})h_w(\sqrt{\zeta})h_w^*), \tag{42}$$

where h_w is the complex Gaussian RV with zero mean and unit variance. Summing as the dual channel capacity, $C_D = C_V + C_H$, yields

$$C_D = \log_2(1 + (\alpha^2(1-\zeta) + \beta^2\zeta + \rho\alpha^2\beta^2\zeta(1-\zeta)h_w h_w^*)(\rho h_w h_w^*)). \tag{43}$$

This is compared to the single (vertical) channel capacity of Figure 15(a),

$$C_S = \log_2(1 + \rho(\sqrt{1-\zeta})h_w(\sqrt{1-\zeta})h_w^*) \tag{44}$$

so $C_D = C_S$ when

$$\alpha^2(1-\zeta) + \beta^2\zeta + \rho\alpha^2\beta^2\zeta(1-\zeta)h_w h_w^* = 1-\zeta. \tag{45}$$

Since h_w is a Rayleigh RV, (43) is simplified by taking the expected value of both sides, and $E\{h_w h_w^*\} = 1$. Then, removing the trivial case of $\zeta = 1$, we restrict $0 \le \zeta < 1$, yielding

$$\alpha^2 + \beta^2 \left(\frac{\zeta}{1-\zeta} \right) + \rho \alpha^2 \beta^2 \zeta = 1, \tag{46}$$

which is a scalar function α of the variables ζ and ρ. This threshold value can be imagined as a surface for α_{min}. Values of $\alpha > \alpha_{min}$ provide greater capacity for the 1x1p compared to the 1x1 configuration.

We next determine the optimum values of α, those that maximize the capacity across the same two variables. For a given channel condition ζ and ρ, we desire to maximize the dual channel capacity (43) under the constraint of (35).

Taking the expected value of (43), we consider the ergodic capacity:

$$E\{C_D\} = E\{\log_2(1 + \rho\alpha^2(1-\zeta) h_w h_w^* + \rho\beta^2 \zeta h_w h_w^* + \rho^2 \alpha^2 \beta^2 \zeta (1-\zeta)(h_w h_w^*)^2)\}. \tag{47}$$

By Jensen's inequality, this can be upper bounded as

$$E\{C_D\} \le \log_2 E\{(1 + \rho\alpha^2(1-\zeta) h_w h_w^* + \rho\beta^2 \zeta h_w h_w^* + \rho^2 \alpha^2 \beta^2 \zeta (1-\zeta)(h_w h_w^*)^2)\}. \tag{48}$$

Because the distribution of $h_w h_w^*$ is χ^2, the $E\{(h_w h_w^*)^2\} = 2$,

$$E\{C_D\} \le \log_2 (1 + \rho\alpha^2(1-\zeta) + \rho\beta^2 \zeta + 2\rho^2 \alpha^2 \beta^2 \zeta (1-\zeta)). \tag{49}$$

Since log is a monotonically increasing function, the capacity is maximized when the dummy variable η is maximized:

$$\eta = \rho((1-\zeta)\alpha^2 + (1-\alpha^2)\zeta + 2\rho\alpha^2(1-\alpha^2)\zeta(1-\zeta)). \tag{50}$$

Expanding these terms and combining yields a fourth and second order term in α. Then taking the derivative with respect to α yields a depressed cubic polynomial,

$$\frac{\partial \eta}{\partial \alpha} = -8\rho^2 \zeta (1-\zeta)\alpha^3 + (2\rho - 4\rho\zeta + 4\rho^2 \zeta (1-\zeta))\alpha. \tag{51}$$

Factoring (51), there are three roots, one of which is $\alpha = 0$. The other two roots are

$$\alpha_{opt} = \pm\sqrt{\frac{(1/2) - \zeta + \rho\zeta(1-\zeta)}{2\rho\zeta(1-\zeta)}}, \tag{52}$$

of which the positive value is used. These α_{opt} are plotted in Figure 20 and reveal a symmetry about the $\zeta = 0.5$ channel condition.

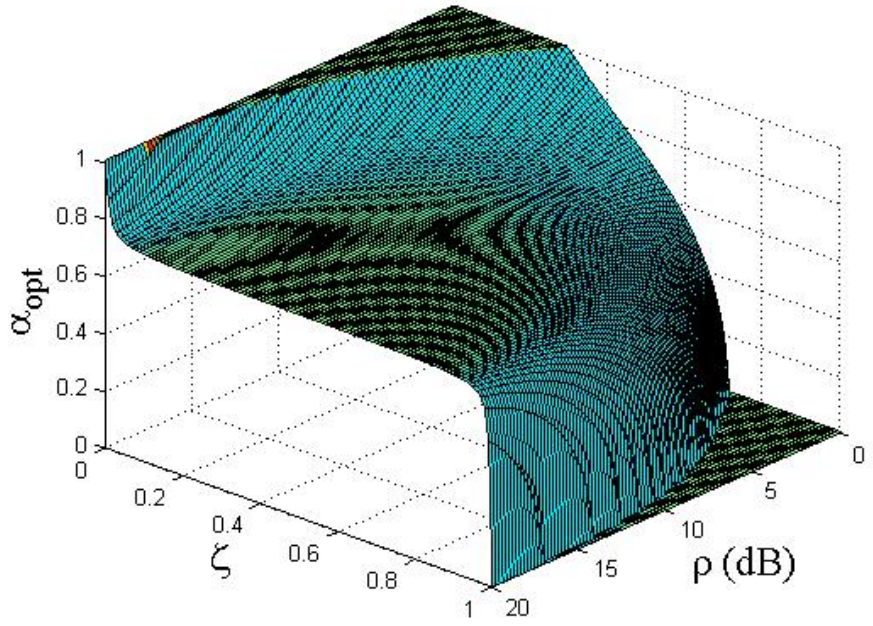


Figure 20. The optimum uplink gain multiplier α for maximizing the uplink capacity of the 1×1p configuration.

Because these values of α and β represent gains in two channels to maximize capacity, it is similar to water-filling [99]. The water-filling algorithm can optimally allocate power between channels of different quality, in this case simply two subchannels. For the polarized channel, each realization has

$$h_1 = \sqrt{1-\zeta}\,h_w \tag{53}$$

$$h_2 = \sqrt{\zeta}\,h_w \tag{54}$$

and these have scalar eigenvalues $\lambda_1 = E\{h_1 h_1^*\} = 2(1-\zeta)$ and $\lambda_2 = E\{h_2 h_2^*\} = 2\zeta$.

The configuration can be evaluated by two metrics, error rate and capacity. We first consider the bit error rate (BER). One may observe that in the derivation of the optimal uplink gain multiplier, the expected values were used. Does the same performance hold for schemes when they are used on each channel realization? To add a second scheme for comparison, we consider the equal power (EP) scheme, where $\alpha = \beta = 1/\sqrt{2}$ for every channel instance. This may be used as a default uplink power setting when channel state information is unavailable. We use a QPSK 2×1 Alamouti-encoded uplink with an uncorrelated Rayleigh channel and Monte Carlo methods $(N = 200{,}000)$ to generate the BER curves in Figure 21. Better error performance is obtained when the channel permits a lower cross-polarization ratio (when ζ increases).

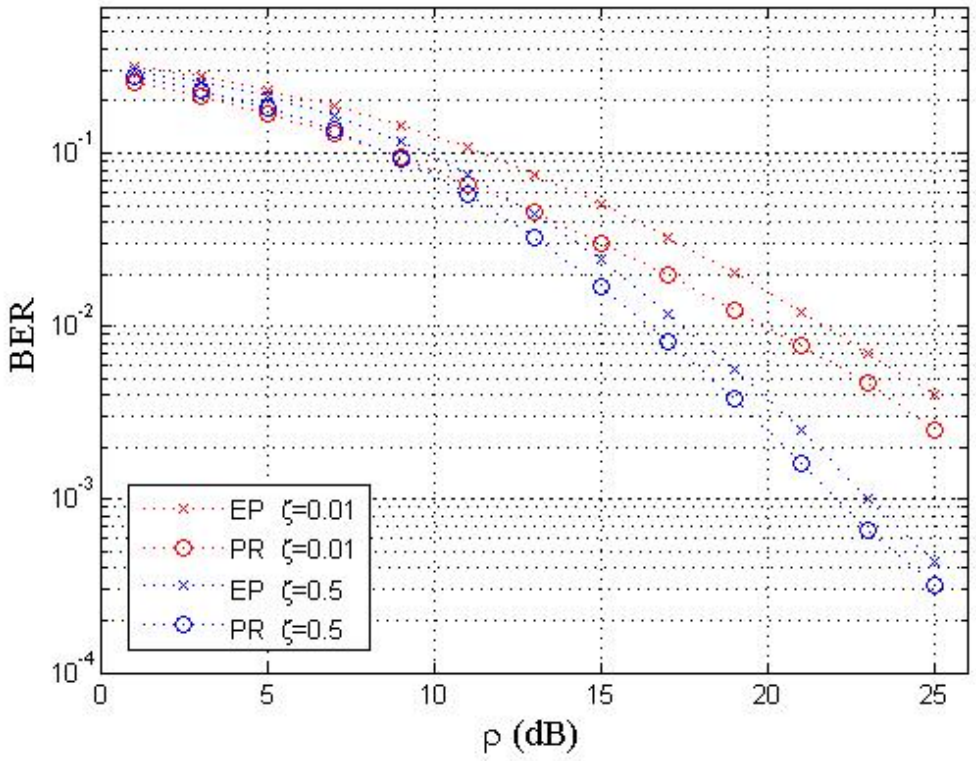


Figure 21. A comparison of BER performance of two schemes for uplink power in the $1\times1p$ configuration.

In contrast to just two values of ζ as shown in Figure 21, the full range of ζ values can determine the ergodic capacity. The ergodic capacity of the $1\times 1\text{p}$ configuration (43) using the uplink multiplier values that maximize capacity (52) is compared to the 1×1 configuration capacity (44). From Figure 22, it can be seen that the dual polarization has an advantage in capacity, especially as ρ increases. The vertically transmitted 1×1 configuration capacity goes to zero when all the power goes to H-polarization at $\zeta = 1$.

Note that the consistency or improvement in capacity as a function of ζ is performed with *the same uplink transmit power* as the 1×1 configuration. The values of ζ and ρ at which the capacities of the two configurations are equal can be found by curve fitting, as in Figure 23.

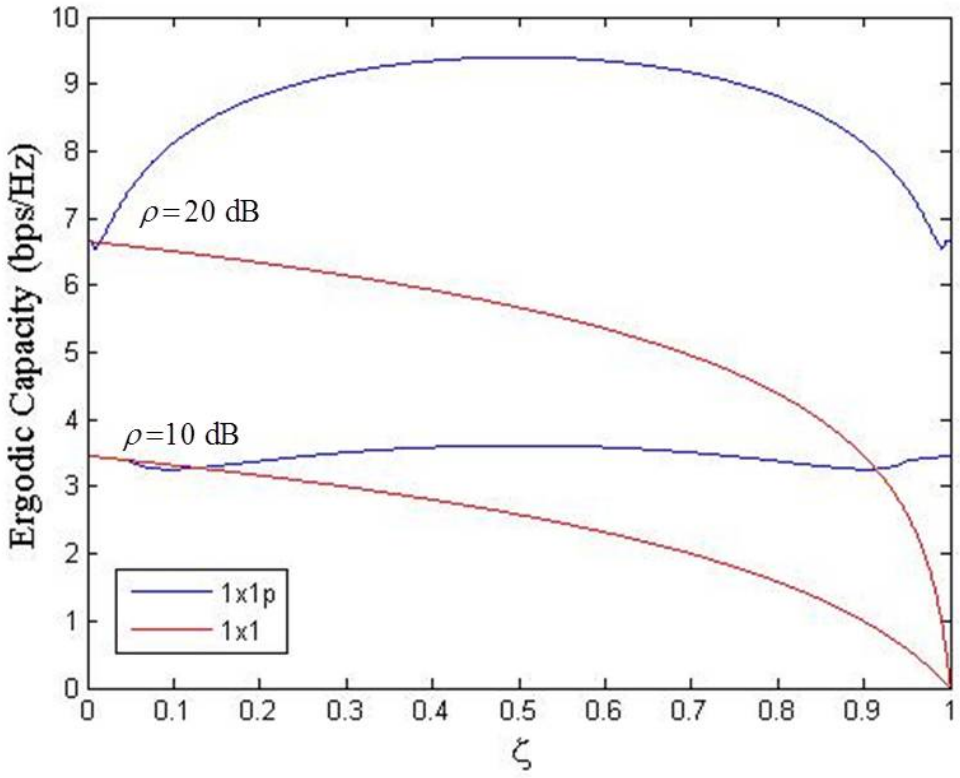


Figure 22. A comparison of the ergodic uplink capacity for $1\times 1\text{p}$ and 1×1 configurations in the polarized channel for varying cross-polarization ratios.

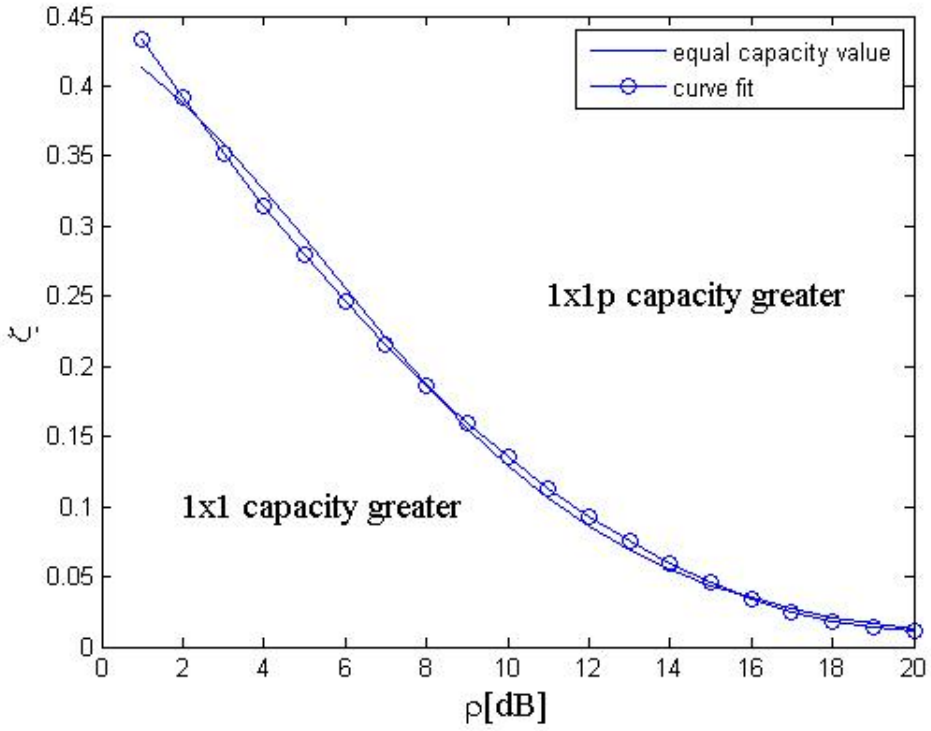


Figure 23. The channel conditions for choosing 1×1 or $1\times1\text{p}$ for greater uplink capacity.

The conditions under which dual-polarization uplink capacity is greater is approximated by the curve fit as

$$\zeta \geq 0.0011\rho_{[dB]}^2 - 0.0452\rho_{[dB]} + 0.4777 \tag{55}$$

Finally, we remark that the downlink capacity of the $1\times1\text{p}$ is maintained at the $\zeta = 0$ value (Figure 22) for all ζ. However, the downlink capacity of the 1×1 is identical to the uplink capacity of Figure 22, decreasing as ζ increases.

4. $2\times1\text{p}$

The addition of diversity paths is expected to improve performance of MIMO systems, even more so with space-time coding (STC). For the downlink, adding another transmitter of identical polarization creates the 2×1 configurations, as shown in Figure

24. In this analysis, the second downlink transmitting antenna sends an identical signal and equally shares the total transmit power.

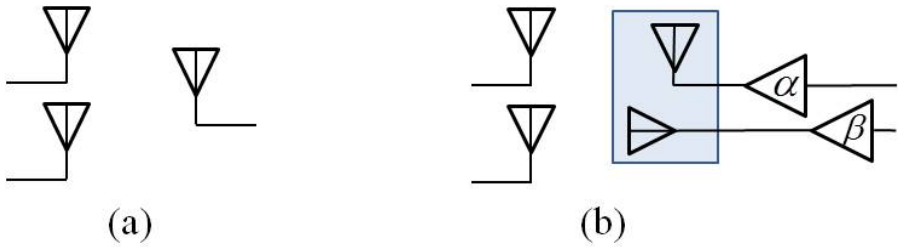


Figure 24. The (a) 2×1 configuration and (b) $2\times1\text{p}$ with uplink gain multipliers in dual polarized channels.

We then consider 2 by 1 Alamouti STC [93], and since the downlink of $2\times1\text{p}$ has four paths, the 2 by 2 Alamouti STC is investigated. Downlink STC effectiveness is highlighted in the unbalanced power distribution among channels, which is a characteristic of these polarized channels. The 2×1 configuration downlink receive antenna has a single output whereas the $2\times1\text{p}$ receiver uses maximum ratio combining (MRC), which requires CSIR. This is expected to maintain performance across changing ζ .

Similar to the previous capacity discussion, the BER will have different curves for the 2×1 configurations' uplinks and downlinks. On downlink without STC, Figure 25 shows that the 2×1 configuration loses the leaked power but maintains a diversity gain of 1 as ζ increases. The similar $2\times1\text{p}$ configuration offsets the power loss of the vertical channel as ζ increases with diversity gain from the horizontal receiver. The $2\times1\text{p}$ achieves a maximum diversity gain of 2 as ζ reaches 0.5.

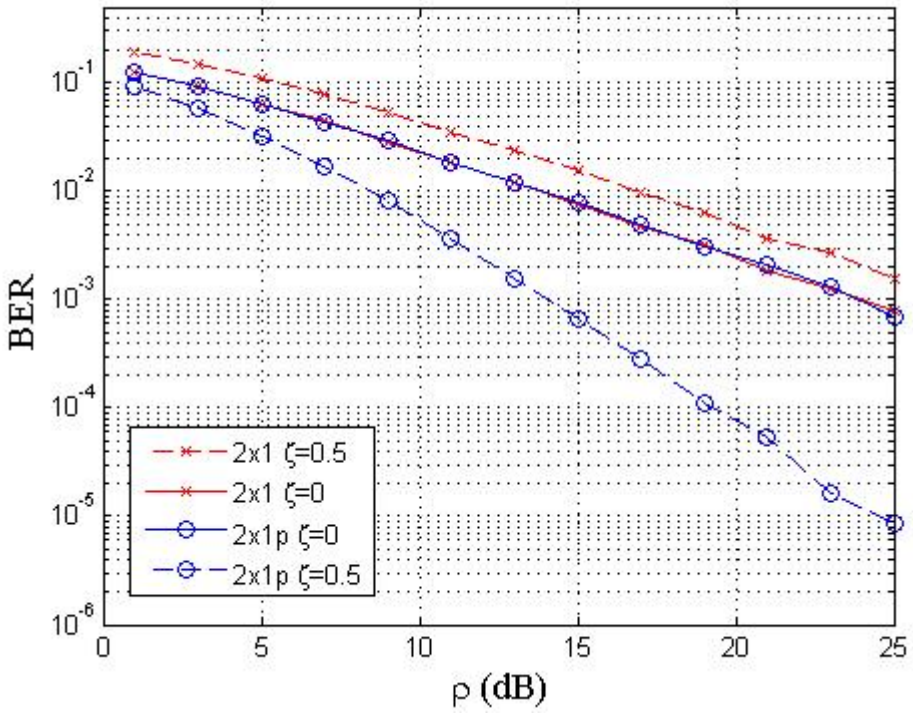


Figure 25. A comparison of BER for 2×1 and $2\times1p$ downlink configurations with no STC in the polarized channel with varying cross-polarization ratio, using QPSK modulation.

If STC is added to the downlink, we also see the generally advantageous effect, as in Figure 26.

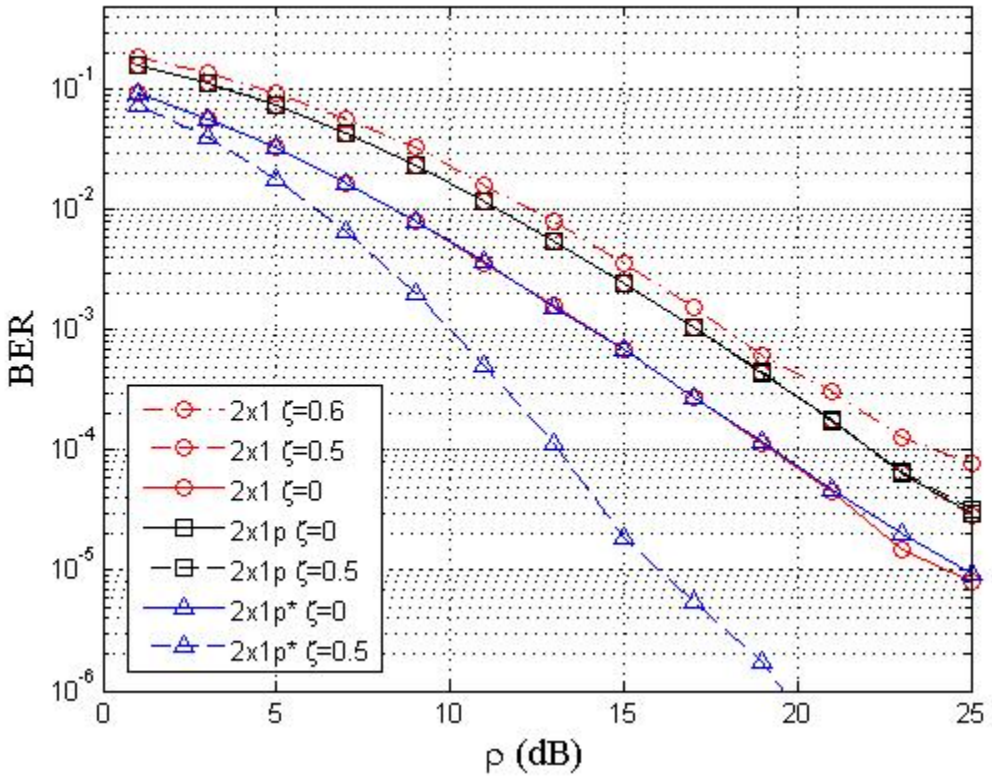


Figure 26. A comparison of BER for three downlink configurations with Alamouti STC in the polarized channel for varying cross-polarization ratio using QPSK modulation.

For the 2×1, the Alamouti STC extracts diversity gain 2 regardless of ζ. But the STC 2×1 performance is affected by the total power gain of the channel, where $\zeta = 0.5$ has reduced the squared Frobenius norm of H, $\|H\|_F^2$ by half [91]. Even with coding, the 2×1 configuration has this leakage of power that decreases performance. Using the 2×1p configuration with 2 by 1 Alamouti STC, the outputs of the V- and H-polarized antennas are non-coherently summed, then processed in an identical manner to 2×1. With our basic assumption of non-coherent addition (two Rayleigh signals) and no equalization, the 2×1p has higher BER than 2×1, and the coding maintains a diversity gain of 2. The principal advantage of 2×1p with 2 by 1 Alamouti STC is that the BER

does not change with ζ. Barring equalization, better performance would come from using the $2\times 1\text{p}$ configuration with selection combining.

If we use the $2\times 1\text{p}$ configuration with 2 by 2 Alamouti coding, we label this as the $2\times 1\text{p}^{*}$ configuration. The transmit side is identical to previous (transmit power, code matrix), but the receiver must add the complexity of the 2 by 2 Alamouti decoding. Rather than the non-coherent summation of two subchannels to one, each antenna separates the two independent subchannels. The 2 by 2 STC combiner outputs can be expressed as [93]:

$$\tilde{s}_1 = \left(\|H\|_F^2\right) s_1 + n_1 \tag{56}$$

$$\tilde{s}_2 = \left(\|H\|_F^2\right) s_2 + n_2 , \tag{57}$$

where s is the input constellation symbol and n is the summary noise term from the STC combiner. At $\zeta = 0$, $h_{21} = h_{22} = 0$ but these increase with ζ, improving the total channel power. The $\zeta = 0$ case will be identical to the 2×1 $\zeta = 0$ case. But as ζ increases, the BER decreases compared to the 2×1 due to additional diversity paths of the V-polarized to H-polarized channels, reaching a maximum diversity gain at $\zeta = 0.5$. Values of $\zeta > 0.5$ then increase the BER.

Now considering the uplink in Figure 27, only the ratio ζ is used for determining α and β; it does not require channel estimation, as explained in Section II.

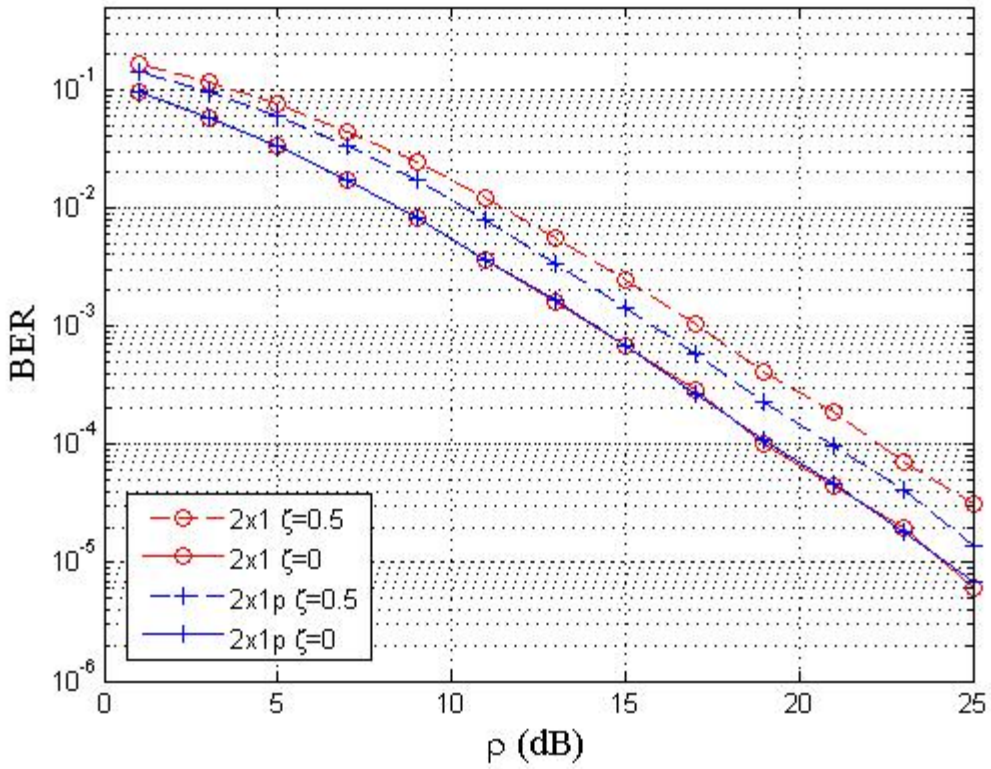


Figure 27. A comparison of BER for 2×1p (PR) and 2×1 uplink configurations.

Similar in concept to the MRC receiver, which multiplies a signal by the complex conjugate of its channel to maximize SNR, this "feedback" scheme uses the stronger channel proportionately. Both 2×1 and 2×1p use the same uplink transmit power, and errors are identical in the $\zeta = 0$ case. But 2×1p with PR has equivalent or better performance in polarizing channels using the non-coherent summations as

$$y_v = \left|h_{11} + h_{12}\right| \tag{58}$$

and

$$y_h = \left|h_{21} + h_{22}\right|. \tag{59}$$

Unique to the asymmetric configuration is the potential for different capacities in downlink and uplink. In the capacity formula (40) the N_t term is used to specify a number of equal power components of the transmit vector [12]. In our configurations, the

"p" suffix indicates the same power as a single antenna split among the cross polarized antennas of an array by some scheme (EP, PR, or WF) but with a transmit power of unity. Explicitly, we count these unique arrays as one antenna with regards to the capacity formula. Hence in the $2\times 1\text{p}$ case the channel matrix is square, but the downlink ($N_t=2$) and uplink ($N_t=1$) capacities are different. The $2\times 2\text{p}$ case has a rectangular channel matrix, but only one capacity formulation ($N_t=2$) for both downlink and uplink. We will see that this also provides a consistent comparison with the same power, single polarization (vertical only) cases. Utilizing (40), the uplink capacities of the respective 2×1 configurations are greater than the downlink capacities, as shown in Figures 28 and 29. They also show the invariance of the $2\times 1\text{p}$ configuration to changes in channel characteristic ζ .

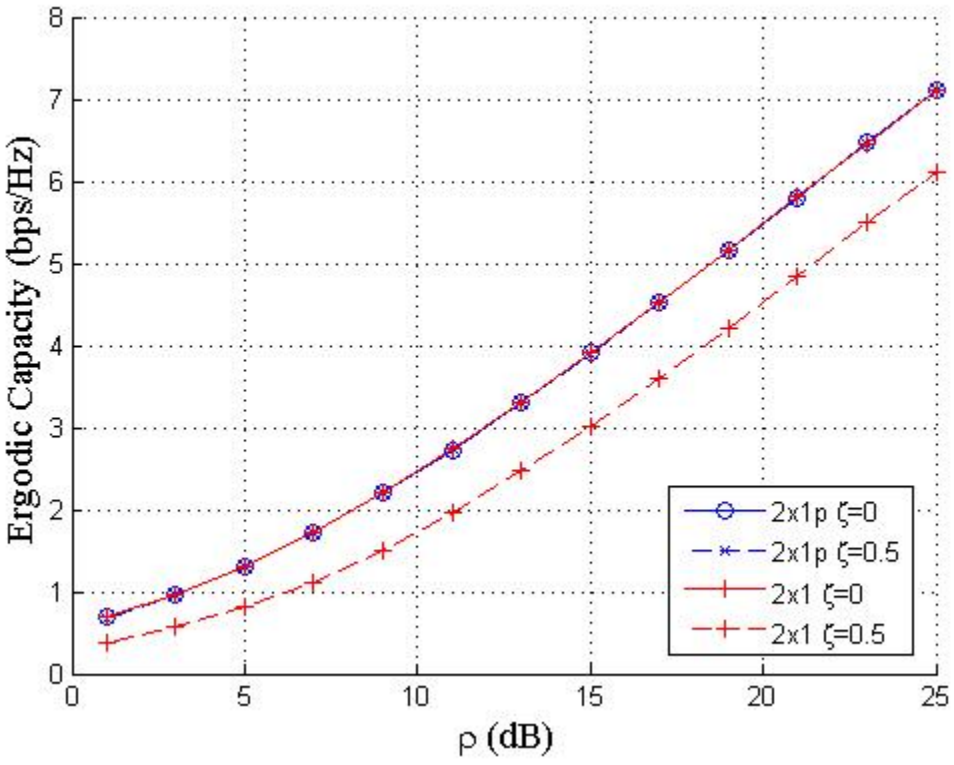


Figure 28. A comparison of downlink capacity for 2×1 and $2\times 1\text{p}$ (EP).

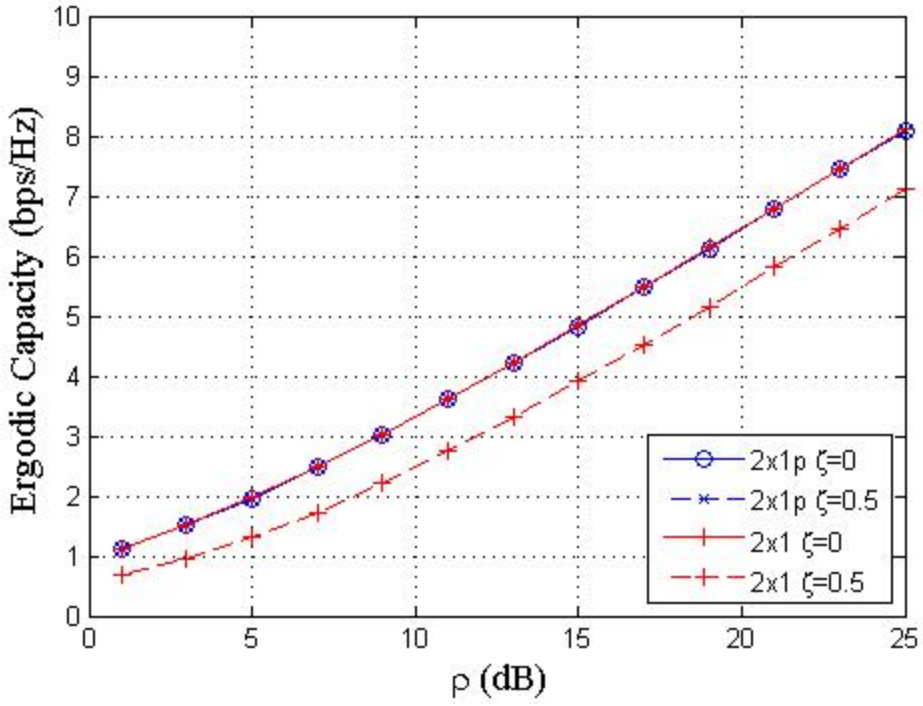


Figure 29. A comparison of uplink capacity for 2×1 and $2\times1\text{p}$ (EP) showing unchanging capacity for the $2\times1\text{p}$ configuration in changing channel conditions.

5. $2\times2\text{p}$

We study only the downlink for the $2\times2\text{p}$ configuration (see Figure 14) compared to the 2×2 configuration, both of which can be configured with and without Alamouti STC.

Without coding, the MRC is used with the downlink receiver with 2 inputs (2×2) or 4 inputs ($2\times2\text{p}$). For the 2×2 configuration, the 3 dB power loss when $\zeta = 0.5$ is a result of the vertical receive antennas. Without coding, the $2\times2\text{p}$ downlink configuration obviously has the advantage of being insensitive to ζ, as shown in Figure 30.

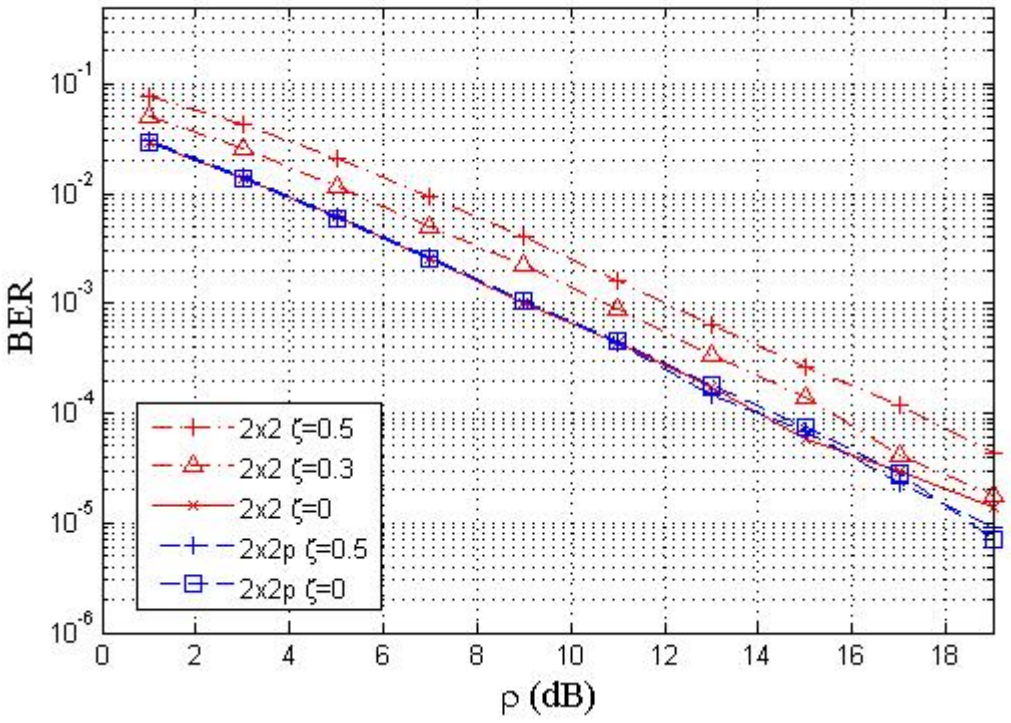


Figure 30. The 2×2 and $2\times2p$ configurations' downlink BER in polarized channels with no STC.

With coding, the 2 by 2 Alamouti STC is used, which is matched to the 2×2 configuration. However, the $2\times2p$ configuration has four antenna receiver outputs on downlink, so the space time decoder is extended to each antenna output. The result is an indifference to the changing ζ of the channel and maximization of diversity gain, as in Figure 31.

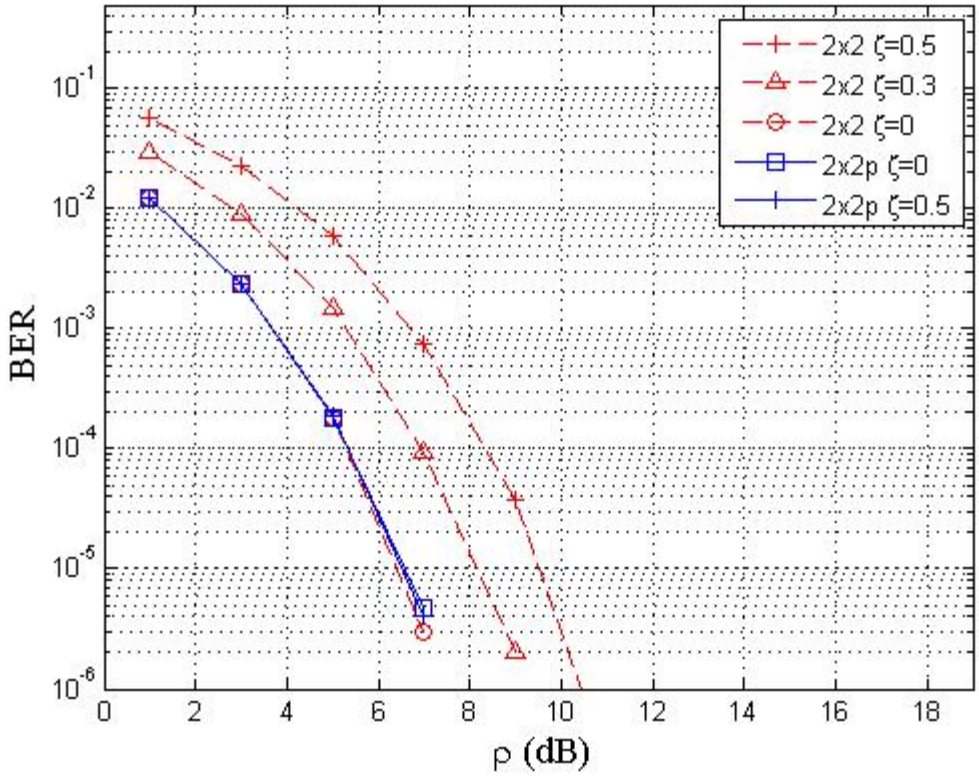


Figure 31. The 2×2 and $2\times2\text{p}$ configurations' downlink BER in polarized channels with Alamouti STC.

As for capacity, we expect that as ζ increases, the $2\times2\text{p}$ configuration will be unchanged in capacity as the horizontal receive antennas capture leaked power. By comparison, the 2×2 configuration has decreased capacity because of this leakage, as in Figure 32. This decrease corresponds to a decreasing squared Frobenius norm of the channel, $\|H\|_F^2$, which will become zero at $\zeta = 1$.

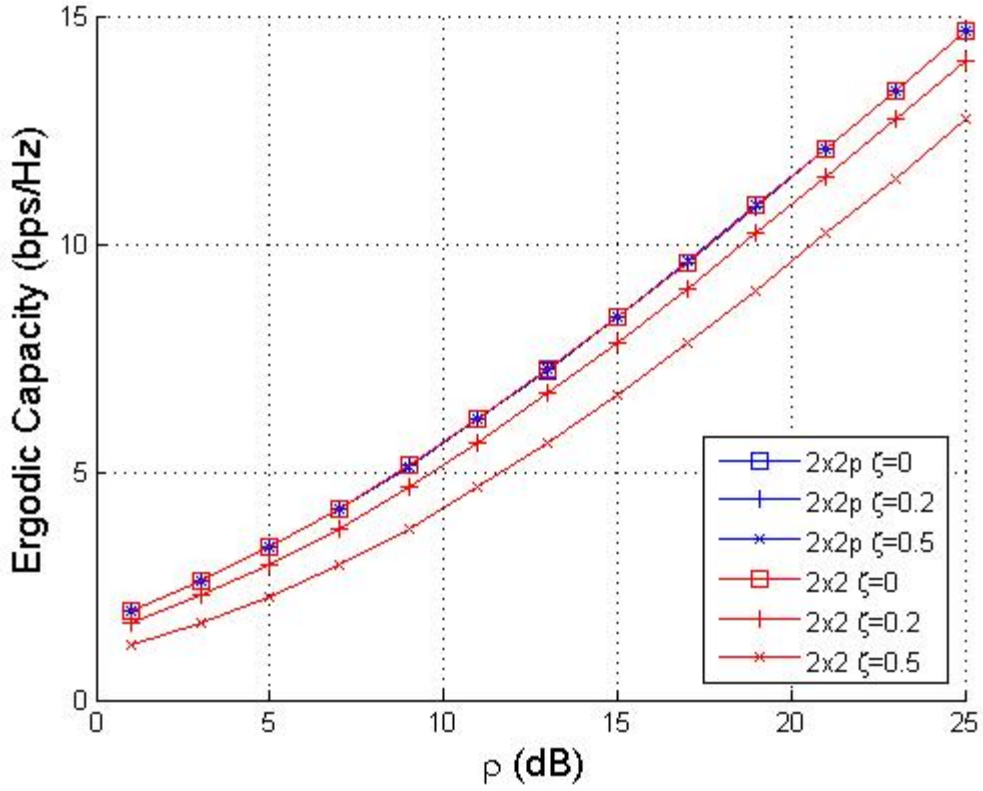


Figure 32. Increasing ζ decreases downlink capacity for the 2×2 configuration while the $2\times 2\mathrm{p}$ maintains the same capacity.

We can study the effect of *spatial correlation* using Monte Carlo methods to compute the ergodic capacity using (23) and (40), as shown in Figure 33. This assumes no CSIT; therefore, equal power is transmitted on the two downlink antennas. Such factors as antenna spacing and the local scattering environment around the transmitter and receiver antennas determine the spatial correlation value [100].

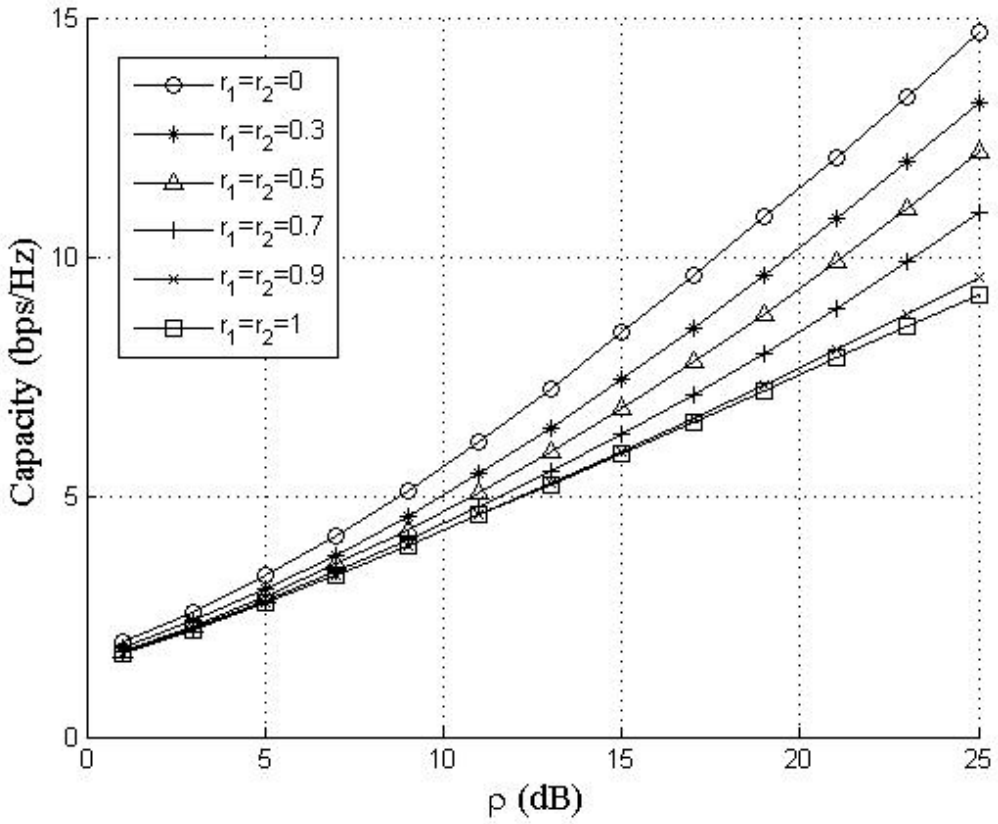


Figure 33. Increasing receive correlation decreases downlink capacity for the $2\times 2\text{p}$ $(\zeta = 0)$ configuration. Identical results for the 2×2 $(\zeta = 0)$ configuration.

As expected, for the hybrid model with $r_1 = r_2 = 1$, the $2\times 2\text{p}$ capacity collapses to the $1\times 2\text{p}$ capacity curve. In this case, the receive correlation always has a negative effect on capacity because it reduces the effective dimensionality of the receiver without increasing the captured power [101]. From Figure 33, it can be seen that the effect is more pronounced at higher values of ρ.

We have seen in this chapter that the polarized channel provides an opportunity for diversity to overcome fading. We developed a hybrid dual-polarized channel model that was used to analyze asymmetric antenna configurations. The dual polarization receivers on downlink had the advantage of capturing transmitted power in all cross-

polarization conditions. The polarization reciprocity concept was introduced and used on uplink. For the simplest configuration, the optimum uplink multiplier to use for differing SNR and channel conditions was derived to give maximum channel capacity. The results show that especially at high ρ, the dual polarization uplink provides *better capacity with the same uplink power* as a single polarization. However, the combined use of Alamouti STC and the polarization reciprocity scheme with the $1\times1\text{p}$ configuration is not recommended. The $2\times1\text{p}$ downlink showed the best diversity gain of 2 whereas the single polarization just lost power. With Alamouti STC, the 2×1 configuration had a diversity gain of 2 even with the reduction in the power of the channel. Adding 2 by 2 Alamouti coding to the $2\times1\text{p}$ configuration ($2\times1\text{p}^*$) *provided the best BER performance*. In the uplink, $2\times1\text{p}$ had equal or better performance compared to 2×1. Capacities for the 2×1 (in uplink and downlink) decreased with channel cross polarization ratio, but $2\times1\text{p}$ capacities remained steady. Similar themes were found in the 2×2 configurations. The $2\times2\text{p}$ downlink BER is insensitive to the cross polarization ratio compared to the 2×2. Also, Alamouti STC greatly improves BER performance through diversity gain. Ergodic capacity of the $2\times2\text{p}$ downlink configuration is also insensitive to the cross polarization ratio. Finally, we confirmed the expected result of decreased capacity with increased received correlation values. These configurations all presented ways to improve the availability of a single communication link utilizing asymmetric polarized MIMO.

V. AVAILABILITY OF UXS NETWORKS

The preceding two chapters showed the many ways to improve the link between two nodes, specifically polarized MIMO. Reviewing the dissertation objective of *maintaining the highest network availability for a MANET with heterogeneous communication capabilities,* we now expand our focus to a system level view of MANET availability.

This chapter starts with a demonstration of adding realistic channel characteristics due to shadowing and its relationship to a communication parameter. A MANET group mobility model that includes the effect of shadowing on each link is then presented. Next, we examine a special case of the MANET, exploration, where only one node is moving and large-scale path loss is simulated. Finally, we examine another special case in which the nodes create formations. In all these cases, the availability metric is used to determine the efficacy of the MANET.

For these models to be relevant to real-world deployment, the relationship between the link availability $\mathrm{p_a}$ and a communication system parameter will be established.

To briefly see how $\mathrm{p_a}$ relates to E_b / N_o, consider a transmission between two nodes. The first node has a fixed location while the second node will be moving. The received power at each node decreases with increasing distance d between nodes and as a function of the path loss exponent n_p. The shadowing path loss effects of this channel are modeled with X_σ, a zero-mean Gaussian random variable with standard deviation σ, as [35]

$$P_r = P_T - L(d_o) - 10 n_p \log\left(\frac{d}{d_o}\right) + \mathrm{X}_\sigma \tag{60}$$

where P_T is the transmitted power, and $L(d_o)$ is the path loss at a reference distance d_o. The received power P_r is used to compute the SNR by

$$SNR = \frac{P_R}{N}, \tag{61}$$

where N is the noise power and is assumed constant.

For digital communications, the probability of bit error P_e of a modulation scheme depends on the ratio of bit energy E_b to noise power spectral density N_o as

$$\frac{E_b}{N_o} = SNR\left(\frac{B}{R}\right), \tag{62}$$

where B is the channel bandwidth and R the bit rate. We now consider the simple case of binary phase shift keying (BPSK) in order to demonstrate the relationship between P_e and E_b / N_o . The bit error probability for BPSK is given by [35]:

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_o}}\right). \tag{63}$$

Figure 34 shows this probability of bit error performance for the BPSK modulation scheme. This P_e is the parameter indirectly related to the link availability $\mathrm{p_a}$.

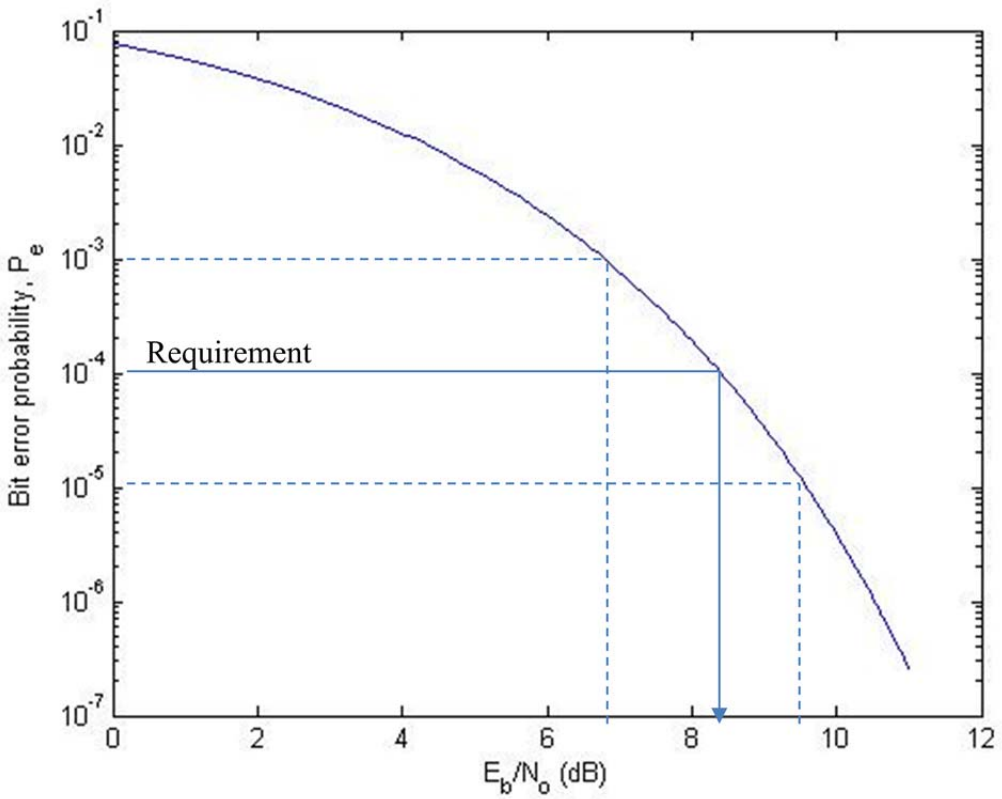


Figure 34. The probability of bit error for BPSK modulation as a function of E_b/N_o.

As an example, choosing a requirement for minimum P_e, e.g., $P_e < 10^{-4}$, then the minimum E_b/N_o, denoted $E_b/N_{o,\min}$ is 8.4 dB using (4). The link is considered available when it meets or exceeds this requirement, which is $\frac{E_b}{N_o} > \frac{E_b}{N_o}_{\min}$. The E_b/N_o received at a certain distance is modeled as a random variable, similar to (60) and is assumed to be Gaussian with non-zero mean, as in Figure 35. The standard deviation $\sigma_{\frac{E_b}{N_0}}$ represents the spread of the received E_b/N_o around the mean, due to shadowing. Since the link is available when it meets or exceeds this requirement, the link probability of availability is represented as:

$$p_a = \int_{\frac{E_b}{N_o}_{\min}}^{\infty} f_{\frac{E_b}{N_o}}\left(\frac{E_b}{N_o}\right) d\frac{E_b}{N_0} \tag{64}$$

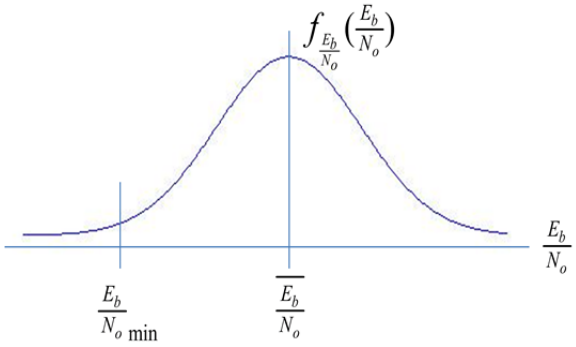


Figure 35. The E_b/N_o PDF is shown with a threshold E_b/N_o minimum.

Now consider the second node moving away from the first node. As the distance between nodes increases, the signal strength decreases, so the mean $\overline{\frac{E_b}{N_0}}$ decreases. As more of the Gaussian density curve is below the unchanging $E_b/N_{o,\min}$ threshold, the p_a will decrease. Conversely, as the second node moves closer to the first node, the p_a will increase. The link probability of availability can be obtained in terms of a *Q-function* as:

$$p_a = Q\left(\frac{\frac{E_b}{N_o} - \overline{\frac{E_b}{N_o}}}{\sigma_{\frac{E_b}{N_o}}}\right). \tag{65}$$

In summary, a communications system parameter E_b / N_o can be related to the link probability of availability p_a based on a requirement of probability of bit error for the link. Having shown that individual link probabilities can be related to real-world communication parameters, we now consider the networks formed by these individual links.

A. MANET GROUP MOBILITY AND AVAILABILITY

Mobile ad hoc networks (MANET) make it possible for nodes to communicate without any permanent infrastructure. The member nodes of a MANET must perform tasks, such as entering and leaving the network, determining neighbors, and providing their own routing. After the network is established in some form, useful data exchange occurs among the nodes with some data traffic pattern. Both the maintenance and data

exchange modes of the MANET require adequate link state. This must be accomplished with dynamic link states that accommodate the mobility of the network's members. This link state is characterized by the signal power at the receiver of the link.

There are many factors inherent to MANETs that collectively cause the link state to change, and only a few are mentioned here. Each node is constrained by transmit power, and this is directly proportional to received power at the other node. The MANET nodes are dynamic and have a relative motion among themselves. As distance increases, the received signal power decreases according to a path loss exponent. Another factor is fading, which can be large- or small-scale. One large-scale fading example is shadowing. As explained in Chapter II, it is caused by the interaction of a signal with the local environment and has a log-normal distribution about the distance-dependent mean [35]. The commonality of these factors is their decrease in the received signal power, generally degrading the link state of each node.

UGV teams are an example application of MANETs, where each system (robot) is considered a node in the network. The received signal power is influenced both by distance between the nodes and their environment; we focus first on distance. To analyze distance among UGV teams in a MANET, we use a description of the movement of each individual UGV as a node. These relative positions over time can be represented by mobility models, which describe the movement of a group of unmanned systems (nodes) in a MANET. A discrete instant in time of a mobility model describes the absolute positions of nodes from which signal strength is found. In summary, the MANET is composed of many nodes each with a link state changing in time. Nodes with a deficient state are unavailable to the network, and we will quantify the metric of network availability. The mission (movement) of the UGV team is described by a mobility model, which will determine network availability in time.

Small teams have spatially dependent mobility [102] hence a group mobility model is appropriate. In general, less emphasis has been placed on group mobility models because they are difficult to analyze except in simulation. The use of group mobility models via simulation is explored in [22], [23]. These use the RPGM model (introduced in the next section) but only in 'groups of groups'. These simulations revealed a broad

relationship that mobility influences the routing protocol. There have been no published attempts at relating communication system performance *analytically* to the RPGM mobility model. We begin with the step of relating the RPGM model with a single group to the metric of network availability. It is network availability that is related to protocol performance but is easier to compute. Full network simulations can determine protocol performance and are able to account for factors, such as data traffic, routing algorithms, and more complex environmental factors. Nevertheless, it is hoped that the analytical work that follows can help explain the relationship between mobility and protocol performance.

1. Reference Point Group Mobility Model

The Reference Point Group Mobility (RPGM) model [103] describes the physical positioning of each node with reference to a group center. The group center, C, moves along an arbitrary path with time, as in Figure 36. Each node, $n_i, i \in \{1 \ldots N\}$ in the group of N nodes is assigned a reference point, which is defined with respect to the group center $C(t)$ by the vector $\overrightarrow{R_i}$. At all times, the node stays within its mobility circle, a radius of r_i from its reference point. At each time step (in a simulation) the node's reference point moves with the group motion vector $\overrightarrow{m_g(t)}$. Figure 36 shows one node's mobility circle at two points in time. The position of each node $p_i(t)$ is found by vector addition of a node motion vector $\overrightarrow{m_i(t)}$ to the reference point:

$$p_i(t) = C(t) + \overrightarrow{R_i} + \overrightarrow{m_i(t)}. \tag{66}$$

The node motion vector is a random vector providing uniform distribution of position across the mobility circle. The node motion vector is independent from the node's previous location. Any node velocity can be represented by varying the time step. Specifically, the diameter of the mobility circle divided by the time step is the maximum node speed that can be represented (relative to $C(t)$). The *absolute* maximum node speed that can be represented would add the group motion vector.

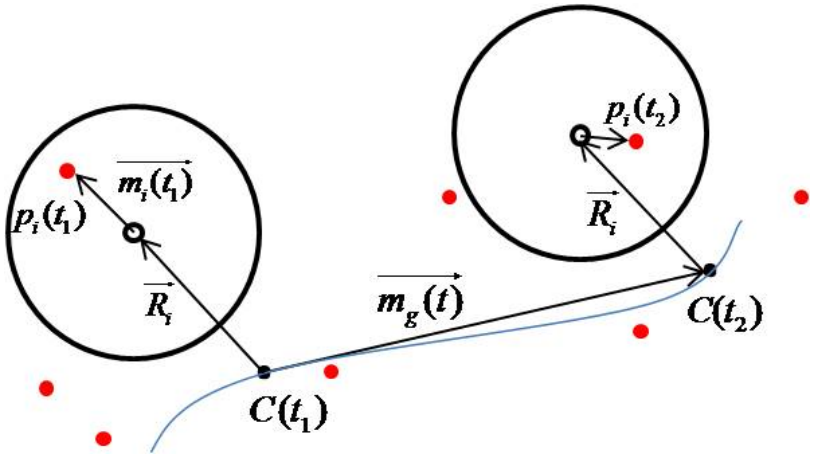


Figure 36. RPGM model with red dots indicating random node positions within a mobility circle. After [103].

The movement of small teams of UGVs can be described by the RPGM model. There are two control topologies of interest that can be represented within the RPGM: clusterhead and mesh. In the clusterhead topology, each node communicates only with the clusterhead node, similar to a sensor network that funnels data to a central node. This could also represent a manned control system that is controlling many UGVs. The clusterhead topology can also describe a larger, more capable robot that communicates directly with a group of smaller robots. Such a centralized team might use a TDMA MAC protocol. The development focuses on this clusterhead control topology and its data traffic model.

The mesh topology allows any number of connections between nodes of the team. This can represent a more autonomous team of unmanned systems exchanging data among other nodes during a mission. With substantial data flow among all nodes, such a decentralized team may use a CSMA MAC protocol.

2. Communication within the RPGM model

Having described the motion of UGVs, we now consider the data flow between these member nodes of the MANET. Unmanned systems that are sensing the environment need to transmit high data rate information, such as video. They also

transmit telemetry and receive control information at a lower data rate. In operation, the telemetry and control information is typically of higher priority than the sensor data. When the communications link is weak, the telemetry, control, and management information must be supported, but the sensor information may tolerate temporary interruption. The communication state diagram for each node is therefore described by Figure 37.

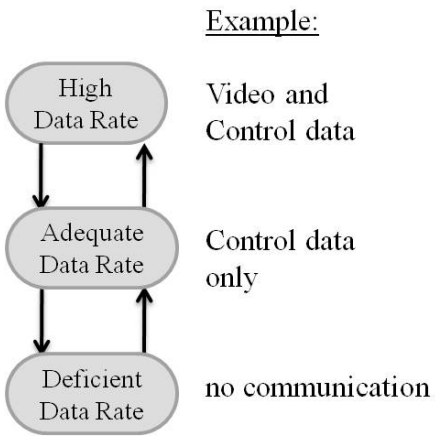


Figure 37. The communication state diagram for each node of a MANET.

Each unmanned system has a threshold data rate (minimum) that determines the difference between the states. These thresholds can also be expressed in terms of signal power at the receiver. Maximum data rates or capacities have a well understood relationship to the SNR and bandwidth. Assuming that bandwidth and noise are fixed, a received signal power at the receiver corresponds to a data rate.

The IEEE 802.11 standard generally serves to demonstrate that the received signal strength is aligned with a data rate. Table 1 listed the modulation coding schemes (MCS) supported in IEEE 802.11a. In the standard, the MCS is dynamically adjusted based on signal strength. For very strong signals, higher data rates can be supported.

As an example, if 8 Mbps were required for control data and 20 Mbps were required for sensor data, there would be two thresholds among the MCS. One between MCS 0 and 1 separating “deficient” from “adequate” communications, corresponding to a minimum sensitivity of -81 dBm. Similarly, a threshold between MCS 3 and 4 would separate “adequate” from “high data rate” communications.

The listed sensitivity values of Table 1 are the minimum signal levels as measured at the antenna connector of the receiver [51]. These minimum signal levels ensure a certain maximum packet error rate.

The signal levels can be converted to a distance based on the widely used path loss equation with a chosen transmit power and path loss exponent [35]. So the data rate corresponds to a signal level and a distance. From this, a communications model based on the clusterhead topology is shown in Figure 38, where the two threshold data rates are now represented by distance via range rings from the center node.

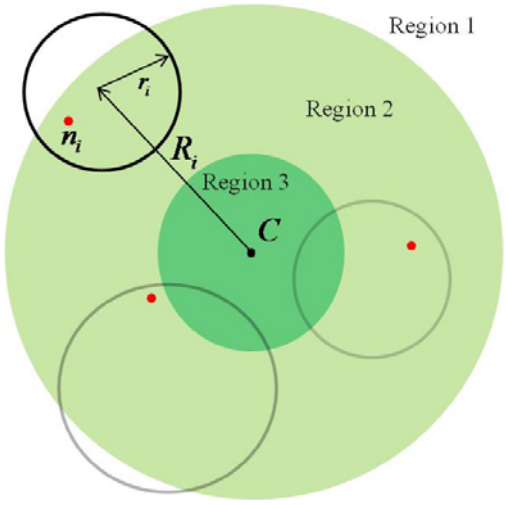


Figure 38. The RPGM communications model showing different regions where communication states exist.

The two threshold values create the range rings that separate the field into three regions. Within the physical region 3, the signal exceeds the upper signal threshold, supporting the 'high' data rate. Region 2 supports the 'adequate' data rate and is separated from region 1 by the lower signal threshold. Region 1 has a 'deficient' data rate. As the node moves in its mobility circle, it may traverse a different region, changing its communication state.

The RPGM model describes the absolute motion of a group of nodes. To represent the clusterhead topology, we place a node at position C. For communications

analysis within the group of nodes that travels together, the $\overrightarrow{m_g}$ vectors can be set to zero. The random motion of the nodes within their mobility circles can continue in simulation time. This preserves the value of interest to the communications analysis: the *relative* motion and positions of the group of nodes. Using this communication model, we next determine each node's coverage. The node coverage κ is the probability that the received signal power P_r exceeds a threshold γ as the node moves in its mobility circle

$$\kappa = \Pr[P_r > \gamma] . \tag{67}$$

We will consider only the lower threshold value in the following discussion.

a. Node Coverage with No Fading

A simplistic, and possibly accurate enough, approximation of node coverage can be illustrated using Figure 38. This approximation would disregard fading and assume the communication regions are based only on range. This approximation assumes that in region 1, the node is not connected to the center node and is therefore disconnected from the network.

From the RPGM model, we assume that each node's position within its mobility circle is uniformly distributed. So the fraction of each mobility circle that is not in region 1 is the node coverage. This value of node coverage based only on range κ_r is

$$\kappa_r = \kappa_{r2} + \kappa_{r3} = \frac{A_2}{A_{total}} + \frac{A_3}{A_{total}} , \tag{68}$$

where A_j **,** $j \in \{2,3\}$ is the area of the mobility circle that is within region j , and A_{total} is the area of the mobility circle. These areas are calculated using the overlapping circles method detailed in the Appendix.

b. Node Coverage with Shadowing

Shadowing has a strong effect on node coverage. As described in Chapter II, the path loss model that includes shadowing uses a RV to have a mean path loss for a

chosen distance between transmitter and receiver. We now extend that one-dimensional distance (shadowing) to a two-dimensional area (coverage).

To consider the effect of shadowing on communications, we first describe the distance between transmitter and receiver. Referring to Figure 38, we consider the case of one point fixed at the center and the other moving within a mobility circle. The distance R from the fixed center to the center of the mobility circle is *not* the mean of the distances to the points in the mobility circle. The probability density function (PDF) of this distance d is the distance PDF. This distance PDF is not symmetric about R but skewed and has a mean greater than R. As shown in the Appendix, an approximation of the distance PDF for a mobility circle of radius r is

$$f_D(d) = \frac{2}{\pi r^2}\sqrt{r^2 - (d-R)^2} \quad . \tag{69}$$

This approximate distance PDF is symmetric and closely resembles the actual distance PDF when the value $R \gg r$. As shown in the Appendix, Figure 52, this increases the radius of curvature R_c making the area B_1 a better approximation of the area $B_1 + B_2$.

To determine the node coverage, we note that the mean value of the received signal changes with distance. The standard deviation of the shadowing also increases with distance. We, however, use the simplifying assumption of a fixed standard deviation σ, as in [104]. The threshold value γ, of course, does not change with distance. The value κ_c is then expressed as an integral of the product of the position PDF and the probability that the received power at that distance is greater than the threshold,

$$\kappa_c = \int_{R-r}^{R+r} f_D(d) Q\left(\frac{\gamma - \mu(d)}{\sigma}\right) \delta d \quad . \tag{70}$$

In (70), the mean of the received power is a function of the transmitted power P_t, the loss at the reference distance R and the path loss exponent n_p:

$$\mu(d) = P_t - L(R) - 10n_p \log\left(\frac{d}{R}\right). \tag{71}$$

Unfortunately, there is no closed form expression for (70), and straight line approximations of the position PDF have lengthy solutions. A fast approximation of κ_c can be obtained by a single calculation at the distance R by

$$\kappa_c \approx Q\left(\frac{\gamma - \mu(R)}{\sigma}\right) = Q\left(\frac{\gamma - (P_t - L(R))}{\sigma}\right). \tag{72}$$

To verify this, a Monte Carlo method is used to give an estimate of the coverage probability, using [35]

$$P_{\hat{r}} = P_t - L(R) - 10n_p \log\left(\frac{d}{R}\right) + \mathrm{X}_{dB}, \tag{73}$$

where the effect of shadowing is represented as a zero mean Gaussian random variable X with standard deviation σ, and so the estimate

$$\hat{\kappa}_c = \Pr[P_{\hat{r}} > \gamma]. \tag{74}$$

In Figure 39, it is shown that the value of $\hat{\kappa}_c$ closely tracks κ_c over values of γ for a typical scenario.

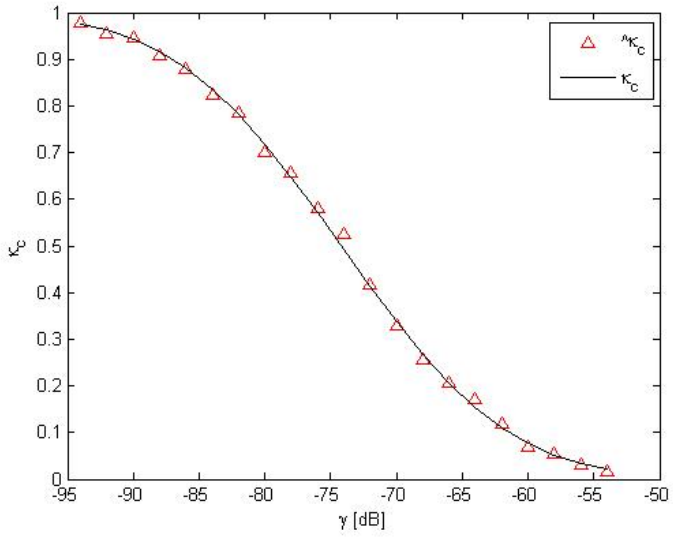


Figure 39. Coverage probability as a function of threshold value for R = 50 meters, r = 20 meters, n_p = 2, f = 2.542 GHz, and σ = 10 dB.

c. Node Coverage with Composite Method

The composite method is based on the observation that in certain regions, the shadowing is negligible. Using the method in the Appendix, we designate the area of the mobility circle that is within region j as A_j, $j \in \{1,2,3\}$, as in Figure 40. Then in region 3, assuming the signal power is sufficient to disregard shadowing, that fraction A_3 / A_{total} for each mobility circle is considered as not experiencing shadowing, with $\kappa_{c3} = 1$. Similarly, in region 1, assuming the signal power is sufficiently weak that shadowing is disregarded, then $\kappa_{c1} = 0$. Then for each mobility circle, the node coverage is a weighted sum:

$$\kappa = \sum_j \kappa_{cj} \kappa_{rj} \tag{75}$$

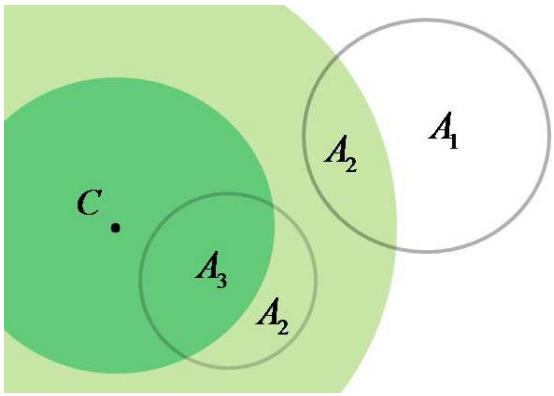


Figure 40. Mobility circles overlaid on regions to compute coverage by the composite method.

Up to this point, we have shown that the probability of coverage for a single node in shadowing can be computed in a variety of ways; next, this is applied to a network of nodes.

d. Network Availability

The node coverage probability gives a measure of the 'connectedness' of one node over one link. The computed values of κ represent a *percentage of area* (mobility circle) in which the received signal power is sufficient for communication, which is above the lower threshold. Because the position is uniformly distributed over this area with time, the values of κ are also a *percentage of time* above the lower signal threshold. These nodes can be connected to each other by different topologies.

For the clusterhead control topology, each of the nodes is only connected to the center node. With regard to the lower signal threshold, each value of κ is the probability that each node is connected to the center node. A network metric from the node coverage is the probability of availability of the cluster. It can be simply found as the product of all the node coverages [48]

$$\mathrm{P_A} = \prod_{n=1}^{N} \kappa_n \tag{76}$$

and represents the probability that all nodes are connected.

The mesh control topology permits nodes to communicate with all other nodes in the group. Although the mesh network probability of availability also uses the κ from each node, the calculation is substantially more complex [48], as outlined in Chapter II. For either topology, if the higher threshold is used to calculate each node's κ value, the P_A value would represent the probability that the MANET was in the 'high data rate' state.

Having provided some analytical answers to availability of a MANET, consider what if the UGV needed to move beyond the range of adequate data rate, into region 1 of Figure 38. We next consider a special case of the MANET for this mission of exploration.

B. MANET FOR EXPLORATION

The realistic use of UXS, such as UGVs, takes place in many challenging military environments. In these different environments, short communication ranges may, for example, prevent an unmanned ground vehicle from completing its mission. In this section, we consider the means of communicating data from an exploring UGV (or robot) back to a GCS in these challenging environments.

1. Introduction

The common need of these exploring robots is a sufficient data rate for real-time sensor data, such as video to be available to the operator. For navigation and sensing, the robot could be autonomous or remotely operated, but the information gathered by the robot should be immediately available to the user. While two-way communication is considered, the greatest challenge is the large data flow from the robot to the control station.

Increasing this challenge is the lack of supporting communication infrastructure in these environments. In response, there have been a number of different approaches to create the communications systems that are needed. Wireless sensor network (WSN) nodes have been studied to support the communications of mobile nodes but have limited energy (short transmission range, sleep cycles, and limited bandwidth). As a result, WSN

cannot support the data rate required for video transmission. Throwboxes are wireless nodes that act as relays in the store-carry-forward paradigm of disruption tolerant networks [105] but cannot fulfill the real-time need of exploring robots. Cabling of a robot provides wired communication and can certainly provide robust communications support. It is commonly used in harsh communications environments (e.g., rescue robots and commercial pipeline inspection). However, the cable restricts movement of the robot to a maximum distance and may not be practical for a multi-robot team.

Another answer within the existing research is to deploy more robots. Teams of robots have been used to maintain connectivity among fixed positions [106], [107]. Another researcher allows for dynamic role changes of robots from exploring to supporting communications [108]. Some projects designate certain robots of a team to be used exclusively for communications support of another robot [109]. More recent trends are towards hybrid strategies, considering simultaneous communications and mobility constraints as a robot control problem [110], [111]. Although teams of robots provide redundancy and can explore most areas faster than a single robot, there are drawbacks. Beyond the obvious expense and operational support of more robots, the human operator could be overwhelmed by monitoring multiple video streams. There is also lost system utility in having a robot equipped with sensors relegated to a communications relay role only. Regardless, this relay role is required and extensively described in literature as cooperative communications. Laneman et al. [112] analyzed the relay schemes of *amplify-and-forward* and *decode-and-forward* and derived criteria for their selection. Boyer et al. [113] show that amplify-and-forward outperforms decode-and-forward for the single user communication. In total, the purpose of these techniques and schema is to maintain high availability across a communication network.

Because our proposal involves wireless relay nodes in a fixed network, the problem resembles WSN node placement. For this case, the solution must be robust, scalable, support (not interfere with) robot movement, and not under-utilize additional robots. Our contribution will be to introduce the deployment and use of stationary network nodes as a cooperative communication method for maintaining system-level availability for an exploring robot.

2. Models

To determine the proper use of a communication relay node, we first need to adequately model robot movement in realistic environments and compute received signal strength during this movement. In our simulations, we assume finite bounded, convex exploration areas with differing environmental characteristics. These areas remain static, and there are no other moving objects. We assume the robot perfectly knows its own position, for the purpose of mapping. The objective is for the robot to search the area until a goal object is found. However, in the simulation, we can omit the goal object and observe the robot's full exploration of the environment.

Rather than computing received signal power for specific cases, propagation path loss is used instead to abstract away from specific radios, transmit power, antenna gains, and modulations. The model used to describe the RF propagation for the indoor environments is the attenuation factor model [114]. In this propagation model, the primary ray between the transmitter and receiver is a straight line, and the cumulative losses from obstructions along the line are summed to obtain a mean path loss $\bar{L}(d)$

$$\bar{L}(d) = \bar{L}(d_0) + 10 n_{SF} \log(\frac{d}{d_0}) + \sum P \quad . \tag{77}$$

Here d is the distance between transmitter and receiver, d_0 is the reference distance, n_{SF} is the single floor path loss exponent, and values of P are the path attenuation factors for various obstructions of the primary ray. Empirical studies have provided the dB loss (P) at certain frequencies for walls of varying thicknesses and composition [35]. For example, we use P values that typify concrete blocks (13 dB) or sheetrock (2 dB) walls as applicable. As the robot changes position, the simulation recomputes this mean path loss.

The model used to describe the RF propagation for the outdoor environment is the well-known COST 231 Walfisch-Ikegami model [115]. This statistical model is appropriate for urban areas and is valid for distances as short as 20 meters. Distinguishing between line of sight (LOS) and non-line-of-sight (NLOS), the model has seven input

parameters that account for numerous propagation effects. We simulate the radios closest to the ground (4 and 1 meters) in a three-story building, medium-sized city.

Five representative environments are used similar to those in [116] as shown in Figure 41.

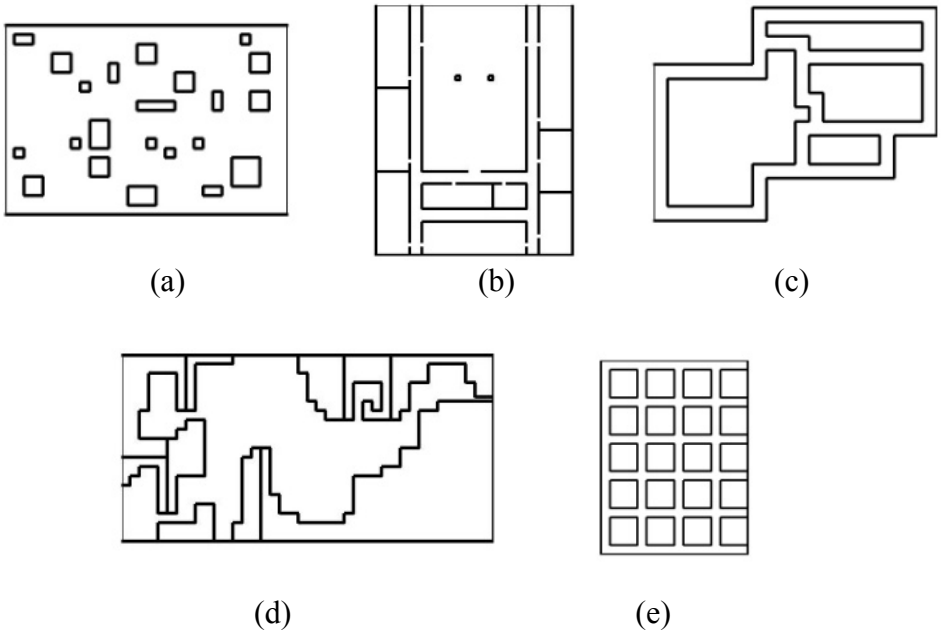


Figure 41. Maps used for the simulation experiments: (a) unstructured, (b) office, (c) corridor, (d) cave and (e) urban environment. After [116].

Table 2. Five environment descriptions for RF propagation models

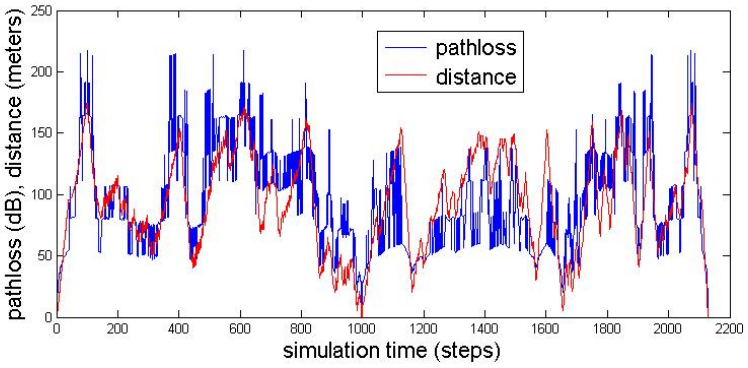
Name	Dimension (m)	Scale (m)	Walls (dB)	Propagation Model
Unstructured	150 x 100	5	13	AF
Office	47 x 60	1	2	AF
Corridor	100 x 75	5	13	AF
Cave	400 x 200	5	20	AF
Urban	320 x 420	10	n/a	COST 231- WI

We next describe the motion of the exploring robot. Starting from a position adjacent to its control station, the robot explores the environment at each time step of the simulation. Among the myriad choices for robot exploration strategies, this research compares an online depth first search (DFS) agent [117] and a frontier search.

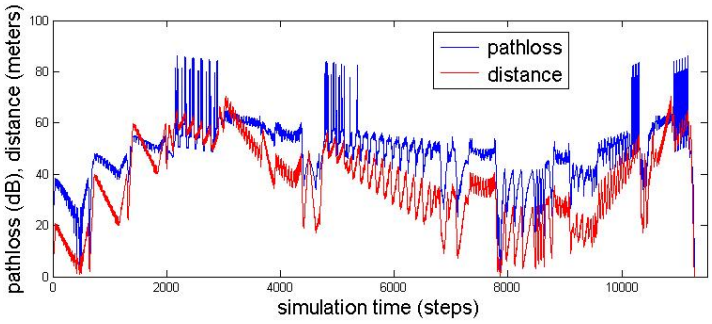
For the DFS, the agent (robot) does not have any map information at the start, and we limit actions from the current state to movement in the four cardinal directions, with priority order: left, up, right, down. The agent has a sensing distance equivalent to one "step," discovers obstacles, and builds its map as it moves. It stops moving after reaching a randomly placed goal on the map. The algorithm will direct the agent to completely explore the environment and return to its starting position if it does not find the goal.

For the online frontier search, the robot must detect open areas and move toward these frontiers. The only difference from DFS is the priority order of movement, which now depends on measurements of the distance to walls in the four directions. It is assumed that the maximum wall sensing range is ten times the goal sensing distance. The robot moves in the direction of the farthest wall sensing distance that it has not already explored. Equal distance "ties" are resolved in the priority order stated above, for reproducibility. This drives the robot into open spaces as opposed to following a left wall (DFS).

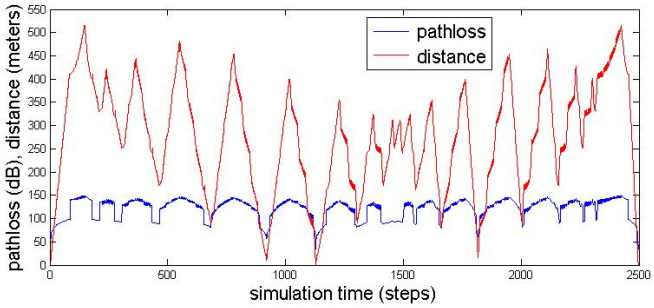
Since the information from the robot must reach the control station, the simulation computes the Euclidean distance and propagation loss during the time along the exploration path. These are shown as a function of time for the three environments in Figure 42.



(a)



(b)



(c)

Figure 42. Distance and path loss from base station over time for: (a) unstructured, (b) office, and (c) urban environments with $\bar{L}(d_0) = 0$ dB, $n_{SF} = 2.76$.

In Figure 42(a) and (b), it is observed that the propagation loss is correlated to distance as expected from (1). But the propagation loss also has sudden "jumps," corresponding to the robot moving behind obstacles. Also, specific robot movement as it sweeps through a room can be observed, as in Figure 42(b), witnessed by the sawtooth distance and path loss traces. The robot's exploration in the urban environment, Figure 42(c), even brought it back near to the control station (distance ~ 0) before backtracking to continue exploring. Last and most important, we observe that the sheer number of

obstacles creates very large propagation losses, which potentially preclude the receiver from functioning. If these propagation losses are too high, the data flow from robot to control station is lost. The undesirable result would be to limit the exploration area.

The receiver at the base station has a minimum received signal power for supporting a particular data rate [118]. Propagation losses exceeding a maximum threshold force the receiver to accept a slower data rate. This propagation loss threshold value T is dependent on a number of factors including transmit power, receiver type, modulation type, error correction, interleaving, antenna gains, etc. For this study, we denote a maximum value of T that supports a particular data rate.

The astute observer can see that aspects of this model are simplistic and do not account for fading, multipath propagation, more realistic robot mobility and sensing, or specific radios. Incorporating all these factors in more complex simulations would certainly yield more accurate values of the path loss. The general trace of path loss and our observations, however, are applicable to multiple environments. Consequently, we next describe the device and a method for the placement of communication relays nodes that will remedy the extreme path loss.

3. Network Nodes

We propose to use network nodes to provide an economical wireless communication between a robot and its control station. We will refer to the deployed communication relay nodes as simply nodes. A node functions to receive, amplify, and re-transmit a signal(s). Being battery powered, it need only function for the duration of the exploration time. Since it has no mobility, this device is initially carried and then deployed by the robot as it moves. Our initial work is with a single robot carrying >1 node. The first node that is deployed relays data between the robot and its control station. If a subsequent node is deployed, generally, it relays to a previously deployed node.

If the environment map were known, the coverage could be calculated as a heat map, and node locations could be planned for placement in “cool’ areas. However, for unknown area exploration, no map is known a priori, so it is important to note that *the optimum placement of a node for area coverage cannot be precomputed*. Since

communications should support, not restrain the mission of exploration, the robot can prioritize its path planning only on exploration and nodes are placed as needed on the robot path.

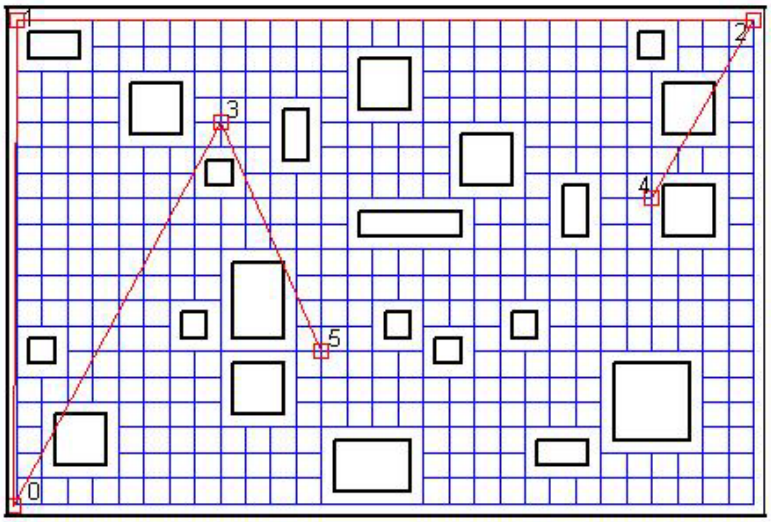
Two other characteristics of the node are important: channel capability and disposability. Should the node be able to be used by more than one robot simultaneously? It would be the subject of a multi-robot study to determine how many separate channels might need to be relayed by one node; we assume one channel in our simulations. A two channel node, if used, would simultaneously relay data from two robots to another node or the mutual control station. Also, a second robot could use a single channel node previously deployed by, but not currently being used by another robot. This added node capability certainly influences robot team exploration.

Also, whether the nodes are disposable or retrievable is of practical concern. Retrievable nodes imply the backtracking of a robot over previously explored terrain, which is possible but not guaranteed in path planning. In our simulations, the extent of backtracking is found to be highly dependent on the environment and a chosen single robot search algorithm. With our focus on communications performance, our treatment of the problem views the nodes as disposable.

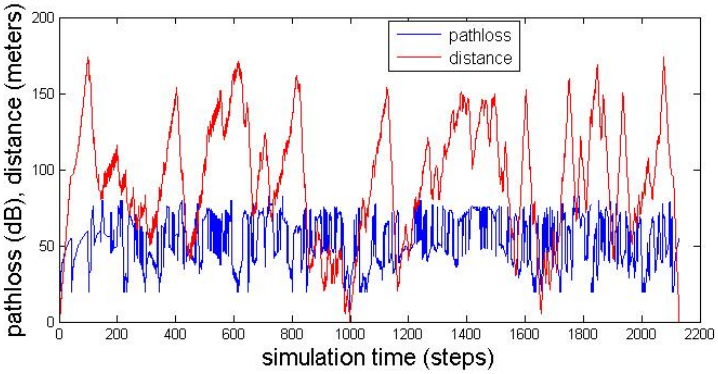
We begin with the basic idea that a new node is deployed when the propagation loss exceeds T . This, however, means that at the current location, T is exceeded, and a node dropped here must compensate for the propagation loss from the previous location to the current location. Assuming that the node has the identical radio as the robot, this could be achieved by increasing its radiated power (e.g., transmit power or antenna gain). From the simulations, it is observed that the propagation loss exceeds T in two manners: gradual or abrupt. A gradual crossing of T results from the first two terms in (77) when there is no change in obstructions; for example, the robot is simply moving down a corridor away from the control station. When the robot turns a corner beyond a dense obstacle, there is an abrupt increase in propagation loss (e.g., 26 dB beyond two concrete walls). Hence the additional radiated power of a node that is more capable than the robot radio would have to account for these worst case scenarios. Since the environment can be estimated but not predicted, there is no guarantee that a node with added radiated power

will perform adequately for the unknown environment. The alternative, given a node with an identical radio and radiated power as the robot, is to back up to the previous location with propagation loss $< T$ and then deploy the node. We follow this simple policy in our simulations. While this does indeed force additional movement on the robot for communications support, there is a distance between the robot antenna and the onboard node antenna before deployment. Possibly matching this to the step size for communication sampling would allow the robot's antenna to have a path loss $> T$ while the node when deployed would be just below T . Regardless, this *"back up and deploy"* for nodes is a policy that is adaptable to any environment as the path loss is sensed and requires no advance planning nor map knowledge.

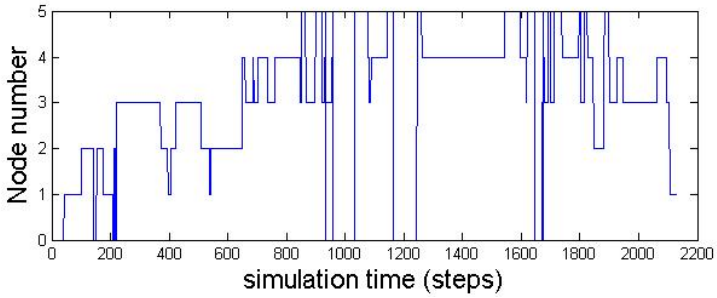
In summary, the chosen movement for a robot that deploys a node is to "back up and deploy," but this implies a movement penalty (that could possibly be minimized). Although the nodes are placed in a geographic location, their deployment is also a function of time. The propagation loss to the current node is monitored over time, and when it exceeds T , the "back up and deploy" policy is used. All nodes then have a 'path loss distance' that is the same, although the physical distance between nodes is different. Intuitively, the lower the threshold T , the more frequently nodes are deployed. The red lines in Figure 43(a) indicate the network formed from the deployment of nodes for a robot exploring the 'unstructured' environment. Node zero is the control station and deployed nodes are numbered sequentially. The path loss between the robot and its closest communication node is shown in Figure 43(b). In contrast to Figure 42(a), these path losses are now below T . Figure 43(c) shows the robot's current communication node, which changes throughout the simulation. We discuss the policy for selection of the current communication node in the next section.



(a)



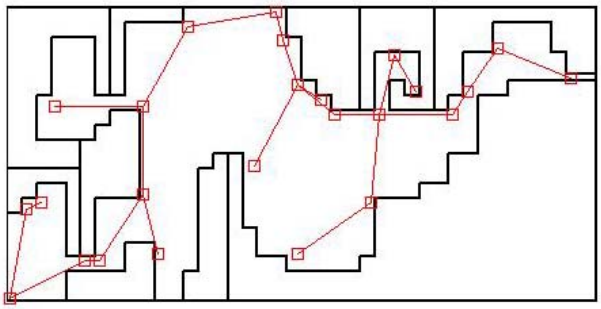
(b)



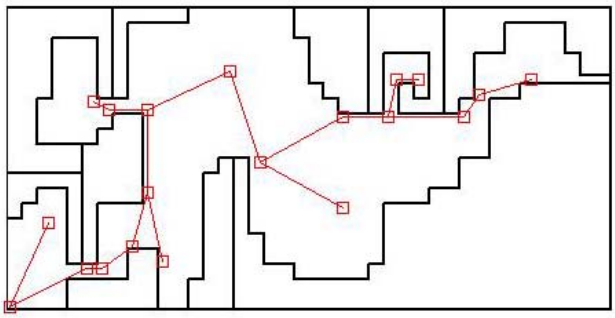
(c)

Figure 43. (a) Fully explored environment (track in blue) with red squares at location of nodes. (b) Path loss to current node and distance to control station. (c) Identification number of current node during the simulation with node 0 as the control station and $\bar{L}(d_0) = 0$ dB, $T = 80$ dB, $n_{SF} = 2.76$.

We briefly mention the contrast between the two exploration strategies. The DFS and frontier exploration strategies have different paths through the environments, placing nodes in different locations. We observe the qualitative difference in Figure 44 by noting that the DFS strategy typically places nodes near to walls, whereas the frontier exploration strategy places some nodes centrally in large open areas.



(a)



(b)

Figure 44. Fully explored cave environment with red squares at location of nodes using the (a) DFS strategy and (b) frontier strategy. The control station is at lower left and $T = 50\,dB$.

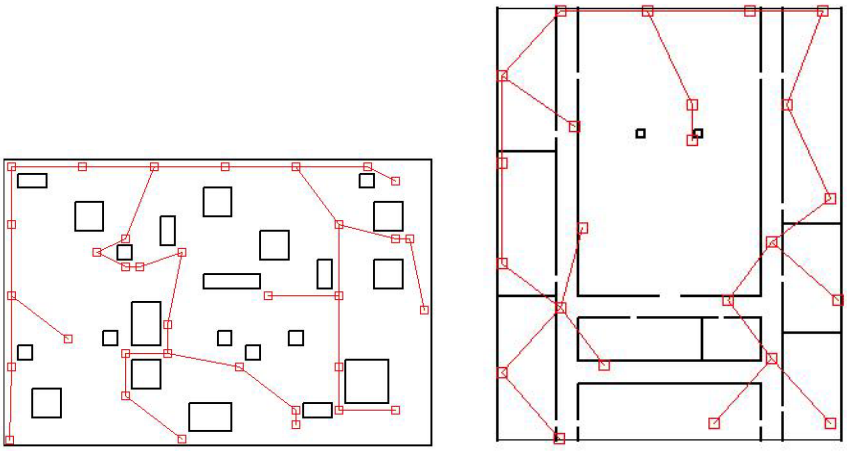
The irregularly shaped inlets of a non-convex map, such as the cave, reveal that neither strategy is always superior in coverage of such complex unknown areas. One obvious quantitative measure distinguishing strategies is the number of nodes deployed. In this particular case, frontier exploration was more efficient, using only 20 nodes, compared to DFS using 24. A network wide measure is developed in the next section.

This fixed placement of nodes is in sharp contrast to strategies using moving nodes for communication relay. Mobile nodes can provide a continually optimized placement, whereas fixed nodes are placed "of necessity" along the robot path. We acknowledge fixed placement strategies are less than optimum but can still be effective.

4. Network Communication Path

We next look at the network structure. As the robot moves away from the control station, it has increasing propagation loss that may reach the value T . Exceeding the T value, the robot executes the "back up and deploy" policy, and the communication network topology is modified to insert the node between the robot and the control station. The robot now communicates multi-hop to the control station through the node. The robot continues exploration and may deploy more nodes in a similar manner, creating a linear topology. Also, as the robot explores, it may switch to different, previously deployed nodes that have a lower propagation loss and potentially then add more nodes. This will change the communications path from robot to control station and may not utilize all of the deployed nodes. We describe nodes that are part of the current communication path, and the links to them, as 'active', otherwise they are 'passive'. If the robot backtracks to retrieve all nodes, there would not be any passive nodes in the network.

As a new node is deployed, *the resulting network topology is a tree with the control station at the root*. The robot deploying the nodes is the only means for increasing the network size. This network can be represented by a graph $G = (V, E)$, where each newly deployed node is a vertex V , and the communication link between nodes, control station and robot is an edge E . Because new nodes are not deployed unless there is a threshold crossing, and there is no other node available, the new node (vertex) has only one edge to the previous node (vertex) and one edge to the robot (vertex). Because of this policy, all active nodes are exactly two-connected, producing a graph that is a tree. Examples of these networks are shown in Figure 45 with exaggerated low T values to deploy many nodes.



(a) (b)

Figure 45. Tree networks formed from deployment of nodes with the control station (root) at lower left. Red squares are nodes, with (a) $T = 40\,\text{dB}$, scale $= 5$ and (b) $T = 32\,\text{dB}$, scale $= 1$.

The robot is utilizing only one communication path at a time to transmit its data to the root (control station). As a consequence, routing and medium access control (MAC) are simplified because each node is within range of a maximum of two other active, known neighbors. Finally, although the entire graph is connected by the means of its construction, the most important connectivity issue is the currently used, active branch of the tree between the robot and the control station.

We next ask the question: what limits the number of nodes that can be used by an exploring robot? Beyond the physical capacity of the robot, there are certainly interference issues for any number of similar devices sharing wireless spectrum. These concerns are lessened by the deployment algorithm, which tests for other nearby nodes in the unknown environment. We assume that the self-interference among nodes is solvable and focus on the theory of their use. Up to now, we have assumed that the communication links between nodes (and the robot and the control station) had an availability of 1. This depended upon the mean of the path loss being less than a threshold value. This would imply that any number of nodes could be used if each link in the path met these criteria. Certainly there is an increase in time delay as each node adds its processing delay to the lengthening hops of the active communication path; such

latencies are unavoidable. We, however, note that continually adding more nodes has another effect on the total communication path availability.

The communication path loss is more fully modeled with shadowing and multipath fading [119], and we consider shadowing here. Shadowing is a variation in received signal strength due to surrounding environmental clutter. Shadowing can be represented as a zero mean log normal random variable (RV) X with a variance σ^2 that depends upon the environment [35]. Typically, the shadowing RV is applied to the signal power; but to consistently speak in terms of path loss, we express the shadowing RV as additive to the mean path loss (in dB):

$$L(d) = \bar{L}(d) + X. \tag{78}$$

The availability $\mathrm{p_a}$, which is the probability that the path loss is lower than the threshold path loss value T, is then expressed using the cumulative distribution function of the normal random variable as [46]

$$\mathrm{p_a} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{T} e^{\frac{-(l-\bar{L}(d))^2}{2\sigma^2}} dl. \tag{79}$$

In (79), $\bar{L}(d)$ is the mean path loss from the appropriate model of Sec. II.A. This equation gives a more realistic availability of the link between the robot and the current communicating node. This value will continually change as the robot explores, generally decreasing with increasing distance. Once a criteria for the deployment of a new node is met, there is a fixed distance between the newly deployed node and its branch node. Therefore, the ith deployed node has a fixed value for availability

$$\mathrm{p_{ai}} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{T} e^{\frac{-(l-\bar{L}(d_{fi}))^2}{2\sigma^2}} dl, \tag{80}$$

where d_{fi} is the fixed distance between node i and its branch node. These values of $\mathrm{p_{ai}}$, where $0 \le \mathrm{p_{ai}} < 1$, can be visualized as edge weights on the node graph.

The series of participating links (via active nodes) have a total path availability. This multi-hop path availability P_A is [120]

$$P_A = p_a \prod_{i \in \{I_{active}\}} p_{ai}, \tag{81}$$

where I_{active} is the set of participating links. Because each additional path link is multiplying previous values, and all individual probabilities are less than one, there is no increase, only a decrease in the path availability with additional active links. This answers the query about a limit to the number of nodes: the inclusion of the probabilistic propagation loss can bring the cumulative communication path availability below an acceptable value. Considering this as the active system-level metric, we next examine the relationship between exploration strategy and network availability.

Resuming our inquiry of exploration strategies, we are equipped to gain quantitative results. The exploration strategies were tested in the four indoor environments, varying the threshold value to deploy different numbers of nodes, with results in Table 3. The value of the multi-hop path availability was averaged across the exploration time.

Table 3. Network availability for two exploration strategies in four environments

Environment	T (dB)	No. of nodes		$\bar{P}_A$	
		frontier	DFS	frontier	DFS
cave	50	19	24	0.147	0.120
	60	15	14	0.524	0.492
	70	14	13	0.512	0.499
	80	11	11	0.558	0.543
	90	8	8	0.317	0.229
office	50	2	3	0.551	0.500
	60	1	1	0.840	0.818
	70	1	1	0.948	0.960
	80	1	1	0.985	0.995
	90	0	0	0.995	0.995
unstructured	50	15	18	0.283	0.201
	60	12	12	0.703	0.491
	70	10	9	0.774	0.662
	80	4	5	0.548	0.677
	90	4	4	0.876	0.786
corridor	50	13	14	0.289	0.302
	60	11	11	0.493	0.531
	70	6	6	0.483	0.386
	80	3	4	0.525	0.621
	90	3	4	0.890	0.810

Using this network availability metric, the frontier exploration strategy generally excelled in open and irregular spaces. For an equal number of nodes deployed, the frontier strategy-deployed network had higher network availability in the cave and unstructured environments. In environments characterized by fewer robot path choices, the DFS and frontier had approximately equal performance. Notably, frontier was superior at the lowest values of T **(**$T < 60\,dB$**)** using fewer nodes in all indoor environments. The corridor environment provided few choices in robot path to differentiate the performance of the two strategies. The office environment was small enough that it was fully "covered" by the control node at $T = 90\,dB$ and no additional nodes were deployed. These results from our coarse simulation seem to slightly favor the frontier exploration strategy.

The criteria used previously for deployment of the nodes was whether the propagation loss (in dB) had exceeded a threshold value T. But knowing the network metric for system-level availability in (80) allows for an incremental optimization by the

placement of each node, which can account for shadowing on the communication path. By now choosing a minimum availability p_m and computing (79) at each step, each newly deployed node will create a network edge with $p_{ai} \geq p_m$. These active edges (hops) form the product of (81). The previous use of a threshold value T was equivalent to using a value of $p_m = 0.5$; we now consider individual hops with higher values of p_m.

Using the minimum availability criterion, the value of P_A is shown for an example robot exploration in Figure 46.

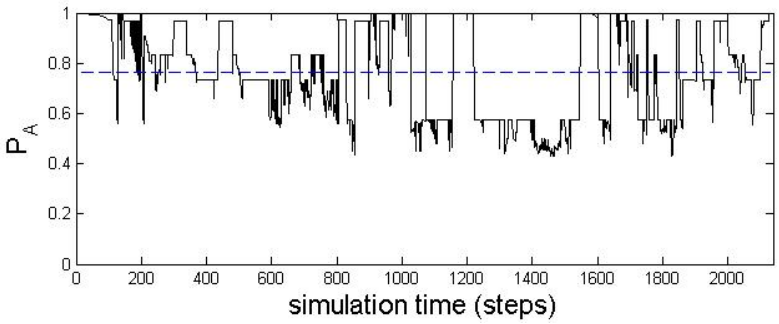


Figure 46. P_A and $\overline{P}_A$ (dashed line) during robot exploration in unstructured environment with $T = 80$ dB, $\sigma = 10$ dB, $p_m = 0.74$.

It is possible for P_A to increase over time, as shown in Figure 46. This happens when, during exploration, the robot moves closer to its current communication node or switches to a communication path with fewer hops. Deploying a node adds another term as (80), so it results in a decrease in P_A. Decisions to deploy a node are based on p_a, not P_A, and P_A only reflects the currently active network condition.

To review, five factors directly influence P_A: distance from robot to the closest node, number of hops in the current communication path, distance between nodes in the current communication path, radio performance (threshold T) and environmental factors (obstacles, path loss exponent, and shadowing variance). Considering the latter three factors unchangeable, we next look at the factors that are within our control.

The robot movement should not be constrained by its communication system requirements, so although 'distance from robot to closest node' is malleable, we decline to modify the robot movement control to stay close to a node. As a result, our focus is on the possibility of decreasing the number of hops in the current (active) communication path.

The importance of decreasing the number of hops is justified in (81). Recall that a new node is deployed when $p_a < p_m$, forcing these links of the communication path to have a very similar value of p_{ai}. The robot to node hop is the last p_a in the sequence of multiplied values. This last link has lower path loss and, therefore, a higher p_a value than the other links, which minimizes its effect on the product P_A. Each of the other hops has a greater effect on P_A than the robot to node hop. Because all branches of the tree have nodes placed in this way, the robot can use the current branch, or other, adjacent branches if available. So the robot should search for any path that has fewer hops back to the control station, thereby increasing P_A.

The algorithm used up to now is passive such that only if $p_a < p_m$ will the robot search for another node with a better p_a; otherwise, it will deploy a node. This is a policy of waiting until it has run out of signal power from its current path, then seeking alternatives. Implementing an aggressive policy adds a continual check: if the robot can communicate with another node and that node has fewer hops to the root than the current node, it will switch nodes. An example is shown in Figure 47, where this aggressive policy has increased the $\overline{P}_A$ by 9.1%.

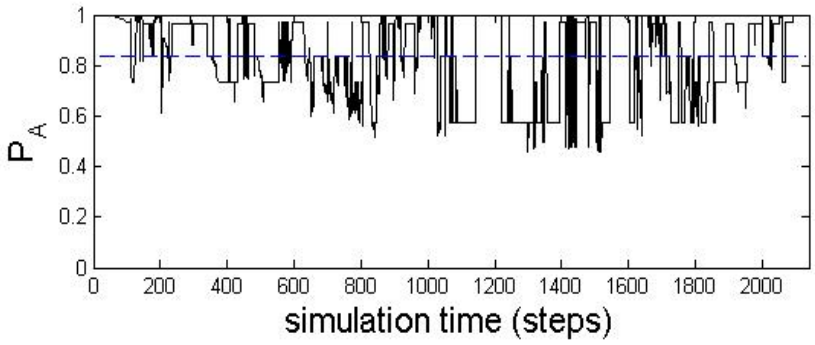
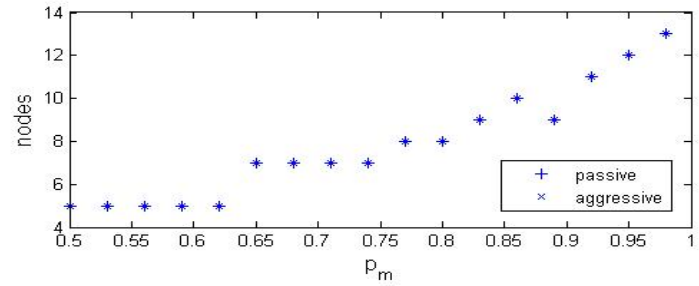
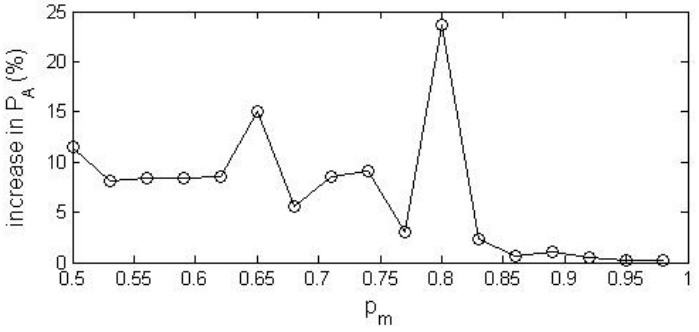


Figure 47. P_A and $\overline{P}_A$ (dashed line) during robot exploration with aggressive routing in the unstructured environment with $T = 80$ dB , $\sigma = 10$ dB**,** $p_m = 0.74$

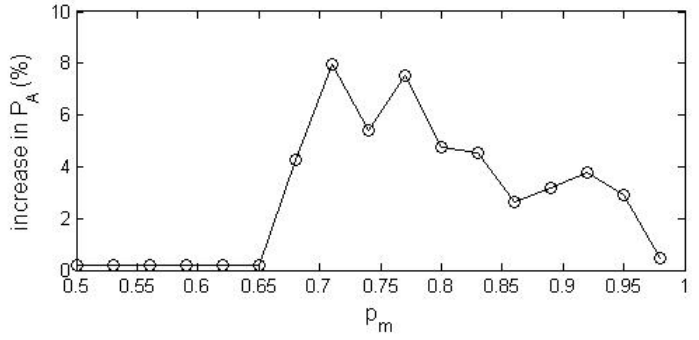
Next, we consider this policy across variations in p_m and some environments. The aggressive routing policy has no effect on the number of nodes deployed, as seen in Figure 48(a). Here, the value of the deployment threshold p_m is increased for each simulation to force an increasing number of nodes to be deployed (similar result for the other environments). But the aggressive policy of continual testing for the lowest P_A path to the root has a cost in terms of energy spent searching for other nodes. The example of Figure 48 included 80 routing switches to lower hop count nodes. This cost may be reduced since the location of deployed nodes is globally known, and only nodes within close range might be queried. Another cost is the potential disruption of real-time data flow as routing changes. However, when discovered, a second path to the root can also be viewed as redundancy that improves network robustness. Above all these considerations, we see the end effect of the aggressive policy via an increase in P_A compared to the passive policy, as shown in Figure 48(b-d) for fully explored environments. The improvement of the aggressive routing policy over the passive routing policy is expressed as a percentage improvement, found by using the average of the P_A values over the total simulation times.



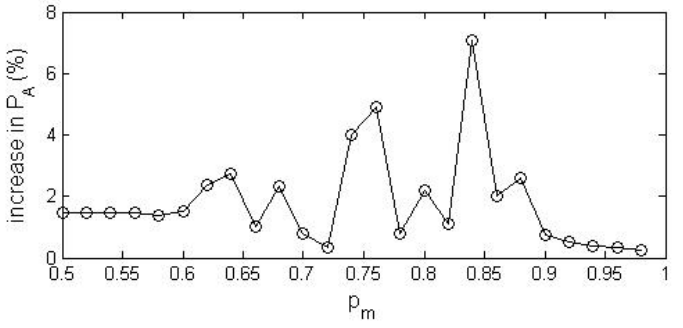
(a)



(b)



(c)



(d)

Figure 48. The (a) number of nodes deployed in the unstructured environment (others similar), (b) percentage increase in $\overline{P}_A$ due to aggressive routing policy in unstructured environment, $T = 80$ dB (c) office, $T = 60$ dB and (d) urban, $T = 100$ dB.

The aggressive policy's greatest benefit is in the unstructured environment with a moderate number of nodes (<10). This scenario provides numerous opportunities for node switching during exploration of the unstructured space. In the office environment, low values of p_m resulted in the deployment of only one node in the simulation, reducing opportunities for node switching. In the urban environment, the crossing street structure rarely exposed the robot to multiple nodes, also reducing the chances for node switching.

Since the map of the environment is unknown in future cases, we can only make general recommendations for the use of the aggressive routing policy from our limited testing. The use of the aggressive policy might best be implemented as a function of environment and the number of nodes already deployed being most important in less structured environments and with a moderate number (1<nodes <10) of nodes deployed. But in all cases, no decrease in P_A was found from implementing the aggressive routing policy.

We have shown how to compute the network availability using this special case of a MANET with an exploring robot and how to improve that availability. The deployed network had a tree topology, and the participating relay nodes formed a linear chain between the exploring robot and the GCS. The next section describes other unique topologies and their effect on availability.

C. AVAILABILITY WITH FORMATIONS

Another special case of the MANET occurs when a particular topology is maintained for a time, which we denote as formations. Formations are common in military scenarios as UXS move in concert to utilize sensors, provide overlapping protection, or transit an area. We can further consider the case of a fixed position formation, such as when WSN nodes are deployed. We use the metric of availability applied to four different extensible formation models, which might be used in a MANET as in Figure 49.

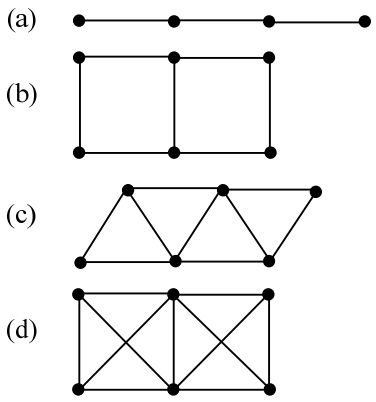


Figure 49. Formation models with interior vertices having $\delta_e(G) = 2, 3, 4,$ and 5 : (a) chain (b) ladder (c) truss and (d) full truss.

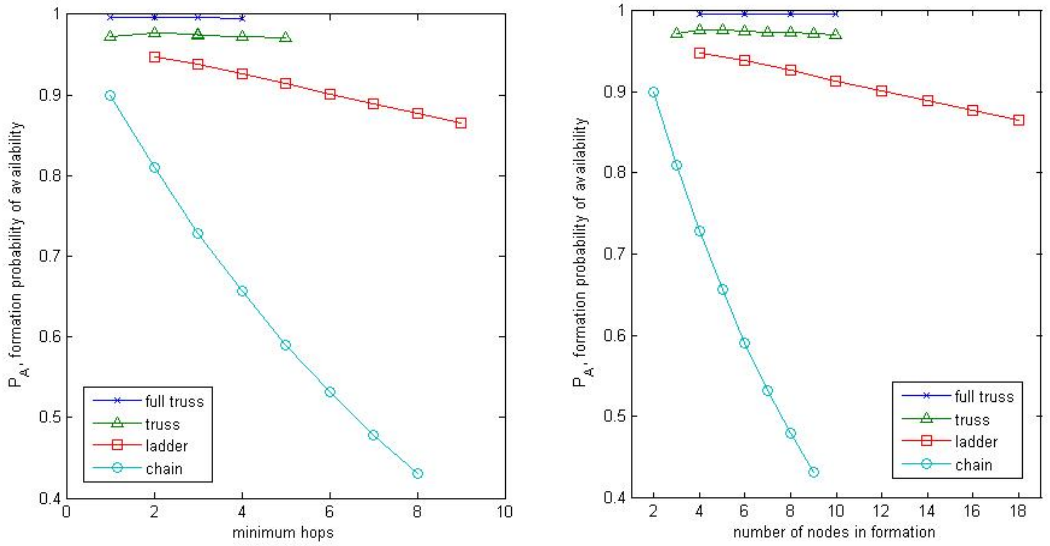
As shown in Figure 49, they resemble a chain, ladder, truss, and full truss, which correspond to all "interior vertices" having an edge connectivity $\delta_e(G) = 2, 3, 4,$ and 5, respectively. These formations could be extended in a direction and can be of variable size depending on the number of nodes used. One measure of the size of these formations is a 'minimum hop' metric, which is found as the minimum number of hops to reach the maximum distance across the formation. In graph theory, this corresponds to the maximum of the eccentricity of all vertices in G . For Figure 49(a-c) the minimum hop metric is 3; for Figure 49(d), the minimum hop metric is 2.

For each of the formation models, as they were extended, it was assumed the p_a of each edge was equal to 0.9. Artificially held to this level, it allows a fair comparison among formations. For each extension in size, the number of nodes, number of edges, minimum hop metric and resulting network availability P_A (using the algorithm of Figure 6) were calculated, and the results are shown in Table 4.

Table 4. Formation P_A of four different formations (Figure 49) with varying size.

	n	e	min. hop	P_A
chain	2	1	1	0.9
	3	2	2	0.81
	4	3	3	0.729
	5	4	4	0.6561
	6	5	5	0.5905
	7	6	6	0.5314
	8	7	7	0.4783
	9	8	8	0.4305
ladder	4	4	2	0.9477
	6	7	3	0.9389
	8	10	4	0.9265
	10	13	5	0.914
	12	16	6	0.9016
	14	19	7	0.8894
	16	22	8	0.8774
	18	25	9	0.8656
truss	3	3	1	0.972
	4	5	2	0.9769
	5	7	2	0.9763
	6	9	3	0.9753
	7	11	3	0.9742
	8	13	4	0.9731
	9	15	4	0.9721
	10	17	5	0.971
full truss	4	6	1	0.9958
	6	11	2	0.9958
	8	16	3	0.9957
	10	21	4	0.9955

The data of Table 4 is plotted in Figure 50, showing that the network availabilities are highest where there is greatest path redundancy, that is, a higher δ_e generally yields a higher P_A. The ladder and full truss formations have a linear decrease in P_A with an increase in the minimum hop metric. The chain formation P_A decreases exponentially with distance, a result of its $\delta_e = 2$ vertices.



(a) (b)

Figure 50. Formation availability P_A as a function of (a) minimum hop metric and (b) number of nodes in formation, for four formation types.

Reversing this analysis, a MANET can also be synthesized. Figure 50 shows the superiority of the full truss formation to the ladder formation, which is achieved by the same number of nodes, but with increased edge connectivity. Formation nodes may change their degree of connectivity based on changing propagation conditions; in this case, the full truss formation would represent an upper bound on P_A, and the chain formation would represent a lower bound on P_A. It should be remembered that although the chain formation P_A decreases rapidly as the number of nodes increases, it may span a required distance in an effective manner as could be measured by another metric, such as physical distance covered per nodes deployed. In general, MANETs can move within these formations keeping values of $p_a \leq 0.9$ and have a P_A at or above the value shown in Figure 50.

Another use for formations is determining some placement criteria for *stationary* nodes. The previous section displayed an extending chain formation with fixed p_a. Instead, choosing a threshold P_A, a number of choices for the quantity of nodes and link

availability for fixed nodes are available. Again, keeping link availability above a value p_a always yields a minimum P_A. Changing values of p_a are shown in Figure 51.

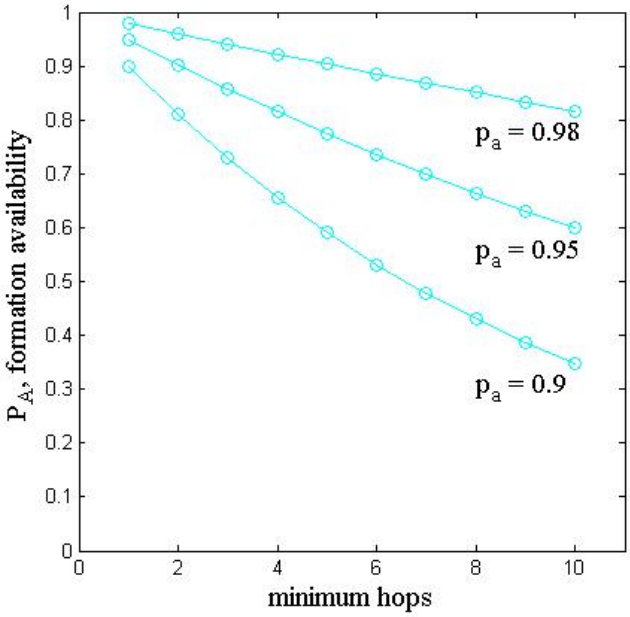


Figure 51. The effect of improving link availability on network availability of the chain formation.

These chain formations were used in the exploring scenario of the MANET.

A system-level view of maintaining high availability for a MANET was shown in this chapter. Such effects as shadowing influenced probabilistic link availability. This effect was incorporated into the MANET availability mobility model and the special case of an exploring robot. The influence of the link availability was also shown for the special case of MANET formations.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

An important and emerging need in unmanned systems is the networking of dissimilar systems. The research first considered some root causes of errors in wireless transmissions, namely the channel characteristics of shadowing, fading, and polarization. Graph theory provided the means of translating the communication links among the UXS to a mathematical representation that included these probabilistic effects of the channel. To evaluate potential changes to the MANET represented by the graph, the metric of availability was introduced, along with its computation. The dissertation objective was to investigate and develop schemes for *maintaining the highest network availability for a MANET with heterogeneous communications capabilities*.

To maximize the availability of a single link, the use of asymmetric polarized MIMO was proposed. The performance here was measured by the metrics of BER and link capacity. Using a hybrid polarized channel model, extensive simulations looked at the performance of a number of asymmetric antenna configurations. The optimum uplink gain multiplier was derived as a function of SNR and the channel cross-polarization ratio. The benefit of STC in these different configurations was demonstrated.

By considering the MANET at the system level, the various movements of UXS nodes of a MANET were studied, and the effect of these movements on the MANET availability were quantified. The general case of the RPGM mobility model showed more realistic UXS movement compared to a random waypoint mobility model and incorporated the multiple data rate needs of the UXS. The RPGM model was overlayed with the effect of shadowing for a clusterhead formation for computing the link and MANET availabilities.

A special case of UXS mobility was exploration with a single UXS. Its MANET was characterized by a linear topology of fixed and mobile nodes. The large-scale path loss effects in five environments which included indoor and outdoor were examined in the research. A simple rule for the deployment of fixed location communication relay nodes was proposed. This rule yielded tree network structures, and then the availability of

the network was derived. Improvements to the MANET availability by means of link availability choice in the environments and aggressive routing were detailed.

The potential for other formations of MANET was outlined in the final section of research. The availability metric was examined for four extensible formation types, using a common link availability for comparison.

A. SIGNIFICANT CONTRIBUTIONS

The network availability is a wide measure of MANET performance, being defined as the probability that there exists a path between any two nodes in a network, whether direct or through a number of connected nodes. We presented a modified depth first search (DFS) method using complete state enumeration for the exact calculation of this value. This robust method assumed a graph theory description of the MANET with probabilistic edge weights.

To consider a single link, we created a unique hybrid polarized channel model for the purpose of analyzing asymmetric antenna configurations. The model isolated the spatial correlation and polarization correlation while also using an absolute reference to study the effect of channel cross polarization ratio. Using the principle of polarization reciprocity, we derived the optimum uplink gain multiplier to be used for the channel cross-polarization ratio and SNR combination. The main result *is increased uplink channel capacity with the same uplink power* using dual polarization configurations under certain conditions.

We proposed a number of unique asymmetric antenna configurations and analyzed them with and without Alamouti STC. The results of simulations showed that the layering of these configurations with Alamouti STC combined with the effect of polarization diversity yielded improved performance compared to single polarization configurations. An example is the $2\times 1\mathrm{p}^{*}$ configuration, with diversity gain improvements as the channel cross polarization ratio increased. To maintain a BER of 10^{-4}, the $2\times 1\mathrm{p}^{*}$ configuration had up to 9 dB lower SNR than the single polarization configuration in the polarized channel. As was presumed, the $1\times 1\mathrm{p}$, $2\times 1\mathrm{p}$, and $2\times 2\mathrm{p}$

configurations remained unchanged or improved performance (BER and capacity) compared to similar single polarization antenna configurations as the cross polarization ratio increased.

Now considering the broader system-level network, the results of a full network simulation can take time and are very specific to one scenario. Our results give a simpler, first-order measure of network performance. To accomplish this, the realistic path loss effects were applied to scenarios that are typically encountered by UXS that compose a MANET: group mobility and exploration. In contrast to simulation approaches to computing network performance, this analytical approach to availability has not been applied to the Reference Point Group Mobility (RPGM) model before. The method presented in the dissertation showed that with the complex, yet realistic movement of the group, the network availability with shadowing could be computed. The method was supported by the geometrical analysis. The method can also be used in a simplified way if shadowing is absent.

The special case of an exploring UXS MANET was also analyzed to determine an effective way to construct the network. This case occurs when one UGS needs to go beyond the effective communication range of the controlling node. It showed a simple and effective method for the deployment and use of stationary network nodes as a cooperative communication path for maintaining system-level availability for the exploring UGS (robot). One outdoor and four indoor environments were created and an appropriate path loss model was applied to the two environment types. The simulations all used propagation path loss to abstract away from specific radio characteristics and concentrate on the wireless channel effects. Our analysis of the path loss of a robot exploring in these environments reasoned for the policy of "back up and deploy" for the fixed position MANET nodes. This causal policy is effective in the *unknown* environment when placement *cannot* be optimized. Combining shadowing with path loss, an expression for the multi-hop path availability was derived. Analysis of factors influencing this showed that an aggressive routing policy can increase availability without incurring the cost of using additional nodes.

Another special case of the MANET was formations. Using graph theory to realize them, these four extensible formation availabilities were computed using an efficient algorithm. The graphical results can be used to analyze and synthesize MANET formations for their network availability.

C. FUTURE RESEARCH

The growing field of unmanned systems research suggests a number of extensions to topics presented in the dissertation.

The work in this dissertation considered the availability measure, also termed as the all-terminal reliability [47], which is a wide measure of network connectivity. Other measures, such as two-terminal reliability may be considered for prioritized communications among a subset of the MANET nodes. Further, the computation of the availability by complete state enumeration had a computational burden that is exponential in the number of edges, similar to combinations visualized as a Pascal's triangle. This research considered up to 25 edges for small MANET teams. The approximation of networks with more edges might be accommodated by representing each row of the Pascal triangle as a probability mass function and then thresholding because the multiple edge failures have decreasing influence.

The validation of the polarization reciprocity scheme with physical antennas could explore the correlation effects more thoroughly. This is because antenna mutual coupling effects were not included in our simulations and these decrease capacity. The uplink capacity for 2x1p was only simulated using equal power and could be extended by simulating polarization reciprocity with different combinations of uplink receivers. Similarly, the 2x2p uplinks (with and without STC) could be studied.

The research expressed a relationship between a communication parameter and the probability of availability. This assumed, for instance, that many of the common error correction mechanisms are not in place. An experimental comparison of the availability of various MCS using a network simulator could be compared to our link availability results.

The exploring robot movements in the simulation environments were thorough, but greatly simplified for the purpose of analyzing the communication link. Real robot exploration missions would use more sophisticated exploration strategies and simulation environment, and the effect on network performance could be investigated. Also, the use of multiple robots for exploration was not pursued due to time constraints, but could be examined.

Due to resource constraints, the modeling of the robot mobility model and exploration were done in simulation. Their validation by experimentation with real mobile robots and fixed communication nodes would verify the effectiveness of the (theoretical) availability approach.

The generic application of MANET formations has many interesting possibilities. The formations synthesis problem must consider sensor range and communication range, which seems more relevant to the proliferation of WSNs. The mobile "stitching" of these fixed nodes appears synergistic with various geometrical sensor patterns on the WSN nodes. With moving nodes, an analysis with a network simulator, such as NS-2, could determine the relationship between network availability and the network measures of a simulator output.

As with many research interests in MANET, the cross-layer solutions appear to be a fruitful domain. We observed this with aggressive routing in the exploring robot scenario. Practical exploring robot systems may require applications that span delay tolerant network concepts.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX. OVERLAPPING CIRCLES METHOD

This appendix presents the calculation of the probability that a uniformly distributed node position within a circle of radius r is within the boundary of a circle of radius R_c. The centers of the two circles are not co-located, as in Figure 52. The calculation is in two parts, corresponding to regions B_1 and B_2. The arbitrary x-axis joins the centers of the two circles.

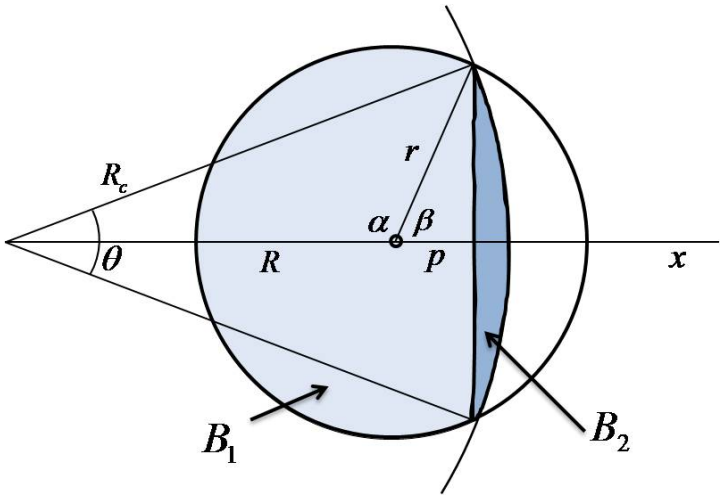


Figure 52. Calculation of uniformly distributed node probability with radius r within boundary of circle of radius R_c.

For the region B_1, the uniform distribution of the node position simplifies the calculation of the marginal density to

$$f_X(x) = \int f_{XY}(x,y)dy = \int_{-\sqrt{r^2-x^2}}^{+\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy \tag{82}$$

where the origin is at the center of the circle of radius r. This reduces to

$$f_X(x) = \frac{2}{\pi r^2}\sqrt{r^2 - x^2} \quad |x| \le r\,. \tag{83}$$

The value of p is defined next. From the triangle with sides r, R, and R_c, the angle α can be found from the law of cosines,

$$\alpha = \cos^{-1}\left(\frac{r^2 + R^2 - R_c^{\,2}}{2rR}\right). \tag{84}$$

The angle β is the supplementary angle of α. Then the distance p can be found by

$$p = r\cos\beta\ . \tag{85}$$

The PDF can then be integrated to find

$$\Pr[B_1] = \int_{-r}^{p} f_X(x)dx. \tag{86}$$

The probability that the node is in region B_2, $\Pr[B_2]$, is found by first finding θ and then finding the area. The angle θ can be found from the law of cosines,

$$\theta = 2\cos^{-1}\left(\frac{R^2 + R_c^{\,2} - r^2}{2RR_c}\right). \tag{87}$$

The area of the circular segment B_2 is found by [121]

$$B_2 = \frac{1}{2}R_c^{\,2}(\theta - \sin\theta). \tag{88}$$

The uniform distribution of the node position allows the computation of the $\Pr[B_2]$ as:

$$\Pr[B_2] = B_2\frac{1}{\pi r^2}. \tag{89}$$

The probability that a uniformly distributed node position within a circle of radius r is within the boundary of a circle of radius R_c is then the sum $\Pr[B_1] + \Pr[B_2]$.

LIST OF REFERENCES

[1] P. W. Singer, *Wired for War.* New York: Penguin Press, 2009.

[2] The unmanned systems integrated roadmap FY2011–2036. Under Secretary of Defense for Acquisition, Technology and Logistics; 2011[http://www.acq.osd.mil/sts/initiatives/]. Available: http://www.acq.osd.mil/sts/initiatives/.

[3] J. Gertler, U.S. unmanned aerial systems. Congressional Research Service. Washington, DC, 2012. Available: www.crs.gov.

[4] Headquarters, Dept. of the Army, "Technical manual, small unmanned aircraft system (SUAS) RQ-11B," Tech. Rep. TM 1–1550–695–13&P, 28 March 2008.

[5] U.S. Department of Defense, "Unmanned systems roadmap, 2007–2032," U.S. Dept. of Defense, Washington, DC, Dec 10, 2007.

[6] D. Walsh, "Ground sensors play key role in battlefield sensing," *Defense Systems*, April 25, 2012.

[7] National Instruments, (6 May 2011), *Wireless Sensor Network Analog Input Nodes, NI WSN-3202*. Available: http://sine.ni.com/ds/app/doc/p/id/ds-293/lang/en.

[8] M. W. Isherwood, "Unmanned Systems and the Joint Team," *Joint Force Quarterly,* vol. 58, pp. 57–61, 2010.

[9] Army networks: Size and scope of modernization investment merit increased oversight, Government Accountability Office, 2013 [http://www.gao.gov/]. Available: https://www.hsdl.org/?view&did=728559.

[10] B. Touchton, T. Galluzzo, D. Kent, and C. Crane, "Perception and Planning Architecture for Autonomous Ground Vehicles," *Computer,* vol. 39, pp. 40–47, 2006.

[11] AS-4c Committee, (March 23, 2010), *AS 6009: JAUS Mobility Service Set*. Available: http://standards.sae.org/wip/as6009a/.

[12] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wirel. Pers. Commun.,* vol. 6, pp. 585–595, Mar. 1998.

[13] R. U. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, and A. J. Paulraj, "Performance of multiantenna signaling techniques in the presence of polarization diversity," *IEEE Trans. Signal Process.,* vol. 50, pp. 2553–2562, Oct. 2002.

[14] M. Coldrey, “Modeling and capacity of polarized MIMO channels,” in *IEEE Vehicular Technology Conference*, pp. 440–444, 2008.

[15] C. Oestges, B. Clerckx, M. Guillaud, and M. Debbah, “Dual-polarized wireless communications: from propagation models to system performance evaluation,” *IEEE Trans. Wireless Commun.,* vol. 7, pp. 4019–4031, Oct. 2008.

[16] T. Camp, J. Boleng and V. Davies, “A survey of mobility models for ad hoc network research,” *Wirel. Commun. Mob. Comput.,* vol. 2, pp. 483–502, 2002.

[17] P. Samar and S. B. Wicker, “On the behavior of communication links of a node in a multi-hop mobile environment,” in *5th ACM Int. Symposium on Mobile Ad Hoc Networking and Computing,* pp. 145–156, 2004.

[18] A. B. McDonald and T. F. Znati, “A mobility-based framework for adaptive clustering in wireless ad hoc networks,” *IEEE J. Sel. Areas Commun. 17(8),* pp. 1466–1487, 1999.

[19] R. Sharma, R. Bhagavathula, K. Namuduri, and R. Pendse, “Link availability and its effects on the capacity of mobile ad hoc wireless networks - an analytical approach,” in *31st IEEE Conference on Local Computer Networks,* pp. 295–299, 2006.

[20] X. Wu, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, “Link lifetime as a function of node mobility in MANETs with restricted mobility: Modeling and applications,” in *5th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks,* pp. 1–10, 2007.

[21] M. Ayyash, D. Ucci, and K. Alzoubi, “An entity stability measure for mobile ad hoc networks,” in *MILCOM,* pp. 1–6, 2006.

[22] P. Prabhakaran and R. Sankar, “Impact of realistic mobility models on wireless networks performance,” in *IEEE Intl. Conf. on Wireless and Mobile Computing, Networking and Communications,* pp. 329–334, 2006.

[23] F. Bai, Narayanan Sadagopan, and A. Helmy, “IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks,” in *INFOCOM,* pp. 825–835, 2003.

[24] Q. Zheng, X. Hong, and S. Ray, “Recent advances in mobility modeling for mobile ad hoc network research,” in *42nd Annual Southeast Regional Conference,* pp. 70–75, 2004.

[25] W. J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, “Modeling Spatial and Temporal Dependencies of User Mobility in Wireless Mobile Networks,” *IEEE/ACM Trans. on Networking,* vol. 17, pp. 1564–1577, 2009.

[26] S. Koenig, “Exploring unknown environments with real-time search or reinforcement learning,” *Adv. in Neural Info. Process. Sys.*, pp. 1003–1009, 1999.

[27] S. Koenig, C. Tovey, and W. Halliburton, “Greedy mapping of terrain,” in *Proc. ICRA*, pp. 3594–3599, 2001.

[28] B. Yamauchi, “A frontier-based approach for autonomous exploration,” in *Proc. CIRA*, pp. 146–151, 1997.

[29] S. Misra, S. D. Hong, G. Xue, and J. Tang, “Constrained Relay Node Placement in Wireless Sensor Networks: Formulation and Approximations,” *IEEE/ACM Trans. Netw.,* vol. 18, pp. 434–447, 2010.

[30] A. Kashyap, S. Khuller, and M. Shayman, “Relay placement for higher order connectivity in wireless sensor networks,” in *Proc. 25th IEEE Int. Conf. on Comput. Commun.,* pp. 1–12, 2006.

[31] N. Pezeshkian, H. G. Nguyen, and A. Burmeister, “Unmanned ground vehicle radio relay deployment system for non-line-of-sight operations,” in *Proc. 13th IASTED Int. Conf. on Robotics & Applications*, pp. 1–7, 2007.

[32] M. R. Souryal, A. Wapf, and N. Moayeri, “Rapidly-deployable mesh network testbed,” in *Proc. IEEE GLOBECOM*, pp. 1–6, 2009.

[33] B. Sklar, *Digital Communications.* Upper Saddle River, NJ: Prentice Hall, 2001.

[34] H. T. Friis. A note on a simple transmission formula. *Proceedings of the IRE* vol. 34 no. 5, pp. 254–256, May 1946.

[35] T. S. Rappaport, *Wireless Communications : Principles and Practice.* Upper Saddle River, NJ: Prentice Hall PTR, 2002.

[36] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry, “A statistical model of urban multipath propagation,” *IEEE Trans. Veh. Technol.,* vol. 21, pp. 1–9, 1972.

[37] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, “A statistical basis for lognormal shadowing effects in multipath fading channels,” *IEEE Trans. Commun.,* vol. 46, pp. 494–502, 1998.

[38] D. C. Cox, “Universal digital portable radio communications,” *Proceedings of the IEEE,* vol. 75, pp. 436–477, 1987.

[39] S. Haykin and M. Moher, *Modern Wireless Communications.* Upper Saddle River, NJ: Pearson Prentice Hall, 2005.

[40] W. Stutzman and G. Thiele, *Antenna Theory and Design.* USA: John Wiley & Sons, 1998.

[41] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.,* vol. 27, pp. 1029–1046, 2009.

[42] K. K. Aggarwal and S. Rai, "Reliability Evaluation in Computer-Communication Networks," *IEEE Trans. Reliab.,* vol. R-30, pp. 32–35, 1981.

[43] R. Ahlswede, C. Ning, S. -. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory,* vol. 46, pp. 1204–1216, 2000.

[44] F. Dai and J. Wu, "An extended localized algorithm for connected dominating set formation in ad hoc wireless networks," *IEEE Trans. Parallel Distrib. Syst.,* vol. 15, pp. 908–920, 2004.

[45] J. L. Gross and J. Yellen, *Graph Theory and its Applications.* Boca Raton: CRC Press, 1999.

[46] C. Therrien and M. Tummala, *Probability and Random Processes for Electrical and Computer Engineers.* Boca Raton, FL: CRC Press, 2012.

[47] C. J. Colbourn, *The Combinatorics of Network Reliability.* New York: Oxford University Press, 1987.

[48] R. Severinghaus, M. Tummala, and J. McEachen, "Network formations to bridge or extend wireless sensor networks," in *5th Intl. Conf. on Signal Process. and Commun. Syst. (ICSPCS)*, pp. 1–6, 2011.

[49] V. Srivastava and M. Motani, "Cross-layer design: a survey and the road ahead," *IEEE Commun. Mag.,* vol. 43, pp. 112–119, 2005.

[50] R. Severinghaus, M. Tummala, and J. McEachen, "Networks for maintaining system-level availability for an exploring robot," *IEEE Syst. J.,* submitted for publication.

[51] Supplement to IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5 GHz band. *IEEE Std 802. 11a-1999* pp. i. 1999.

[52] Y. Chen, Q. Zeng, and D. P. Agrawal, "Performance analysis and enhancement for IEEE 802.11 MAC protocol," in *Proceedings of IEEE Int'l Conf on Telecommunications*, pp. 860–867, 2003.

[53] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” *IEEE J. Sel. Areas Commun.,* vol. 18, pp. 535–547, 2000.

[54] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *Information Theory, IEEE Transactions on,* vol. 46, pp. 388–404, 2000.

[55] C. E. Perkins and E. M. Royer, “Ad-hoc on-demand distance vector routing,” in *Proceedings of 2nd IEEE Workshop on Mobile Computing Systems and Applications,* pp. 90–100, 1999.

[56] E. Gelal, G. Jakllari, S. V. Krishnamurthy, and N. E. Young, “Topology Management in Directional Antenna-Equipped Ad Hoc Networks,” *Mobile Computing, IEEE Transactions on,* vol. 8, pp. 590–605, 2009.

[57] C. E. Shannon, “A mathematical theory of communication,” *BSTJ,* vol. 27, pp. 379–423, 623–657, 1948.

[58] C. Park, K. Lahiri, and A. Raghunathan, “Battery discharge characteristics of wireless sensor nodes: an experimental analysis,” *IEEE Conf. Sensor and Ad Hoc Commun. and Networks (SECON),* pp. 430–440, 2005.

[59] J. Keller, “The time has come for military ground robots,” *Mil. and Aerospace Electron.,* vol. 20/6, June 1, 2009.

[60] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela, “Lifting the limits on high speed wireless data access using antenna arrays,” *IEEE Commun. Mag.,* vol. 39, pp. 156–162, 2001.

[61] T. ElBatt and A. Ephremides, “Joint scheduling and power control for wireless ad hoc networks,” *IEEE Trans. Wireless Commun.,* vol. 3, pp. 74–85, 2004.

[62] N. Pradhan and T. Saadawi, “Energy efficient distributed power management algorithm with directional antenna for wireless sensor networks,” in *34th IEEE Sarnoff Symposium*, pp. 1–6, 2011.

[63] R. F. Mills and G. E. Prescott, “Detectability models for multiple access low-probability-of-intercept networks,” *IEEE Trans. Aerosp. Electron. Syst.,* vol. 36, pp. 848–858, 2000.

[64] Z. Zhang, “DTRA: directional transmission and reception algorithms in WLANs with directional antennas for QoS support,” *Network, IEEE,* vol. 19, pp. 27–32, 2005.

[65] R. Choudhury, X. Yang, R. Ramanathan, and N. Vaidya, “Using directional antennas for medium access control in ad hoc networks,” in *MOBICOM ‘02,* Atlanta, Georgia, Sep 23–28, 2002.

[66] R. R. Choudhury and N. H. Vaidya, "Deafness: A MAC problem in ad hoc networks when using directional antennas," in *Proceedings of the 12th IEEE Int. Conf. on Network Protocols,* pp. 283–292, 2004.

[67] H. Dai, K. Ng, and M. Wu, "An overview of MAC protocols with directional antennas in wireless ad hoc networks," in *Int. Conf. on Wireless and Mobile Commun.,* pp. 84–84, 2006.

[68] O. Bazan and M. Jaseemuddin, "A Survey On MAC Protocols for Wireless Adhoc Networks with Beamforming Antennas," *IEEE Commun. Surveys & Tutorials,* vol. PP, pp. 1–24, 2011.

[69] J. Mitola, "The software radio architecture," *IEEE Commun. Mag.,* vol. 33, pp. 26–38, 1995.

[70] F. K. Jondral, "Software-defined radio: basics and evolution to cognitive radio," *EURASIP Journal on Wireless Commun. and Networking,* pp. 275–283, 2005.

[71] W. H. W. Tuttlebee, "Software-defined radio: facets of a developing technology," *IEEE Pers. Commun.,* vol. 6, pp. 38–44, 1999.

[72] J. Mitola III, "SDR architecture refinement for JTRS," in *IEEE MILCOM 2000,* pp. 214–218, 2000.

[73] Vanu Inc. *Vanu products- Anywave base station*. Available: http://www.vanu.com/products/anywave/.

[74] H. Harada, "A small-size software defined cognitive radio prototype," in *IEEE 19th Int'l Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC), 2008.* Cannes, France, pp. 1–5, 2008.

[75] J. L. Shanton, "A software defined radio transformation," in *IEEE MILCOM 2009,* Boston, MA, pp. 1–5, 2009.

[76] M. Woh, Sangwon Seo, S. Mahlke, T. Mudge, C. Chakrabarti, and K. Flautner, "AnySP: Anytime Anywhere Anyway Signal Processing," *IEEE Micro,* vol. 30, pp. 81–91, 2010.

[77] S. Haykin. Cognitive radio: Brain-empowered wireless communications. *IEEE J. Sel. Areas Commun.* vol. 23 no. 2, pp. 201–220. 2005.

[78] M. McHenry, D. McCloskey, and G. Lane-Roberts. Spectrum occupancy measurements location 4 of 6: Republican national convention, new york city, new york, august 30, 2004-september 3, 2004. Shared Spectrum Company. 2005Available: http://www.sharedspectrum.com/wp-content/uploads/4_NSF_NYC_Report.pdf.

[79] IEEE, “Standard for information technology- part 22: Cognitive wireless RAN medium access control (MAC) and physical layer (PHY) specifications: Policies and procedures for operation in the TV bands,” IEEE, New York, USA, 2011.

[80] C. Cordeiro, K. Challapali, D. Birru, and N. Sai Shankar, “IEEE 802.22: The first worldwide wireless standard based on cognitive radios,” in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN) 2005.* Baltimore, MD, pp. 328–337, 2005.

[81] D. Qiao, S. Choi, and K. G. Shin, “Goodput analysis and link adaptation for IEEE 802.11a wireless LANs,” *IEEE Trans. Mobile Comput.,* vol. 1, pp. 278–292, 2002.

[82] D. Cabric, S. M. Mishra, and R. W. Brodersen, “Implementation issues in spectrum sensing for cognitive radios,” in *Conf. Record of the 38th Asilomar Conf. on Signals, Systems and Computers,* pp. 772–776, 2004.

[83] R. C. Dorf, Ed., *The Electrical Engineering Handbook.* Boca Raton: CRC Press, 1993.

[84] I. Gupta and A. Ksienski, “Effect of mutual coupling on the performance of adaptive arrays,” *IEEE Trans. Antennas Propag.,* vol. 31, pp. 785–791, 1983.

[85] J. H. Winters, “Smart antenna techniques and their application to wireless ad hoc networks,” *IEEE Wireless Communications,* vol. 13, pp. 77–83, 2006.

[86] A. Paulraj, R. Roy, and T. Kailath, “Estimation of signal parameters via rotational invariance techniques- esprit,” in *19th Asilomar Conference on Circuits, Systems and Computers,* pp. 83–89, 1985.

[87] T. K. Paul and T. Ogunfunmi, “Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment,” *IEEE Circuits Syst. Mag.,* vol. 8, pp. 28–54, 2008.

[88] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications.* UK: Cambridge University Press, 2007.

[89] E. Telatar, “Capacity of Multi-antenna Gaussian Channels,” *Eur. Trans. Telecommun.,* vol. 10, pp. 585–595, 1999.

[90] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, “Great expectations: the value of spatial diversity in wireless networks,” *Proc. IEEE,* vol. 92, pp. 219–270, 2004.

[91] A. J. Paulraj, R. U. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications.* Cambridge, UK: University Press, 2003.

[92] D. Gesbert, M. Shafi, Da-shan Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.,* vol. 21, pp. 281–302, 2003.

[93] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.,* vol. 16, pp. 1451–1458, Oct. 1998.

[94] R. G. Vaughan, "Polarization diversity in mobile communications," *IEEE Trans. Veh. Technol.,* vol. 39, pp. 177–186, Aug. 1990.

[95] H. Asplund, J. Berg, F. Harrysson, J. Medbo, and M. Riback, "Propagation characteristics of polarized radio waves in cellular communications," in *Proc. of IEEE 66th Vehicular Technol. Conf.*, pp. 839–843, 2007.

[96] M. Toeltsch, J. Laurila, K. Kalliola, A. F. Molisch, P. Vainikainen, and E. Bonek, "Statistical characterization of urban spatial radio channels," *IEEE J. Sel. Areas Commun.,* vol. 20, pp. 539–549, 2002.

[97] C. Oestges and B. Clerckx, *MIMO Wireless Communications.* Oxford, UK: Academic Press, 2007.

[98] H. Mott, *Antennas for Radar and Communications: A Polarimetric Approach.* John Wiley and Sons, Inc., 1992.

[99] C. N. Chuah, J. M. Kahn, and D. Tse, "Capacity of multi-antenna array systems in indoor wireless environment," in *IEEE Global Telecomm. Conf.*, pp. 1894–1899 vol.4, 1998.

[100] C. N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory,* vol. 48, pp. 637–650, 2002.

[101] A. M. Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory,* vol. 51, pp. 2491–2509, 2005.

[102] F. Bai and A. Helmy, "A survey of mobility models in wireless adhoc networks," in *Wireless Ad Hoc and Sensor Networks,* Kluwer Academic Publishers, pp. 1–29, 2004.

[103] X. Hong, M. Gerla, G. Pei, and C. Chiang, "A group mobility model for ad hoc wireless networks," in *Workshop on Modeling and Simulation of Wireless and Mobile Systems (MSWiM),* pp. 53–60, 1999.

[104] S. Y. Seidel, T. S. Rappaport, S. Jain, M. L. Lord, and R. Singh. Path loss, scattering and multipath delay statistics in four european cities for digital cellular and microcellular radiotelephone. *IEEE Trans. Veh. Technol. 40(4),* pp. 721–730, 1991.

[105] W. Zhao, Y. Chen, M. Ammar, M. Corner, B. Levine, and E. Zegura, "Capacity enhancement using throwboxes in DTNs," in *Proc. IEEE MASS*, pp. 31–40, 2006.

[106] M. A. M. Vieira, R. Govindan, and G. S. Sukhatme, "Towards autonomous wireless backbone deployment in highly-obstructed environments," in *Proc. IEEE ICRA*, pp. 5369–5374, 2011.

[107] H. C. H. Chiu, B. Ryu, H. Zhu, P. Szekely, R. Maheswaran, C. Rogers, A. Galstyan, B. Salemi, M. Rubenstein, and W. Shen, "TENTACLES: Self-configuring robotic radio networks in unknown environments," in *Proc. IEEE/RSJ IROS,* pp. 1383–1388, 2009.

[108] Y. Pei, M. W. Mutka, and N. Xi, "Coordinated multi-robot real-time exploration with connectivity and bandwidth awareness," in *Proc. IEEE ICRA*, pp. 5460–5465, 2010.

[109] H. G. Nguyen, N. Pezeshkian, M. Raymond, A. Gupta, and J. M. Spector, "Autonomous communication relays for tactical robots," in *Proc. 11th Int. Conf. Adv. Robotics*, pp. [CD-ROM], 2003.

[110] M. M. Zavlanos and G. J. Pappas, "Distributed connectivity control of mobile networks," *IEEE Trans. Robot. 24(6),* pp. 1416–1428, 2008.

[111] J. Le Ny, A. Ribeiro, and G. J. Pappas, "Adaptive communication-constrained deployment of unmanned vehicle systems," *IEEE J. Sel. Areas Commun. 30(5),* pp. 923–934, 2012.

[112] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory 50(12),* pp. 3062–3080, 2004.

[113] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Trans. Commun. 52(10),* pp. 1820–1830, 2004.

[114] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Antennas Propag. 40(2),* pp. 207–217, 1992.

[115] EURO-COST 231, Digital mobile radio towards future generation systems, COST 231 final report, 1998[Online]. Available: http://www.lx.it.pt/cost231/.

[116] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Trans. Robot. 21(3),* pp. 376–386, 2005.

[117] S. Russell and P. Norvig, *Artificial Intelligence A Modern Approach.* Upper Saddle River, NJ: Prentice Hall, 2010.

[118] R. Severinghaus, M. Tummala, and J. McEachen, "Availability of ad hoc wireless networks of unmanned ground vehicles with group mobility," in *Proc. 46th Hawaii Int. Conf. on System Sciences*, pp. 5097–5105, 2013.

[119] Y. Mostofi, A. Gonzalez-Ruiz, A. Gaffarkhah, and D. Li, "Characterization and modeling of wireless channels for networked robotic and control systems–a comprehensive overview," in *Proc. IEEE/RSJ IROS,* pp. 4849–4854, 2009.

[120] H. Pham, "Commentary: steady-state series-system availability," *IEEE Trans. Reliab.,* vol. 52, pp. 146–147, 2003.

[121] D. Zwillinger, Ed., *CRC Standard Mathematical Tables and Formulae.* Boca Raton: CRC Press, 1996.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California