Implementation and Analysis of a Hybrid-ARQ Based Cooperative Diversity Protocol

Sheetu Dasari

Problem Report submitted to the College of Engineering and Mineral Resources

at West Virginia University in partial fulfillment of the requirements

for the degree of

Master of Science

in

Electrical Engineering

Matthew C. Valenti, Ph.D., Chair

Daryl S. Reynolds, Ph.D.

Natalia A. Schmid, Ph.D.

Lane Department of Computer Science and Electrical Engineering

2005
Morgantown, West Virginia

Keywords: Relaying, Hybrid-ARQ, HARBINGER, Cooperative communications, Throughput.

Implementation and Analysis of a Hybrid-ARQ based Cooperative Diversity Protocol.

Sheetu Dasari

ABSTRACT

In this project, a hybrid Automatic Repeat reQuest (hybrid-ARQ) based cooperative diversity protocol is implemented and analyzed using Reed-Solomon (RS) codes. This protocol is suitable for wireless sensor networks that are decomposed into clusters of several radio nodes including a source, a destination, and one or more relays. A general approach, that involves the coordination of both the direct and relayed transmissions and which corresponds to the classical *relay channel* model is followed. In these networks, the spatial diversity advantage is obtained with relaying. The protocol is an *incremental redundancy* scheme, which increases the *throughput efficiency* by adapting the error correcting code redundancy to the varying channel conditions. The *normalized throughput* of the protocol is simulated for both AWGN and *block fading* environments with BPSK modulation and RS codes. The performance of the protocol is studied and analyzed from these simulation results. The results indicate that, contrary to previous information-theoretic studies, hybrid-ARQ based cooperative diversity is not always better than conventional multihop routing.

ACKNOWLEDGEMENTS

I would like to acknowledge many people without whom this project would not have been possible. First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Matthew C. Valenti, and thank him for his invaluable patience and esteemed guidance throughout this project. The project would not have been possible without his constant support and motivation. Also, I would like to thank Dr. Daryl Reynolds and Dr. Natalia A. Schmid for their support as committee members.

On a personal note, I am greatly indebted to my parents for their constant support and encouragement to pursue higher studies.

Table of Contents

Chapter 1: Introduction

The recent past has seen rapid developments in the field of wireless technology. The proliferation of wireless devices, which in part is due to the overall increase in connectivity, digitization and sharing of information globally, is a significant trend that most experts predict will continue. A world is envisioned where wireless embedded networks hold a great promise for deployment in residential and commercial building-automation, industrial plant monitoring, and other wireless sensing and control applications. These embedded networks of distributed sensors and actuators must operate at extremely low power, should be inexpensive and small. To meet the challenges of next generation wireless system design, fundamentally new methods are needed to exploit all available dimensions of communication channels and network. A proposed solution to meet some of these challenges is *cooperative communication* [1].

This project deals with the analysis and implementation of an energy efficient wireless ad hoc network protocol using the concept of *relaying*. The present chapter gives a brief overview of the wireless channel and then discusses some of the strategies used to overcome problems that arise in wireless environment. The chapter concludes with the problem statement and an overview of the reminder of the problem report.

1.1 Wireless Channel

Wireless transmission is nothing but propagation of electromagnetic waves through space. These electromagnetic waves represent the information being communicated. The main effects on the information transmission in a mobile radio channel arise due to the reflection, diffraction, and scattering mechanisms behind the electromagnetic waves propagation. The transmission path in a wireless channel varies from a simple Line-Of-Sight (LOS) to one that is severely obstructed by buildings, mountains & foliage. The wireless medium is highly random and prone to transmission errors.

Due to multiple reflections, the electromagnetic waves travel along different paths. The interaction among these waves causes multipath fading. Also, as the signal travels in space the strength of it decreases with distance. Various propagation models are suggested to predict the average received signal strength at a given distance from the transmitter, as well as the variability of the signal strength in close spatial proximity to a particular location [2]. These models are characterized into large scale propagation models and small scale propagation models. Large scale propagation models or the path loss models predict the mean signal strength over large transmitter-receiver separation distances (many wavelengths). Small scale or fading models characterize the rapid fluctuations of the received signal strength over very short travel distances. The signal fades rapidly as the receiver moves, but the local average signal changes much more gradually with distance.

1.1.1 Path Loss Models

Let a transmitted signal be represented as *s(t)* and the received signal as *r(t)*. Suppose *s(t)* of power P_t is transmitted through a wireless channel, with corresponding received signal *r(t)* of power P_r, where P_r is averaged over any random variations due to fading and shadowing. Then, the *path loss*, which represents signal attenuation as a positive quantity measured in dB, can be defined as the dB ratio between the effective transmitted power and the received power as [2]:

$$PL(dB) = 10\log\frac{P_t}{P_r}\ (dB). \qquad (1\cdot1)$$

The simplest model for signal propagation in a wireless medium is free space path loss. The power received by a receiver antenna in free space without any obstructions is given by the *Friis free space* equation as [2]:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L}, \qquad (1.2)$$

where P_t is the transmitted power, P_r (d) is the received power which is a function of transmitter-receiver separation, G_t is transmitter antenna gain, G_r is receiver antenna gain, d is transmitter-receiver separation distance in meters, L is system loss factor not related to propagation, λ is wavelength in meters. When antenna gains are excluded, the antennas are assumed to have unity gain. Thus, the received signal power is inversely proportional to the square of the distance between transmit and receive antennas. The *Friis* free space model is a valid predictor for P_r, for values of d which are in the far field of the transmitting antenna.

The *far-field* or *Fraunhofer* region, of a transmitting antenna is defined as the region beyond the far field distance d_f, which is related to the largest linear dimension of the transmitter antenna aperture *(D)* and the carrier wavelength (λ) as [2]:

$$d_f = \frac{2D^2}{\lambda}. \qquad (1.3)$$

The path loss models use a close in distance d_0 as a known received power reference point, such that $d_0 > d_f$. The received power at any distance $d>d_0$ is related to P_r at d_0 as:

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^2, \quad d \ge d_0 \ge d_f \ . \qquad (1.4)$$

Usually, $d_0 = 1$.

Most mobile communication systems operate in complex propagation environments which cannot be accurately modeled by free-space path loss or ray tracing. Various path loss models have been suggested to predict path loss in typical wireless environments. These models are mainly based on empirical measurements over a given distance in a given frequency range and a particular geographical area or building. The empirical path loss *PL (d)* for a given environment (e.g. a city, suburban area, or office building) is defined as the average of the local mean attenuation measurements at distance d, averaged over all available measurements in the given environment. Some examples of the empirical models developed are Okumura model, Hata model, and the piecewise linear model.

For system analysis, a simple generic model that captures the essence of signal propagation without resorting to complicated path loss models is often used. The *log distance path loss model* is commonly used for system design. The average large scale path loss for an arbitrary transmitter-receiver separation for this model is expressed as a function of distance by using a path loss exponent *(n)* as:

$$\overline{PL}(d)[dB] = PL(d_0) + 10n\log\left(\frac{d}{d_0}\right)\ dB. \tag{1.5}$$

The bars in equations denote the ensemble average of all possible path loss values for a given value of d. The value of n depends on the specific propagation environment. For example, in free space n is 2, and in the presence of multipath reflections will have a larger value. The average received signal level is estimated by using path loss model as a function of distance given by equation (1.1). By knowing the average received signal it becomes possible to predict the Signal-Noise-Ratio (*SNR*) for a wireless communication system, which is a frequently used parameter for system design.

Shadowing effects are described by the *log normal shadowing model*. These effects occur over a large number of measurement locations which have the same transmitter-receiver separation, but different levels of clutter on propagation path. This implies that measured signals at a specific transmitter-receiver separation have a Gaussian (normal) distribution about the distance dependent mean of equation (1.5) and is given as:

$$PL(d)[dB] = PL(d_0) + 10n\log\left(\frac{d}{d_0}\right) + X_{\sigma}. \tag{1.6}$$

1.1.2 Fading

Fading is caused by interference between two or more versions of the transmitted signal which arrive at the receiver at slightly different times. These multipath waves combine at the receiver antenna to give a resultant signal which can vary widely in amplitude and phase depending on the distribution of the intensity, relative propagation time of the waves and the bandwidth of the transmitted signal. Though multipath effects are captured in the ray-tracing models for deterministic channels, in practice deterministic channel models are rarely available, and thus multipath channels are to be statistically characterized. The small scale variations of a mobile radio channel can be modeled by a random time-varying impulse response [2]. The received signal *r (d, t)* at position *d* and time *t* can be expressed as convolution of *s (t)* with *h (d, t)* given as:

$$r(d,t) = s(t) \otimes h(d,t)\,, \tag{1.7}$$

where *h* (*d, t*) is channel impulse response and *s (t)* being the transmitted signal. Factors influencing small scale fading are multipath propagation, speed of mobile, speed of surrounding objects, and transmission bandwidth of the signal.

The type of fading experienced by a signal propagating through a mobile radio channel depends on the nature of the transmitted signal with respect to the characteristics of the channel. Depending on the relation between the signal parameters (such as *bandwidth*, *symbol period*, *etc*.) and the channel parameters (such as *rms delay spread* and *Doppler spread*), different transmitted signals will undergo different types of fading. The time dispersion and frequency dispersion mechanisms in a mobile radio channel lead to four possible effects [3]. Time dispersion due to multipath causes the transmitted signal to undergo either *flat* or *frequency selective* fading whereas frequency dispersion leads to either *slow* or *fast* fading. If the mobile radio channel has a constant gain and linear phase response over a bandwidth which is greater than the bandwidth of the transmitted signal, then the received signal will undergo *flat fading*. The distribution of the instantaneous gain of flat fading channels is commonly Rayleigh. When the channel has a constant gain and linear phase over a bandwidth which is smaller than the bandwidth of the transmitted signal, then it is *frequency selective fading*. If the channel impulse response changes rapidly with in the symbol duration then it is called *fast fading* and when the channel impulse response changes at a rate much slower than the transmitted base band signal then it is *slow* fading.

Let a signal transmitted through wireless channel be received as signal Z, a complex Gaussian i.e. $Z=X+jY$. Let the real and imaginary parts of Z be independent, have zero mean and equal variances σ^2. The random variable R is defined to be the magnitude of Z. $R=|Z|=\sqrt{X^2+Y^2}$ is Rayleigh and has the following probability density function [4]:

$$f_R(r)=\frac{r}{\sigma^2}\exp\left\{-\frac{r^2}{2\sigma^2}\right\}u(r), \tag{1.8}$$

where, $u(.)$ is the unit step function.

Rayleigh variables model the received envelope of signals that are transmitted through a fading channel when there is no direct line of sight (LOS) between the transmitter and receiver. When the magnitude of Z is Rayleigh, then the phase of Z will be uniformly distributed over the range (0, 2π).

Let, $P=|Z|^2$ be the instantaneous power of Z and $\sigma^2=\frac{1}{2\lambda}$. Since, Z is Rayleigh then P is *exponentially* distributed with probability density function given by [4]:

$$f_P(x)=\lambda e^{-\lambda x}u(x), \tag{1.9}$$

where, λ is a parameter of the random variable. The mean and variance of P are given by: mean $(m)=\frac{1}{\lambda}$ and variance $\left(\sigma^2\right)=\frac{1}{\lambda^2}$.

Hence, exponential random variables can be used to model the power of signals received through fading channels with no direct LOS. If the fading channel has a fixed LOS component which often occurs when the transmitter and receiver are close, then a Rician random variable is used to model the received amplitude of signals. Assuming the imaginary component Y has a zero mean but the real component X has a positive mean $m_X=\beta$.

The signal envelope in this case is given by Rician distribution as:

$$f_R(r) = \frac{r}{\sigma^2} \exp\left\{ \frac{-(r^2+\beta^2)}{2\sigma^2} \right\} I_o\left(\frac{\beta r}{\sigma^2} \right) u(x), \tag{1.10}$$

where, $I_o(.)$ is the zero order modified Bessel function of the first kind and is defined as:

$$I_o(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x\cos\theta} d\theta. \tag{1.11}$$

Rician random variables are characterized by *Rician-K- Factor* given by:

$$K = \frac{\beta}{2\sigma^2}. \tag{1.12}$$

Some experimental data do not fit well into either of these distributions. Thus, a more general fading distribution, called *Nagakami-m* fading distribution, whose parameters can be adjusted to fit a variety of empirical measurements, is suggested. The Nagakami-m distribution closely approximates Rician probability density function and is given by:

$$f_R(r) = \frac{2r^{2m-1}}{\Gamma(m)} \left(\frac{m}{\Omega} \right)^m \exp\left(-\frac{mr^2}{\Omega} \right) u(r), \tag{1.13}$$

where, $\Gamma(.)$ is the Gamma function and $\Omega = E[R^2]$ is the average power, and $m \geq 1/2$ is a constant called the *fading figure*. The Rayleigh condition is met for *m*=1, the *Rician-K-Factor* and *Nagakami-m* fading figure are related by:

$$m = \frac{(K+1)^2}{2K+1}. \tag{1.14}$$

Rayleigh block fading channel is the model commonly considered [5]. The channel remains constant for the period of time required to transmit a contiguous block of symbols transmission. So the received signal power throughout this period remains same and varies from block to block as a Rayleigh random variable.

In summary, there are three main effects caused by wireless channel i.e., distance loss effect, shadowing and multipath fading. The shadowing and distance loss are relatively constant and can be compensated by power control and automatic gain control. Hence, the multipath fading becomes the dominant effect of concern for a wireless communication system design and analysis.

1.2 Coded Protocols

The wireless channel, being highly unreliable due to noise, interference, fading, and other factors, is error-prone and suffers location-dependent, time-varying, and burst errors. The two basic error control strategies are Automatic Repeat reQuest (ARQ) and Forward Error Correction (FEC). ARQ is an error detection and retransmission strategy whereas FEC is an error correction strategy. ARQ is efficient when the channel condition is good and as the channel deteriorates FEC is more suitable. However, ARQ/FEC neither is practically efficient for block fading channels. A combination of FEC and ARQ called *Hybrid ARQ/FEC* (HARQ) has been suggested for unpredictable environments such as wireless for achieving better throughput efficiency. In HARQ, the receiver attempts to correct errors first and, if the errors are uncorrectable, retransmission of the packet is requested [6].

A considerable amount of research has been devoted to developing energy efficient multihop protocols. Conventional multihop protocols are constructed from the sequential use of point-to-point links, where the links are viewed at the network protocol layer. Here, the broadcast-oriented nature of radios is essentially not being used. For scenarios as in sensor networks, and military battlefield communication networks, the radios are substantially power constrained and new techniques are being explored.

The significant probability of a channel being in a deep fade, and the need to develop energy efficient protocols has motivated researchers to look at various diversity techniques to improve on the error performance. Diversity combats multipath fading by providing the receiver with redundant signal information. This information is provided through uncorrelated diverse channels and thus allowing the receiver to average individual channel effects. The most common forms of diversity are space, time and frequency.

Diversity over time is obtained by coding and interleaving. The coded symbols are dispersed over time in different coherence periods. Hence, these codewords experience independent fades. Diversity over frequency can be obtained if the channel is frequency-selective. Spatial diversity is achieved by the transmission and/or reception of multiple copies of a signal from physically different points in space (i.e. multiple transmit and/or receive antennas). The separation between adjacent antennas should be large enough to achieve spatial diversity. This is required for the signals from different antennas to undergo independent fading.

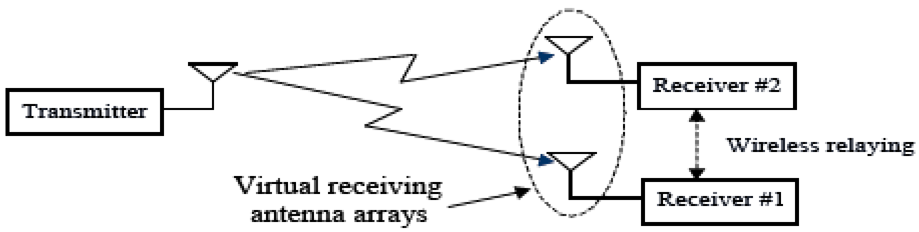

Figure 1.1: Cooperative diversity through virtual antenna array from [7].

With cooperation, the radios in the network can essentially share their antennas and other resources to create a “virtual array” through distributed transmission and signal processing to realize spatial diversity. This type of scheme is called *cooperative communication*. Cooperation has shown promise in increasing throughput and providing better power efficiency in wireless networks [8]. So by cooperation, nodes can transmit more information across the network or improve performance dramatically. In essence, the broadcast nature of the radio and the multipath fading that occurs in a wireless channel are used here.

Some of the strategies employed by the cooperating radios are to detect and repeat received bits (decode-and-forward) or amplify and forward received signal [8]. *Coded cooperation* [9] which is a combination of cooperation and channel coding is also one of the strategies used lately. In general, various channel coding methods can be used within the coded cooperation framework. For example, the overall code may be a block or convolutional code, or a combination of both. The partitioning of the code bits into various frames may be achieved through puncturing, product codes, or other forms of concatenation. In addition, some form of space-time coding may be used for transmission of the next frame [10]. Some of the suggested codes are rate compatible punctured convolutional codes, turbo codes, low density parity check codes, Reed-Solomon (RS) codes. RS codes are a good option for block fading environments as they provide excellent error-correction capability in terms of bursty error suppression.

1.3 Problem Statement

The goal of the problem report is to implement and analyze a hybrid-ARQ based cooperative diversity protocol using Reed-Solomon codes. This protocol is suitable for wireless sensor networks that are decomposed into clusters of several radio nodes including a source, a destination, and one or more relays. Each cluster works cooperatively to transmit information from source to destination. A general approach, that involves the coordination of both the direct and relayed transmissions and which corresponds to the classical *relay channel* model is followed. In these networks, the spatial diversity advantage is obtained with relaying. The relaying protocol studied is cross-layer, combining the functionalities of both medium access control (MAC) and routing. The cooperative diversity protocol considered for study is the HARBINGER protocol proposed by Zhao and Valenti [11]. The work by Zhao mainly dealt with the information theoretic perspective of the cross layer protocol and did not consider the performance of any specific codes or modulation. Here in this project we study, implement and analyze the performance of the protocol using BPSK modulation and Reed Solomon codes.

In the protocol, signaling is over a random phase block interference channel, and transmission from the various nodes are non coherent. The key concept is that the wireless networks considered are viewed as a generalization of a block fading channel with hybrid FEC/ARQ [12], where the retransmission does not necessarily need to come from the originating device but can come from any relaying node that has overheard the transmission.

In most of the simulations the channel noise is additive Gaussian, and the channel gains are independent Rayleigh distributed random variables that vary for each block transmission. This is a block fading channel, where the channel is assumed to be constant for a transmission period of a block of symbols. The path loss model considered is the log distance path loss.

The project proceeds with a discussion of point-point Hybrid ARQ using Reed-Solomon codes in Chapter 2. Also, a throughput analysis and implementation of the HARBINGER protocol for point to point case is studied for both AWGN and Fading environments. In Chapter 3, a generalization of Hybrid ARQ to multiple terminal cooperative networks is provided. The HARBINGER protocol is studied for multi terminal case. Finally, we conclude in Chapter 4.

Chapter 2: Point to Point Hybrid ARQ

An adaptive *hybrid* of FEC and ARQ (*HARQ*) using Reed-Solomon (RS) codes is considered for implementation with the HARBINGER. In our protocol, the RS code rate is adaptively varied according to the channel condition to minimize the retransmissions & maximize the throughput performance. This chapter focuses on the analysis of point to point HARQ over AWGN and Rayleigh block fading channels using Reed-Solomon codes. A brief overview of ARQ, FEC, linear codes, and RS codes is provided before proceeding to the analysis.

2.1 Automatic Repeat Request

An ARQ system is a communication system that employs a two way communication link and a means for the receiver to detect transmission errors. If there are any errors, the receiver requests a retransmission from the transmitter. The error detection strategy involves adding parity symbols to the information sequence and then checking the parity equations for errors at the receiver. Cyclic Redundancy Check (CRC) codes are typically used for error detection.

ARQ is well suited for communication networks where the transmission delay is small and information occurs in blocks [13]. All ARQ schemes partition the information sequence into blocks, add parity and control symbols to these blocks prior to transmission, and retransmit these blocks as required until the data is estimated to be correctly received. Control symbols are added to identify the message sequence number, source and destination address of the message. The error detection code will not be able to detect all error patterns. It is usually characterized by its rate and its probability to correctly detect message transmission errors. The acknowledgement of correct reception is called an ACK and is returned to the transmitter if no error occurs or if an undetectable error pattern occurs. If a detectable error pattern occurs during transmission, a negative acknowledgement (NAK) is returned to the transmitter and the message is retransmitted.

Performance metrics for an ARQ strategy include its throughput (η) and complexity(C). *Throughput* is defined as the average number of correct user data bits accepted at the receiving end in the time required for transmission of a single code bit. *Complexity* involves coding and decoding operations for error detection as well as the retransmission logic and the message buffers necessary to implement the ARQ protocol.

There are three basic ARQ strategies namely, Stop-and-Wait, Go-Back-N, and Selective Repeat. A brief description of the three ARQ strategies follows assuming the feedback channel to be error free.

2.1.1 Stop-and-Wait

Stop-and-Wait (SW) ARQ is the simplest of the three ARQ strategies. The SW ARQ transmitter simply waits for the receiver to acknowledge correct reception before transmitting the next message. The message is held in a buffer until the transmitter gets a corresponding ACK signal for the message. When channel errors occur, the message has to be retransmitted and the throughput is reduced.

2.1.2 Go-Back-N

In SW, the transmitter must sit idle while waiting for the receiver to reply with an ACK. If there are long delays in the channel, then the time that the transmitter is idle is wasted and degrades the throughput. Go-Back-N (GBN) ARQ provides better throughput relative to SW by eliminating transmitter idle time that was spent waiting for transmission delays. The idle time delays are eliminated by permitting the transmitter to transmit messages continuously until the receipt of a NAK indicating a transmission error. When transmission errors occur and are detected by the receiver, the transmitter backs up to the incorrectly received message. The transmitter retransmits that message and all subsequent messages. However, lengthy periods are encountered when the transmitter is repeating messages that have previously been received correctly. No message buffering is required within the receiver for GBN since all messages after an incorrectly received message are repeated. However, message buffering is required in the transmitter. The size of the transmitter buffer is a function of the transmission delay and the channel error probability. GBN is only better than SW if there is a delay in the channel.

2.1.3 Selective Repeat

Selective Repeat (SR) ARQ retransmits only the messages that are incorrectly received rather than the incorrect message and all subsequent messages. SR systems include the logic and message buffering necessary to reduce retransmissions to only incorrectly received messages. The throughput is improved from the previous schemes at the cost of an increase in system complexity and message overhead. Messages are to be numbered in sequence, and message buffering is required in both the transmitter and receiver. Messages are stored in the transmitter until they are acknowledged as correctly received by the receiver. The receiver must also buffer so that all messages can be delivered to the user in the correct order. Ideally, infinite size buffers are to be provided but practically the buffer size is limited thus making buffer overflow a possibility for certain message error sequences. The buffer size is selected as a function of the transmission delay, the message error probability, and the maximum tolerable probability of buffer overflow.

Out of all the ARQ schemes considered, the SR ARQ is the best ARQ scheme followed by GBN ARQ and SW ARQ with respect to throughput. If system complexity is considered, SW ARQ has the minimal complexity.

2.2 Forward Error Correction

FEC coding is a signal processing method that improves data reliability by introducing redundancy to the data sequence prior to transmission using a predetermined algorithm. The redundant data is a function of many original information bits. The receiving device has the capability to detect and correct less than a predetermined number of bit or symbol errors caused by corruption from the channel and the receiver. This coding technique enables the decoder to correct errors without requesting retransmission of the original information. FEC generally improves error performance by trading off a reduction in data rate. The maximum data rate supported by a channel is called the *channel capacity(C)*. *Shannon's theorem* states $C=W(\log_2(1+S/N))$, where W is the bandwidth and S/N is the Signal to Noise Ratio.

2.2.1 Classes of Codes

The two main categories of FEC are *block* codes and *convolutional* codes.

Block codes are based on finite field arithmetic and abstract algebra. The information sequence is divided into blocks of length *k*. Each block is mapped into channel inputs of length *n*. The mapping is independent from previous blocks, i.e., there is no memory from one block to another block. Some of the commonly used block codes are Hamming codes, Golay codes, BCH codes, and Reed Solomon codes, which use non binary symbols. Reed-Solomon codes, which are of interest here, are widely used for error control in various computing and communications media which include magnetic storage, optical storage, high-speed modems, and data transmission channels. The RS codes are of particular interest for transmission through wireless channels because of their burst-error correcting capabilities.

In *Convolutional codes*, each block of *k* bits is mapped into a block of *n* bits which are determined by the present *k* information bits and also by the previous information bits. This dependence can be captured by a finite state machine. The *Viterbi algorithm* [14] is the main decoding strategy used for convolutional codes [15]. These codes are primarily used for real time error correction and find usage in IEEE 802.11a/g standards, GSM standard, CDMA standard and also in Digital Video Broadcasting (DVB).

A class of *concatenated codes* is formed by combining block and convolutional codes. The convolutional code does most of the work and the block code (usually Reed-Solomon) corrects any errors made by the convolutional decoder. These codes are used mainly for satellite and deep space communications.

Advancements in the convolutional coding scheme have resulted in *trellis coded modulation* (TCM) [16], *low density parity codes* (LDPC) [17], and *turbo codes* [18]. TCM improves spectral efficiency by combining coding and modulation into a single operation. Turbo & LDPC codes are revolutionary in the sense that they enable reliable communications with power efficiencies close to the theoretical Shannon limit. A *turbo code* is a parallel concatenation of two convolutional codes separated by a random interleaver. The performance of a turbo code is partly due to the randomness introduced by the use of interleaver. However, there is enough code structure in it to decode efficiently. Turbo codes are useful in deep space and satellite communications, cellular and personal communication services.

2.2.2 Basics of Binary Linear Block Codes

An *(n, k)* binary linear block code is a *k-dimensional* subspace of the *n-dimensional* vector space where *n* is called the length of the code and *k* the dimension. These codes are based on Galois fields (*GF*) or finite fields. The code words can be represented in a vector notation as:

$$\mathbf{c_n} = \{(c_0, c_1, ..., c_{n-1}) | \forall c_j ; c_j \in GF(2) \}.$$

The rate of the code*(r)* can be defined as the ratio of number of information symbols *(k)* to the length of the code *(n)*. Mathematically, $r=k/n$.

A code **C** is *linear* if the XOR sum of any two codewords is also a valid codeword. If $\mathbf{c_i}$, $\mathbf{c_j}$ are two codewords, then $\mathbf{c_i}+\mathbf{c_j}$ is also a codeword. Since $\mathbf{c_i}+\mathbf{c_i}=0$, all zeros codeword must be present in all linear codes. An *(n, k)* linear block code can be specified by any set of *k* linear independent codewords. A *generator matrix* **G** for **C** is formed by arranging the *k* codewords into a $k \times n$ matrix. A *systematic* form of generator matrix can be obtained by permuting the columns and by doing some row operations on generator matrix **G**. This has the form of $[\boldsymbol{I_k} \;\; \boldsymbol{A}]$ where, $\boldsymbol{I_k}$ is a $k \times k$ identity matrix. A code is said to be *systematic* if the code words can be partitioned into the *data* and the *parity* bits.

Let **u** be the input vector of *k* data bits represented as $(u_0, u_1, ..., u_{k-1})$. Where, $u_j \in GF(2)$. Then, if **C** is linear, codeword $\boldsymbol{c} = (c_0, c_1, .. c_{n-1})$ is given by $\mathbf{u}\boldsymbol{G}$, where the multiplication is performed over the binary field GF (2). A parity check for **C** is an equation of the form:

$$a_0c_0 \oplus a_1c_1 \oplus \ldots \oplus a_{n-1}c_{n-1} = 0,$$ which is satisfied for any $\boldsymbol{c} \in \boldsymbol{C}$.

The collection of all vectors $a = (a_0, a_1, ..., a_{n-1})$ forms a subspace of $\boldsymbol{C_n}$ called the dual code of $\boldsymbol{C}$, denoted as $\boldsymbol{C}^{\perp}$. Any generator matrix of $\boldsymbol{C}^{\perp}$ is a parity-check matrix for $\boldsymbol{C}$ and is denoted by $\boldsymbol{H}$, which satisfies $\mathbf{c}\mathbf{H}^{\mathrm{T}} = 0$ for any $\mathbf{c} \in \mathbf{C}$. If $\boldsymbol{c}$ is the transmitted codeword and $\boldsymbol{y}$ the binary received codeword. The vector $\boldsymbol{e}$ called the error pattern vector is given as $\boldsymbol{c} \oplus \boldsymbol{y}$. Hence, $\boldsymbol{y} = \boldsymbol{c} \oplus \boldsymbol{e}$ and defined is a *syndrome vector* $\boldsymbol{s}$ as $\boldsymbol{yH}^T$,

$$\boldsymbol{s}=(\boldsymbol{c} \oplus \boldsymbol{e})\, \boldsymbol{H}^T = \boldsymbol{cH}^T \oplus \boldsymbol{eH}^T = \boldsymbol{0} + \boldsymbol{eH}^T = \boldsymbol{eH}^T.$$

The *Minimum distance* (d_{min}) is the smallest Hamming distance separating any two codewords i.e. it is the number of symbols in which the two codewords differ. The codewords are to be as distinct as possible for the minimum distance to be large. A code with minimum distance d_{min} can *correct* any combination of up to t errors.

Here, $t = \lfloor (d_{min} - 1)/2 \rfloor$ is called the *error correction capability* of a code. Alternatively, the code can detect any combination of up to $(d_{min} - 1)$ errors. A code $\boldsymbol{C}$ is *cyclic* if any cyclic shift of any code word is also a valid code word. The type of linear codes which are of importance to the project are Reed-Solomon codes. Reed Solomon codes are a subset of Bose-Chaudhuri-Hocquenghem (BCH) code which is a specific type of cyclic code.

Since RS codes are non-binary BCH codes, they use M-ary symbols instead of bits. A Reed-Solomon code is specified using the notation RS (n, k) with m-bit symbols. The encoder takes k data symbols of m bits each and adds parity symbols to make an n symbol codeword. A RS codeword is generated by multiplying an information polynomial $I(x)$ of degree $k-1$ by an appropriately chosen generator polynomial $g(x)$ of degree $n-k$. Hence, $C(x) = I(x)\, g(x)$.

There are $n-k$ parity symbols of m bits each. The received codeword $y(x)$ is the sum of the original codeword $c(x)$ and the error polynomial $e(x)$. The error polynomial is constructed from the estimates of the error locations and is used along with the received codeword to reproduce the transmitted codeword. Typically, a RS encoded codeword is decoded using *Berlekamp-Massey Algorithm* (BMA) [19], although more recently a new decoding algorithm by Sudan-Guruswami (SG) [20] has attracted attention due to its coding gains.

A Reed-Solomon decoder can correct up to t symbol errors or $2t$ erasures in a codeword, where $2t$ is equal to $n-k$. An erasure occurs when the position of an error symbol is known and this information is usually provided by a demodulator. Given symbol size (m), the maximum codeword length (n) for a Reed-Solomon code is given by $2^m - 1$. For example, the maximum length of a code with 8-bit symbols (m=8) is 255 bytes. For t=4, we have (n, k) = (255,247). In the example the number of code bits is 2040 but this depends on the product of n & m. RS codes can be used in conjunction with M-ary modulation or with binary modulation. A *symbol error* occurs if any one of the m bits in a symbol is incorrect. When used with binary modulation, RS codes can correct burst errors.

Reed-Solomon codes are a class of *Maximum Distance Separable* (MDS) codes [21]. The singleton bound states that given a linear code **C** with length n and dimension k, then d_{min} satisfies $d_{min} \leq (n-k+1)$. A code is said to be MDS if and only if it satisfies the Singleton bound with equality. MDS codes have the largest possible d_{min} for any code with the same values of n and k. The MDS codes have significant properties such as strong *separability* [22], strong *invertibility,* and excellent reliability performance when used for error detection and correction. When the MDS codes are punctured they result in codes which are also MDS. A code is punctured through the consistent deletion of parity coordinates from each codeword in the code. Using these properties of the RS codes, one can adaptively vary the code rate according to the channel requirements [23], there by, achieving better throughput performance.

As a figure of merit to evaluate the performance of the RS codes we analyze and plot the Codeword Error Rate (CER) vs. Energy per symbol. The CER is the ratio of the codewords that are uncorrectable after decoding to the total number of codewords transmitted at particular symbol energy strength. The results have been obtained for an AWGN channel with BPSK modulation. The probability of successfully decoding a RS codeword is calculated here. Assuming, Binary Phase Shift Keying modulation and a symbol size of 8 bits. The CER was found using the following calculations.

Probability of bit error for BPSK $(P_b) = Q\left(\sqrt{2\left(\frac{E_s}{N_o}\right)}\right)$. $\quad (2.1)$

Probability of a bit being correct $\left(\overline{P_b}\right) = 1 - P_b$ $\quad (2.2)$

Probability of all bits in a symbol being correct $= (\overline{P}_b)^m = (\overline{P}_b)^8$. (2.3)

Probability of a symbol being correct (P_s) = Probability of all bits in a symbol being correct.

Probability of a symbol being in error $(\overline{P}_s) = 1 - P_s = 1 - (\overline{P}_b)^8$. (2.4)

If, t_M is the error correcting capability of the RS *(n, k)* code then the probability of successfully decoding the codeword is:

$$P(t_M) = \sum_{i=0}^{t_M} \binom{n}{i} (\overline{P}_s)^i (P_s)^{n-i} . \tag{2.5}$$

Where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ is the binomial coefficient.

Figure 2.1 shows a plot of CER vs. SNR (E_s/N_0), for some combinations of RS *(n, k)*.

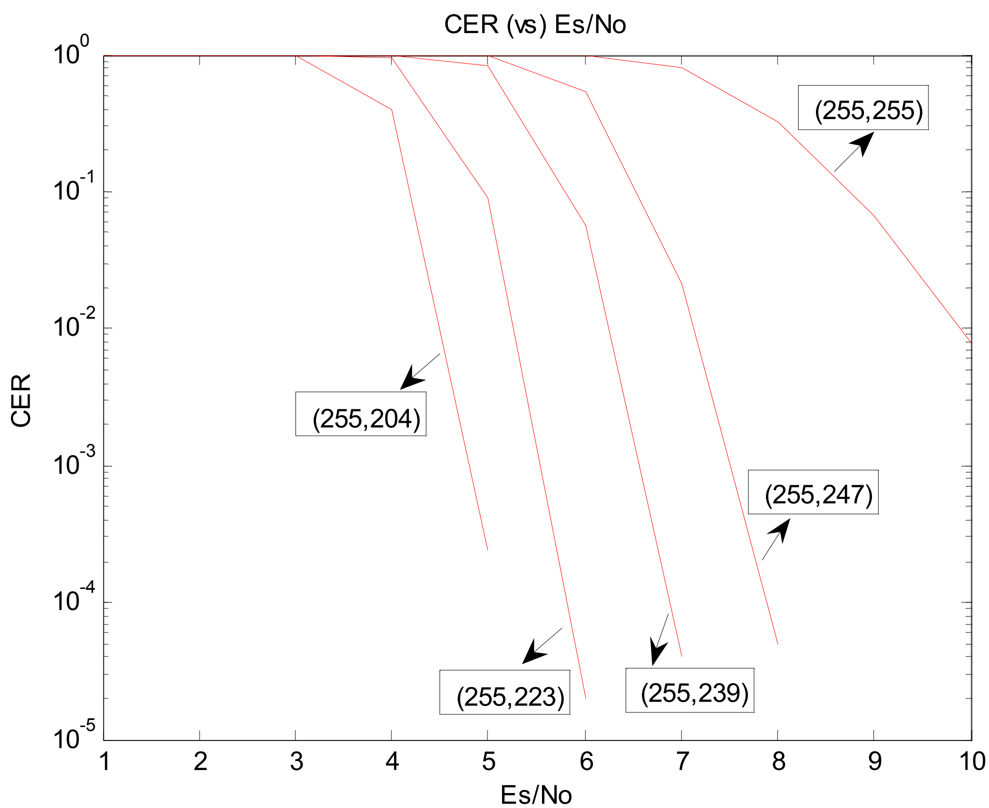

Figure 2.1: Codeword Error Rate vs. Es/No for various RS (n, k) combinations

As we increase the number of parity symbols more errors are corrected and hence the codeword error rate is reduced. This is evident from the plots.

2.3 Hybrid ARQ

Pure FEC transmits the whole codeword before transmitting the next message. Often, all the redundancy provided by the code to successfully decode the message is not required. In scenarios as in wireless channels, where the channel conditions fluctuate over a range of signal strengths (SNR), a combination of FEC and ARQ, called *Hybrid ARQ/FEC* (HARQ), is a good option. The basic Hybrid ARQ methods are *repetition coding* and *incremental redundancy*.

Repetition coding: Here, initially a codeword is transmitted and the receiver tries to decode it. If the receiver isn't able to decode the message then the same codeword is retransmitted. Multiple copies, from the retransmission of the coded packets, are *diversity combined* i.e. the various copies of the same codeword received through independent fading channels are combined to decode the message.

Incremental redundancy (IR): Here, additional redundant information is transmitted in each retransmission request. The IR scheme adapts to the varying channel conditions by essentially adapting the error correcting code rate. The receiver stores the information received in its previous transmissions, and uses it with the current redundant information to decode. This decoding operation is performed after each retransmission. At the transmitter, the information bits and CRC error detection bits are encoded by a systematic low rate mother code. Initially only the systematic part of the codeword and selected parity bits are transmitted. These bits combined together form a punctured mother code. Decoding operation is performed at the receiver. Each time a retransmission is requested, the transmitter sends additional selected parity bits possibly under different channel conditions. Each retransmission produces a codeword of stronger code [24]. This scheme is also called *code combining*. The procedure is repeated until all the parity bits of the mother code are transmitted or till the decoding is successful. The Figure 2.2 illustrates the IR HARQ process.

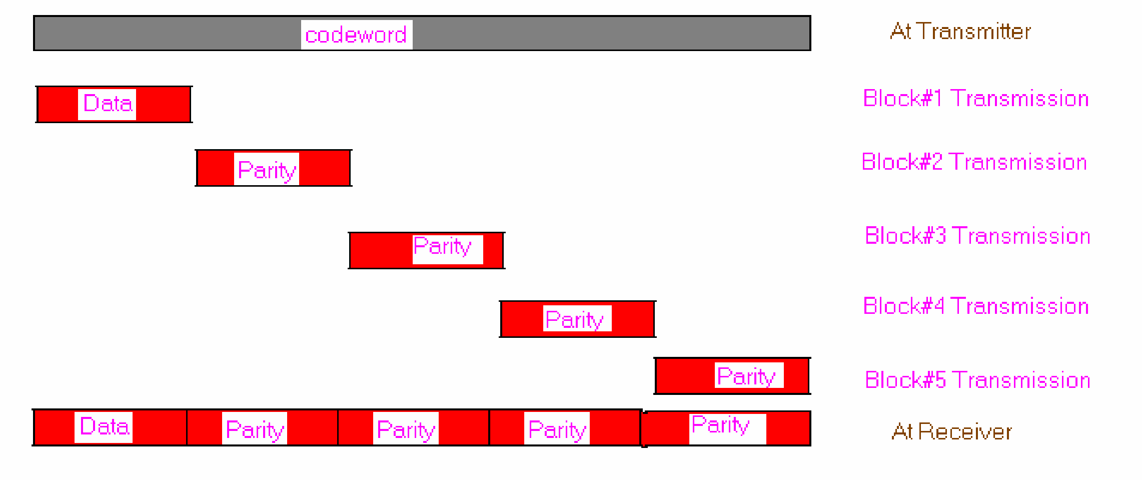

Figure 2.2: Incremental redundancy HARQ protocol.

Evaluation of the HARQ schemes is usually done by their throughput performance. First, consider the information theoretic limits on the performance of hybrid-ARQ [12]. Let $I(\gamma)$ be the mutual information between the input and output of a channel with instantaneous SNR γ. For Gaussian noise and inputs $I(\gamma) = (1/2)\log_2(1+\gamma)$. Let $I_j[m]$ denote the mutual information accumulated by node j during first m transmissions. For repetition coding, which implies that diversity combining is performed at the receiver. Hence, $I_j[m] = I\left(\sum_m \gamma_j[m]\right)$.For IRHARQ $I_j[m] = \sum_m I\left(\gamma_j[m]\right)$[12].

As $\sum I\left(\gamma_j[m]\right) \geq I\left(\sum_m \gamma_j[m]\right)$, the IR scheme is at least as good as repetition coding [11].

The *HARBINGER* protocol which is being studied in this report is an *IR HARQ* scheme. Considering the advantages of the RS codes [25, 21], we analyze and investigate the RS codes for the *HARBINGER* protocol [11]. The scheme considered here is general enough to be used in conjunction with any of the ARQ schemes, not just HARBINGER. But throughput performance and complexity may depend on the underlying ARQ protocol, and so the SRARQ scheme is suggested. Here, the protocol is analyzed for a point to point case. Later in the next chapter, the analysis is extended to the *multi terminal* case. In the following discussion it is assumed that the modulation scheme used is BPSK.

The basic scheme of our protocol is as follows [11]:

A message is encoded using a RS (*n, k*) code and is divided into *M* blocks, each of length $L = n/M$ and rate $R=k/L$. Each block can be obtained by puncturing a rate $r_M=R/M$ mother code. One block is transmitted at a time. If the receiver isn't able to decode the corresponding transmission with code combining, the subsequent block is requested. A different part of the codeword is transmitted each time, and after the i^{th} block, a receiver will pass the rate R/i code that it has until then received through its decoder. If not able to successfully decode even after M^{th} block transmission, then a transmission failure is counted and the system moves on to next message. The protocol reacts to an increase in channel noise by sending additional parity symbols in the very next packet transmission.

A throughput analysis is carried out for the Rayleigh block fading channel, where the channel remains constant for the period of a block transmission and the signal strength varies as a Rayleigh random variable. Before proceeding with the block fading case, we study the protocol for AWGN channel with a particular value of symbol energy strength. The analysis done here for the HARBINGER with RS codes, is an extension of the analysis provided for the codeword error probability of RS codes provided previously in this chapter. In the analysis, it is assumed that after M=5 block transmissions if the codeword is still not decodable then a failure is reported. A record of the number of successful and failed codewords is maintained to calculate the average number of block transmissions required for a successful message reception.

The throughput is given by the following equation:

$$\text{Throughput}(\eta) = \frac{k}{(t_b N_R)} P(success),$$

where, k is the number of information bits, t_b is the total time taken for transmission of a block, $P(success)$ denotes the probability of a message being successfully decoded, N_R is the average number of block transmissions required for a message being successfully received. However, we deal with *normalized throughput* (η_{norm}) here, which is given as:

$$\eta_{norm} = \frac{RP(success)}{N_R}, \tag{2.6}$$

where, $R=k/L$ and is in units of successfully transmitted data bits per channel use.

N_R can be calculated by performing the following manipulations. Let X be a random variable that represents the number of block transmissions before the message is correctly decoded. Here, X= l means that the message is correctly decoded after the l^{th} block transmission. Let the probability that a message is correctly decoded after the l^{th} block transmission be $P_X[l]$.

Then, the mean of X is $N_R = E[X] = \sum_{l=1}^{\infty} l P_X[l]$. $\quad (2.7)$

Let t_l be the correcting ability with l block transmissions and $P(t_l)$ be the probability of successfully decoding a codeword with error correcting ability of t_l which is given by equation (2.5). Since in the hybrid-ARQ discussed, the error correcting capability of the codeword increases with every block transmission, the probability of a message being correctly decoded after the l^{th} block transmission is given by:

$$P_X[l] = (1 - P(t_1))(1 - P(t_2))\ldots(1 - P(t_{l-1}))P(t_l). \quad (2.8)$$

In the protocol, it is assumed that only a maximum of M block transmissions are allowed. Hence, the average number of block transmissions in equation (2.7) is a summation not up to infinity, but only until M. If the message is still not decoded after M block transmissions a failure is counted and this count has an effect on *P (success)*.

In the protocol considered, message size k=51 symbols and each symbol is 8 bits long. The maximum codeword size one can have with that combination is 255 symbols. So a maximum of 102 errors are correctible with a (255, 51) RS codeword. Initially the transmitter sends the original message of 51 symbols without any redundant symbols. If the channel is good enough, then message is received without any errors and we proceed with the transmission of the next message.

However, if errors are not correctible, a retransmission request is made. More redundant information is sent with each retransmission request till the lowest possible code rate is achieved. The block size considered here is L= 51 symbols. Since M =5, the lowest possible code rate=1/5. For each retransmission request of the same message an additional 51 symbols of the codeword are transmitted. So the error correcting capability increases with each additional transmission. Here the size of *n* varies with each retransmission due to the addition of the redundant symbols with each transmission. As M=5, we have t_1, t_2, t_3, t_4, $t_{5;}$ where, t_1=0, t_2=25, t_3=51, t_4=76, t_5=102.

The following plots show simulation results for the protocol in AWGN. Figure 2.3 shows a histogram of the number of block transmissions as a function of the SNR (E_s/N_0), and gives an idea of the minimum number of block transmissions required for the successful decoding of a codeword at particular signal strength. We observe that as the signal strength (channel condition) improves, the number of transmissions required for a codeword to be successfully decoded, decreases.

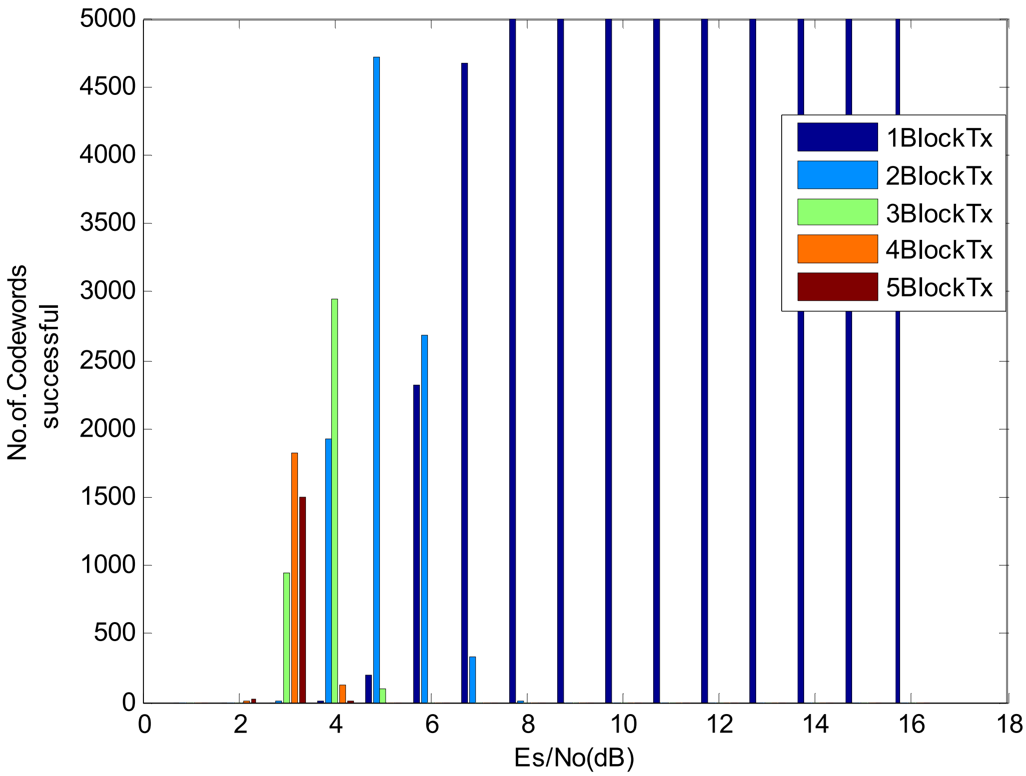

Figure 2.3: No. of codewords successful in each block transmission with a total of 5000 codewords being transmitted at each E_s/N_0

A major limitation of performance with the protocol is that there are no parity bits in the first block. Because of this, at least two blocks must be transmitted at all but for the highest E_s/N_0. This problem can be alleviated by transmitting a few parity bits in the first frame. For instance, some parity from each of the last four frames can be borrowed and sent in the first frame. In this way parity symbols can be introduced, but this is not advised, as it adds to the complexity of the protocol because now the block size is variable. Another option is to reduce the message size k and introduce some parity information in the first frame itself.

A normalized throughput plot is shown in Figure 2.4 for the three cases where the number of information bits k in the codeword is varied. Message sizes of 51, 47 and 39 symbols are considered and the corresponding number of parity symbols in the first frame is 0, 4, and 12, respectively. All the message symbols are sent in the first frame. By using smaller values of k we obtain throughput advantage at low SNR. However, for SNR values of 2-4 dB a message size of 51 performs better than other two. This is because at those values at least two or three block transmissions are required to decode the message and so with a message size of 51 symbols we transmit more information.

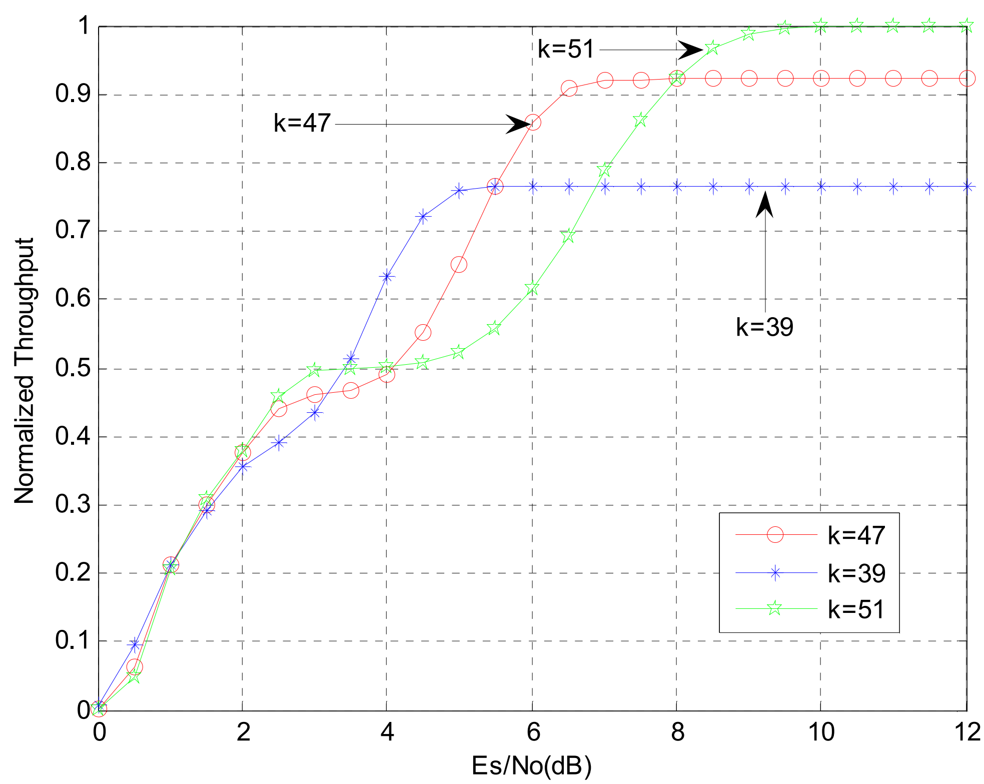

Figure 2.4: Normalized Throughput vs. Es/No (AWGN Channel)

For a block fading channel, the channel remains constant for the period of a block transmission i.e. the instantaneous SNR varies as a Rayleigh random variable for each block transmission. However, one can obtain the average SNR at a particular distance from the transmitter by using the log distance path loss model. Throughput is obtained for a particular value of average SNR.

The throughput calculation is the same as that of AWGN case except that the instantaneous received power varies exponentially for each block transmission as described for Rayleigh fading channel in Chapter 1. The throughput plot obtained for the hybrid ARQ with RS codes and BPSK modulation in a block fading and AWGN channel is as shown in Figure 2.5. Here, the codeword considered is with a message size of 51 symbols and M being 5. The strongest code obtained has a code rate of 1/5.

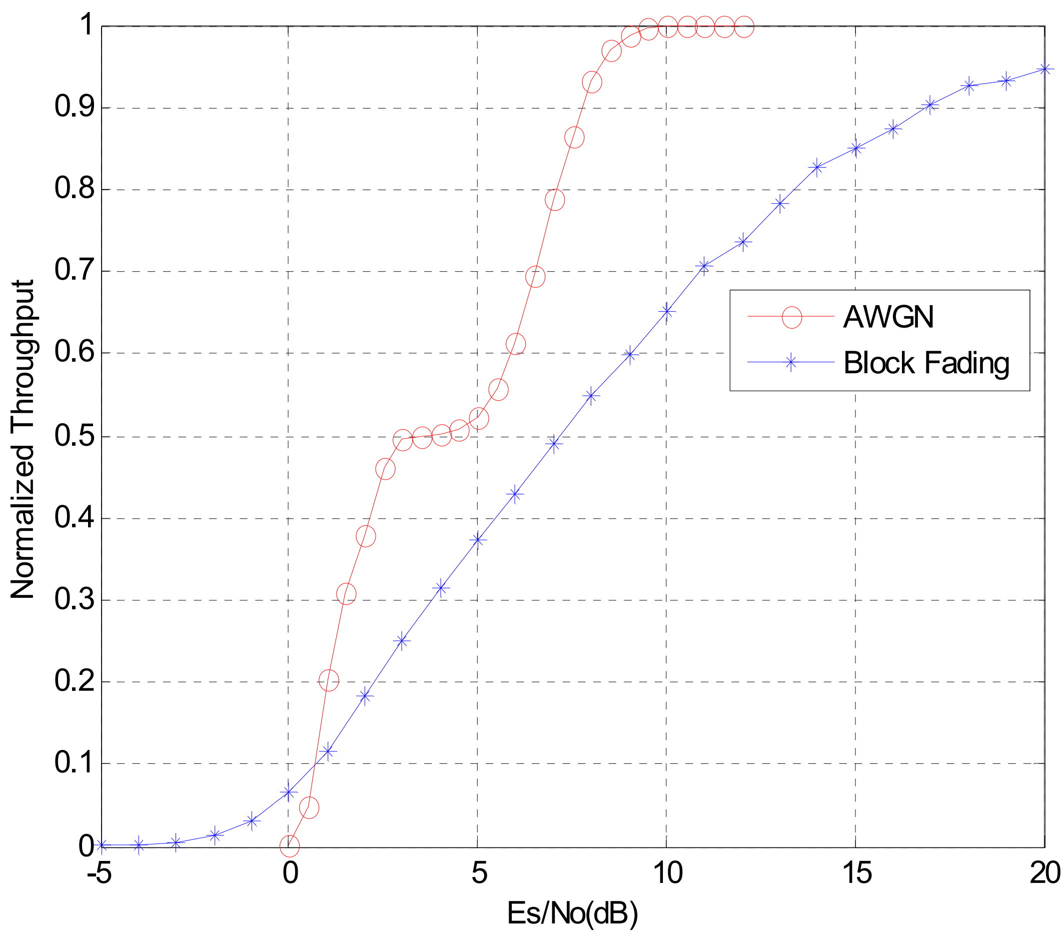

Figure 2.5: Normalized Throughput vs. Es/No (Block fading and AWGN channel).

More variations of M, k could be explored. Different channel coding schemes with various combinations of modulation could also be studied. The turbo codes, LDPC codes and fountain codes [26] such as raptor codes [27] are some of the suggested coding methods. Fountain codes are also a good option for implementation with HARBINGER as stronger codes can be obtained through puncturing. Fountain codes have no lower limit on the mother code rate R_m, or equivalently no maximum limit on the number of blocks M.

Other variations in the protocol like 51 frames of 5 symbols and 255 frames of 1 symbol can also be considered. Here, in the protocol if after certain number of transmission attempts the codeword is still un-decoded, it is counted as a failure and protocol moves on to the next codeword transmission. This is not an ideal case, but for making the throughput calculations easier and for stopping the simulation from running to infinity we do this. A more robust protocol could perform some other operation after the M^{th} block is transmitted, such as switching to a diversity combining strategy, but this adds to the overall complexity.

Chapter 3: Multi Terminal Hybrid ARQ

In this chapter, we extend our study to the multi terminal case in which we first generalize the concept of hybrid ARQ and discuss the HARBINGER protocol for the multi terminal case using RS codes and BPSK modulation. Some of the results comparing the throughput performances of the multihop, relay, and no relay case are provided. Initially we deal with the simple case of three terminals and then move to relay case with more relays.

3.1 Generalized Hybrid ARQ & HARBINGER

As the number of nodes in a network increases and when the network is constrained to use only point-to-point links, the average throughput furnished by each node decreases [28]. Relaying uses the distributed spatial diversity found in wireless networks to obtain better performance. A practical way to implement relay networks is to generalize the concept of hybrid-ARQ [11], where in the retransmitted packets could come either from the original source radio or from the other relays nearby that overhear and decode the transmission. The destination combines the information received from both the source and relay, thereby achieves diversity. Relay networks considered, comprise of a source, a destination and one or more relays.

In a multiple relay case, the scheduling of the relays becomes a main issue. A viable solution for the relay scheduling problem is the GeRaF protocol proposed by Zorzi and Rao [29, 30]. The assumptions made in the GeRaF protocol are:

- All the nodes know their own as well as that of the destination position.
- A contention scheme exists to determine the forwarding node.

In GeRaF, the source broadcasts to all nodes within range. The node that is closest to the destination and has successfully decoded the message is selected to serve as the relay and transmits the message next. These multiple relays operate over orthogonal time slots. HARBINGER is a combination of Hybrid ARQ and GeRaF. It adds to the distributed transmit diversity advantage to GeRaF by allowing the nodes to maintain previously received information concerning each message [11]. HARBINGER is an incremental redundancy strategy.

3.1.1 System model [11]

Consider a group of nodes N within a geographical region R. Each node is assumed to have a single half duplex radio and a single antenna. The scenario considered has a source, a destination, and some interconnected relay nodes. The nodes all are distributed linearly in the region R. The nodes are numbered starting from 1 up to N. The source and destination are assigned 1 and N respectively. All the relay nodes are numbered starting from 2 according to their distance from the source, with the farthest relay node being assigned N-2. This is as shown in Figure 3.1 below.

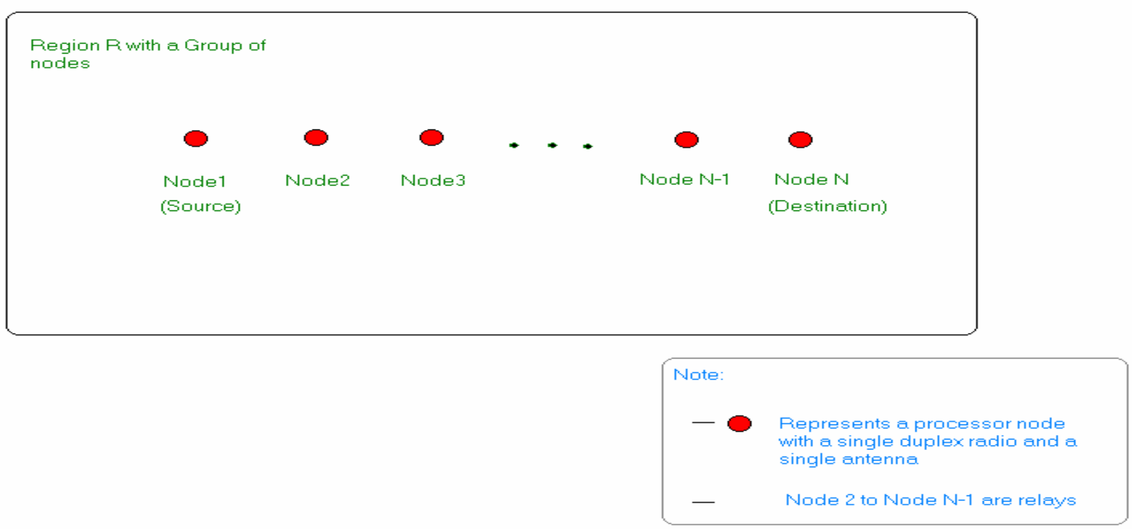

Figure 3.1: System model.

It is generally assumed that the nodes are placed at an equal distance from each other unless otherwise mentioned. In the model, the transmit frequency f_c=2.4 GHz, path loss coefficient μ=3, reference distance d_0=1 meter, and noise floor = -100dBm. The channel is assumed to be block fading with each block transmission having a different fading gain. For protocol simulation we consider RS codes and BPSK modulation. All the nodes are assumed to be at a fixed position in the region R. The relays are decode-forward relays, which forward messages only when they are able to successfully decode the message. Time is divided into slots *s* of equal duration, during which a node must either transmit or receive but not both. The hybrid ARQ protocol concept discussed in the previous chapter is combined with relaying to obtain better throughput performance in a network.

The source begins by encoding a message into a *(n, k)* RS codeword of symbol size *m* bits. The RS codeword size considered is 255 symbols. The codeword is punctured and broken into five blocks each of 51 symbols. Let $S = \{s_1, s_2, \ldots, s_5\}$ denote the five contiguous time slots over which the five blocks are sent. The node that transmits during the slot *s* is called *transmitting node*, and is denoted by *T(s)*. This node is determined from a set of nodes belonging to the *decoding set D(s)*. Decoding set *D(s)* is a collection of nodes that have successfully decoded the message or have knowledge of the codeword, at the start of slot *s*. The node in the decoding set that is closest to the destination is selected as the *transmitting node*. All the nodes, except the ones in the decoding set receive and process the information to decode the message correctly.

Initially only the source has knowledge of the codeword and is the only member in the decoding set. Hence, during the first time slot the source starts transmitting the first 51 symbols of the systematic RS codeword. All the nodes receive the blocks and try to decode the message. The nodes which are successful in decoding the message correctly are included in the decoding set. The decoding set is scanned for the destination node. If the destination node is found to be a member of the decoding set, a success is indicated and a track of the number of retransmissions is maintained. If the destination isn't able to decode the block transmission then a retransmission request is made by it. After the first block and at the start of the i^{th} block transmission, the decoding set has the source and all the relays that have previously received blocks with a sufficiently low error count such that the message can be successfully decoded. During subsequent slots s_i, for $2 \le i \le 5$, *any* node in the decoding set that is closest to the destination re-encodes the message and transmits the next block of the mother code. These block transmissions occur until all the five blocks of the codeword are exhausted or the destination has successfully decoded the message.

After the five blocks have been transmitted and if still the destination isn't able to decode the message correctly, the steps mentioned above in the protocol are repeated again with the first block transmission. However, the decoding set is not refreshed over again. The source now is the node in the decoding set that is closest to destination. This repetition occurs till a maximum latency period has been exceeded or the destination is successful in decoding the message. If the destination isn't able to decode still after reaching the maximum latency limit, a failure is indicated.

The destination can decode the message if the received block is within the error correction capability of the RS code and, if successful, will broadcast an ACK. Otherwise more parity information is sent in the next transmission by the relaying node. The node in the decoding set *D(s)* that is closest to the destination is selected for relaying, this is equivalent to picking the node whose channel to the destination has the highest average SNR. The average SNR is predicted using the log distance path loss model.

With every retransmission request either the source or relay transmits a new frame of parity information, thereby, increasing the error correcting capability of the received codeword. Hybrid ARQ transmits new blocks of the codeword until the destination successfully decodes the message or all blocks have been transmitted. However, after a maximum latency has been exceeded, a transmission failure is indicated. The throughput calculation is similar to the procedure followed in Chapter 2.

3.2 Three Terminal Case

First consider the three node model with a source, destination and a single relay. A source transmits a block initially and is the only node in decoding set. The relay and destination receive a frame with a number of errors that depends on the instantaneous SNR, which varies from block to block. The model is similar to the one shown in Figure 3.2.

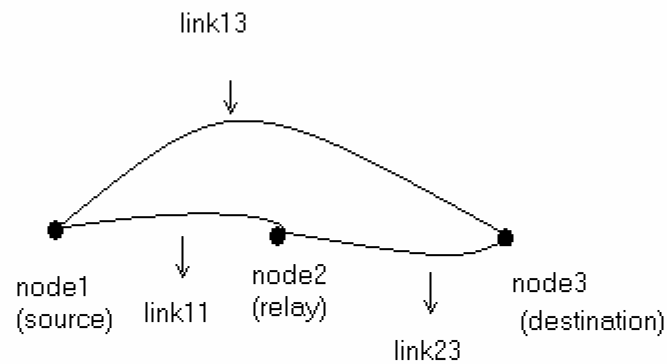

Figure 3.2: Three terminal model.

Node1, node2, node3 are source, relay, and destination, respectively. The channel link between each of the nodes is variable and is assigned as shown in Figure 3.2.

Figure 3.3 shows the normalized throughput performance for the no relay case and the relay case with various source-destination separation distances. The relay considered here is midway between the source and destination. For small separation distances there is hardly any difference in performance of no relay case and relay case. But, as the distances increase the usage of relay certainly shows an improvement in throughput performance as compared to when no relay is used. We determine the best relay position for the relay case in the next section.

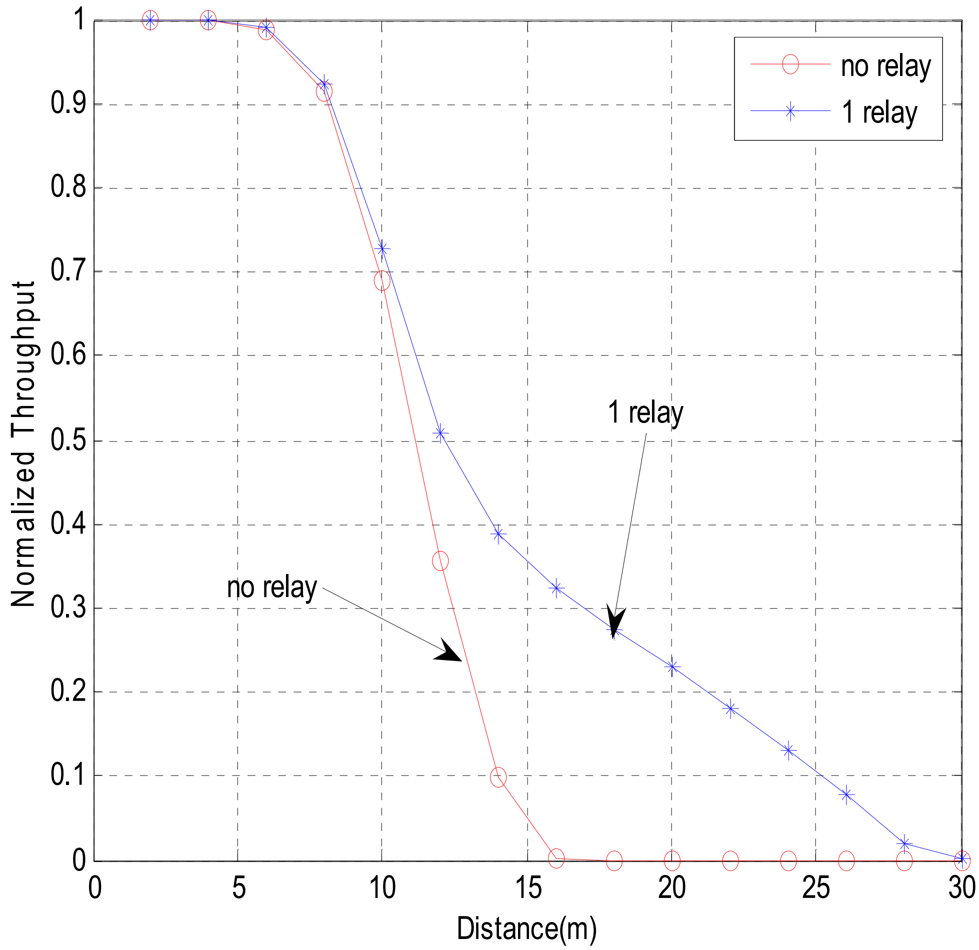

Figure 3.3: Normalized Throughput vs. Source-Destination separation distance in meters for the relay and no relay case.

3.2.1 Variations in the Relay Position

Figure 3.4 compares the throughput performances for various positions of the relay. The position of the relay between the source and destination is varied. Let the separation distance between the source and destination be d. We consider the following four cases:

a) No relay case.

b) Relay is placed at a distance $\frac{d}{2}$ from the source, which is relay midway.

c) Relay is placed at a distance $\frac{d}{3}$ from the source, which is relay closer to source.

d) Relay is placed at a distance $\frac{2d}{3}$ from the source, which is relay closer to destination.

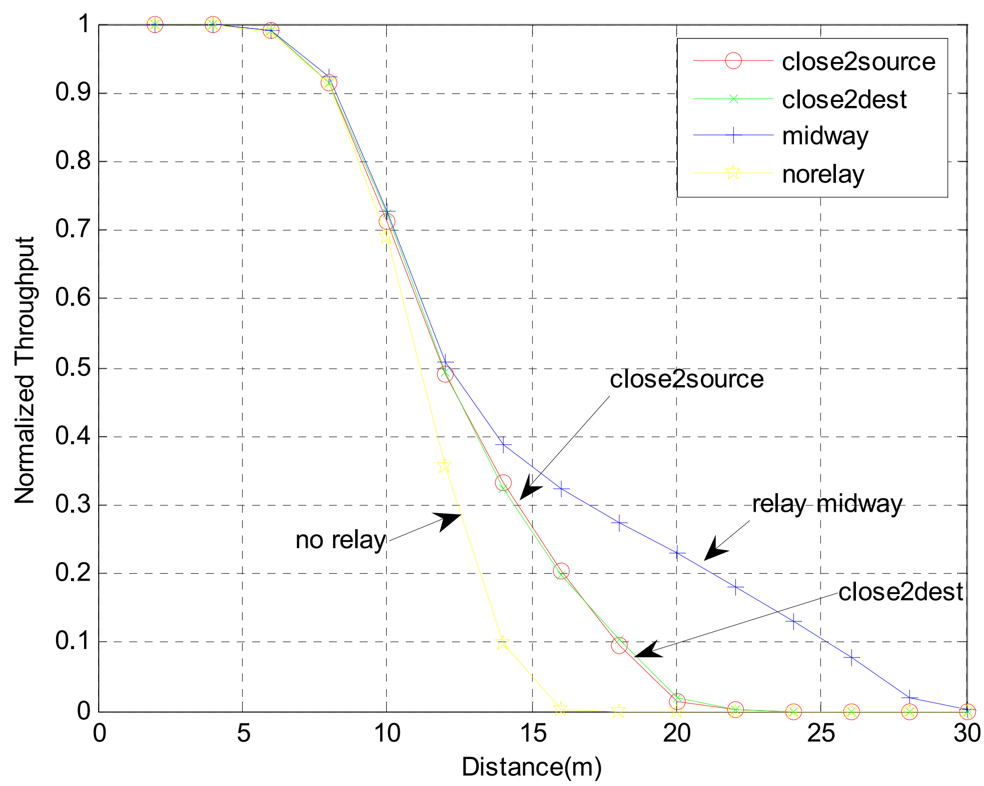

Figure 3.4: Normalized Throughput vs. source-destination separation distance in meters.

The relay has a better throughput performance when in midway between source-destination than when it is nearer to source or destination. Also, placing a relay is certainly beneficial than placing no relay irrespective of the relay position.

3.2.2 Comparison with Multihop

Figure 3.5 compares the normalized throughput performance with source-destination separation distances for the no relay, relay, and multihop case. We observe that for small distances, both the no relay case and the relay case are better than multihop. However, using the relay proves to be advantageous than the no relay case for larger distances. But the multihop performs better than the relay case at larger source-destination separation distances. This can be described by considering the following example which happens with high probability.

For HARBINGER:

a) Node 1 transmits first frame of the message and Node 2, Node 3 listen.
b) Node 2 decodes the message. So Nodes 1, 2 are in decoding set. But Node 3 doesn't decode successfully due to errors.
c) Node 2 transmits 2^{nd} frame with parity but Node 3 again is unsuccessful in decoding the message due to the parity provided by 2^{nd} frame is not enough to correct errors occurred in 1^{st} frame transmission.
d) So the Node 2 transmits 3^{rd} frame with more parity and hence now the Node 3 is able to decode message now.

Hence three frame transmissions are required for successful message decoding at Node 3.

For multihop:

a) Node 1 transmits first frame of the message and Node 2, Node 3 listen.
b) Node 2 decodes the message. So Nodes 1 and 2 are in decoding set.
c) Node 2 starts the protocol again with the 1^{st} frame transmission and the Node 3 decodes successfully.

Hence, only two frame transmissions required for multihop. Thus, the multihop is performing better than HARBINGER considered at large inter-relay distances.

What is happening is that the link between node 1 and node 3 is so bad, that there are too many errors in the transmission of the first frame from node1 to node 3. The link from node 2 to node 3 is much better, and so there are much fewer errors in the second frame from node 2 to node 3. However, node 3 must decode using both the first frame (which has lots of errors) and the second frame (which only has a few errors). Despite the good second frame, the code is simply not powerful enough to make up for the errors in the first frame.

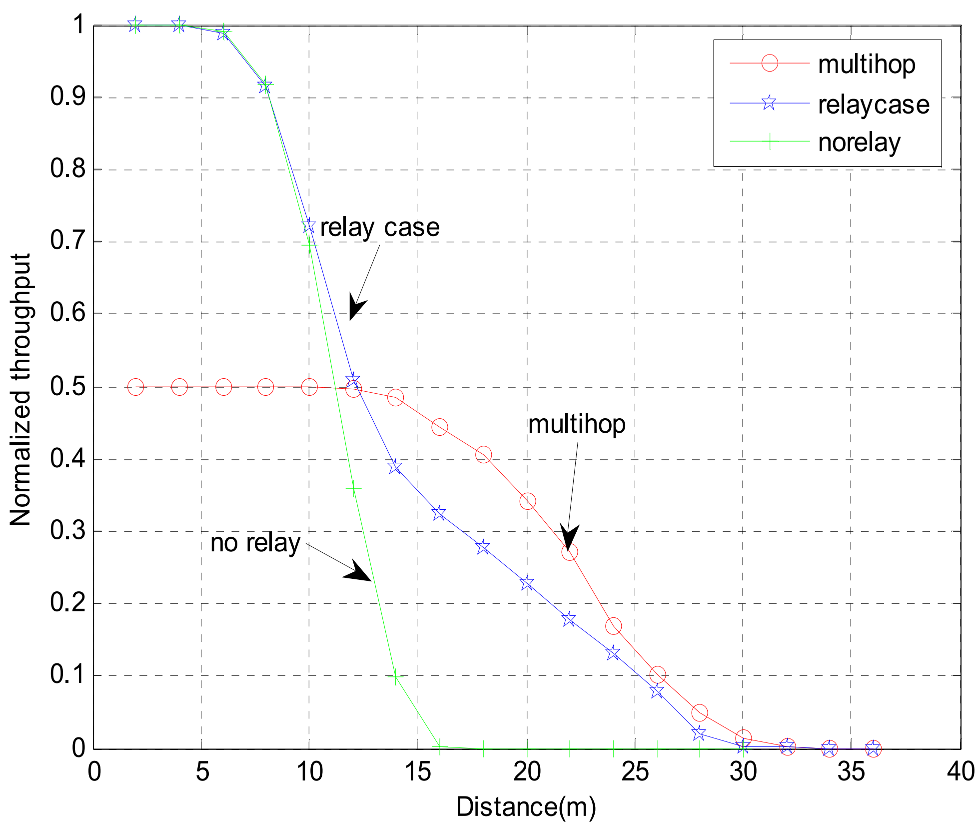

Figure 3.5: Normalized throughput vs. source-destination separation distance in meters for relay, multihop, and no relay case (3 terminal case).

3.3 MultiTerminal Case

3.3.1 Variation in Number of Relays

We now consider the multi-terminal case where more than three nodes are distributed linearly in a region R. A comparison is made in Figure 3.6 between the normalized throughput performance and the source-destination separation distance with variations in the number of relays between the source and destination. The following cases are considered:

a) Ten relays are placed between source and destination.
b) Five relays are placed between source and destination.
c) Three relays are placed between source and destination.
d) Single relay is placed between source and destination.

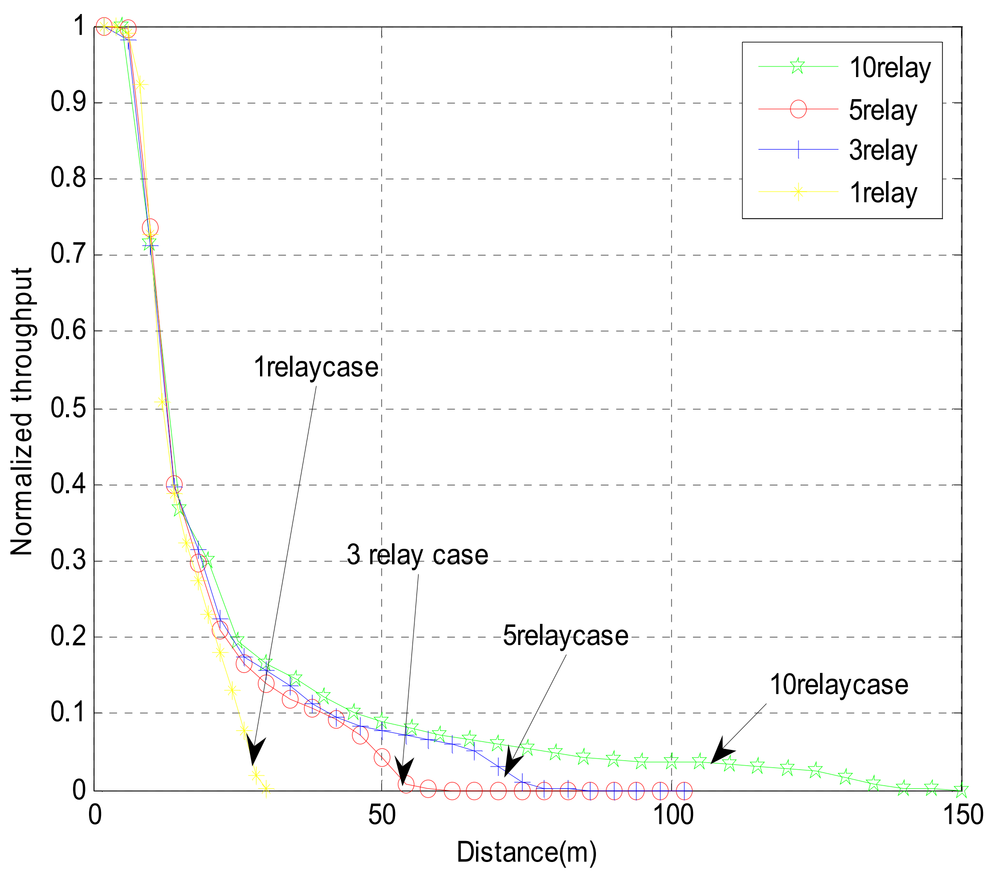

Figure 3.6: Normalized Throughput vs. source-destination separation distance in meters for variations in the number of relays between the source & destination.

For small source-destination distances, an increase in the number of relays between source and destination doesn't show any performance improvement. However, an increase in the number of relays between source and destination yields better throughput for higher source-destination distances. So at higher source-destination distances placing more relays is beneficial as is evident from plots.

3.3.2 Comparison with Multihop

A throughput performance comparison is made for multihop, no relay, and relay case with 10 relays in Figure 3.7. Ten relays are placed between the source and destination such that the distance between each of the nodes is same. For small source-destination distances, the no relay case and relay case perform better than the multihop as in multihop each and every node in the predetermined route has to decode.

Multihop here is a cascade of point-to-point links. Until certain source-destination distances, the relay performs better than multihop and the no relay case. But for large distances multihop performs better than the other two. The fact that multihop performs better than HARBINGER can be explained using an argument similar to that given for the three terminal model.

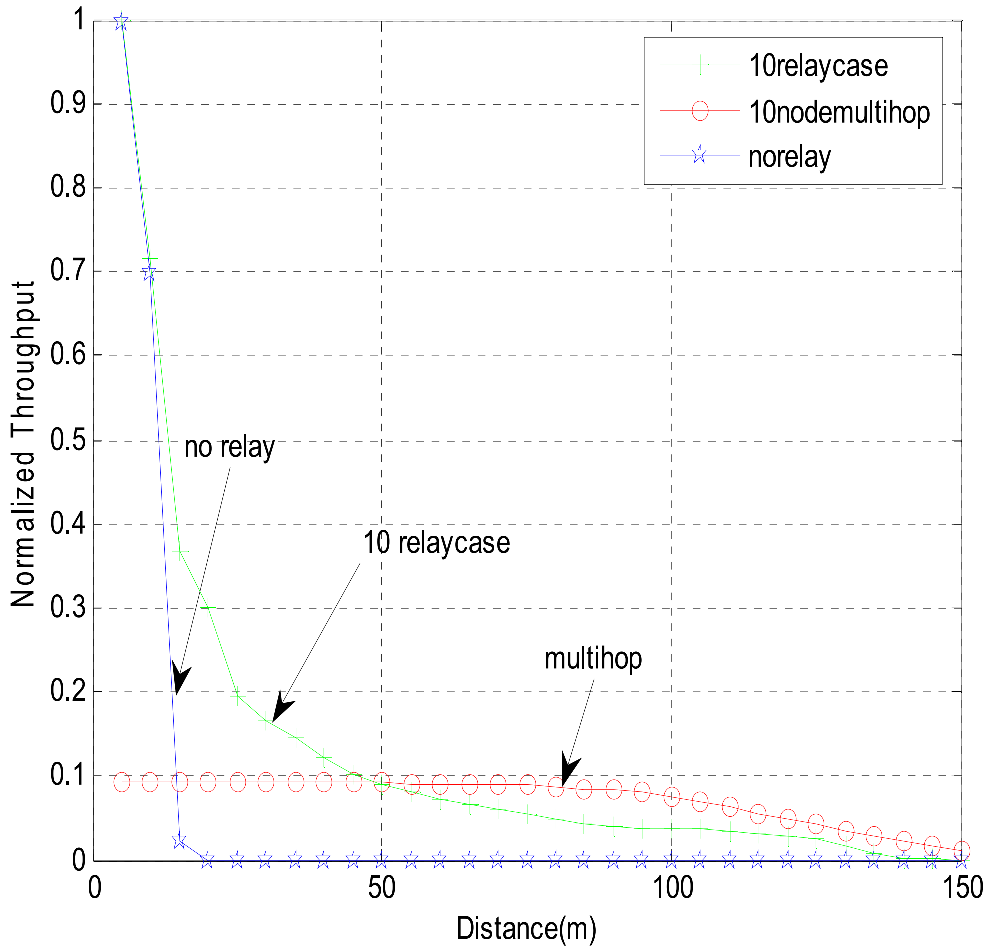

Figure 3.7: Normalized throughput vs. source-destination separation distance in meters for relay, multihop, and no relay case (10 relay case).

Chapter 4: Conclusion

The project mainly deals with the study and the practical implementation of the HARBINGER protocol. The HARBINGER protocol is analyzed with the simple Reed-Solomon codes and BPSK modulation. The work by Bin Zhao [11] mainly dealt with information theoretic perspective of the protocol. Here the Reed-Solomon codes are considered since they have good error correcting capabilities in block fading environments as found in wireless networks and also because there are many readily available commercial RS *codecs* that can be applied to low cost networks. BPSK modulation scheme is considered as it has the minimal complexity in its practical implementation and lends itself to easy analysis. The focus of the study is on a general class of wireless embedded networks that are decomposed into clusters comprised of low cost radio devices including a source, a destination, and one or more relays.

The advantage with relaying is that no routing information is required at the nodes as the best route is determined adaptively based on the channel conditions and the geographical location of the nodes. HARBINGER effectively expands its coverage area through type II hybrid-ARQ retransmission. Reed-Solomon codes are a good choice as they belong to the family of MDS codes and can form systematic codes. Incremental redundancy can be practically implemented by puncturing the RS codes.

HARBINGER is considered for point-to-point case with the actual RS codeword being punctured and broken to five blocks. With each retransmission request, extra parity information is sent to decode the message, thereby increasing the chance of the message being correctly decoded. With no parity information sent in the first frame transmission, at least two frame transmissions are required for the decoding of a message except at high received SNR's. This situation can be improved by modifying the protocol and sending some parity in the first frame transmission itself. In the multi terminal analysis of the protocol it has been found that the relay outperforms the no relay case. Also placing the relay midway between source-destination proves to be more efficient than placing the relay closer to source or destination. So, equally spaced relays are suggested. Increasing the number of relays for a certain source-destination separation distance leads to a better throughput performance at higher separation distances.

Comparing the performance of multihop and relay case for the HARBINGER protocol with the RS codes leads us to the conclusion that relaying is not always better than multihop. Multihop outperforms relaying at higher source-destination distances. This is because the link quality degrades so much that it becomes highly impossible for the retransmission from the relay to correct all of the errors. This is actually a very interesting characteristic, and something that is not observed when dealing with the information theoretic results as studied by Zhao. But when using a real code, this is a very important practical consideration. This result is due to the limited capability of a code to correct errors and also mostly a consequence of using hard decision decoding. A soft decision decoder would place more weight on the “good” blocks and less weight on the “bad” blocks. The weakness with this protocol is that it is limited to a fixed number of blocks per codeword (M). One of the solutions to this problem is the use of *fountain codes* which are *rateless* [27].

References

[1] A. Nosratinia, T. E. Hunter and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine,* pp. 74-80, 2004.

[2] T. S. Rappaport, *Wireless Communications Principles and Practice,* 2nd Edition, Prentice Hall, 2002.

[3] B. Sklar, "Rayleigh fading channels in mobile digital communication systems part I: Characterization," *IEEE Comm. Mag.*, vol. 35, pp. 90- 100, July 1997.

[4] M. C. Valenti, "Random variables in wireless networks", Notes*: Wireless Networking CPE 493g*, 2005.

[5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol.44, pp. 2619–2692, Oct. 1998.

[6] S. Lin and P. S. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," *IEEE Trans. Commun.*, vol.COM-30, no.7, pp.1701-1719, July 1982.

[7] M. C. Valenti and B. Zhao, "Scaling cooperative diversity to large networks," *NATO Workshop on Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks*, June 2-3, 2004.

[8] 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*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[9] T. Hunter and A. Nosratinia, "Cooperation diversity through coding," *Proc. IEEE Int. Symp. on Inf. Theory,* Lausanne, Switzerland, Jun./Jul. 2002, p. 220.

[10] J. N. Laneman and G. W. Wornell, "Distributed space–time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, pp.2415–2425, Oct. 2003.

[11] B. Zhao and M. C. Valenti, "Practical relay networks: A generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Communications (Special Issue on Wireless Ad Hoc Networks)*, vol. 23, no. 1, pp.7-18, Jan. 2005.

[12] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. IT-47, pp.1971 – 1988, July 2001.

[13] R. E. Ziemer and R. Peterson, *Introduction to Digital Communication*, 2nd Edition, Prentice 2001.

[14] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT–13, pp. 260–269, April 1967.

[15] G. D. Forney Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, March 1973.

[16] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets part I: introduction," *IEEE Communications Magazine*, vol. 25-2, pp. 5-11, 1987.

[17] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, pp. 399-431, Mar. 1999.

[18] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes," *ICC*, pp 1064-1070, 1993.

[19] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.

[20] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon codes and algebraic geometry codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1757-1767, Sept. 1999.

[21] S. B. Wicker and V. K. Bhargava, *Reed-Solomon Codes and Their Applications*. Piscataway: IEEE Press, 1994.

[22] J. Du, M. Kasahara and T. Namekawa, "Separable codes on Type-II hybrid ARQ systems," *IEEE Trans. Commun.*, vol.36, no.10, pp.1089-1097, Oct.1988.

[23] C. V. Verikoukis and J. J. Olmos, "An efficient type-II hybrid-ARQ protocol using punctured RS codes for wireless ATM networks," *Proc. IEEE VTC-Fall*, Amsterdam, the Netherlands, pp.1725-1729, Sept.1999.

[24] E. Soljanin, R. Liu, and P. Spasojevic, "Hybrid ARQ with random transmission assignments," *DIMACS Workshop on Network Information Theory,* May 2003.

[25] S. B. Wicker, "Type-II hybrid-ARQ protocols using punctured Reed-Solomon codes," *Proc. of the 1991 IEEE Mil. Commun. Conference,* McLean, Virginia, pp. 52.2.1– 52.2.6, Nov. 1991.

[26] J. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1528-1540, October 2002.

[27] R. Palanki and J. S. Yedidia, "Rateless codes on noisy channels". Available at: http://www.merl.com/reports/docs/TR2004-038.pdf

[28] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, pp. 388–404, Mar. 2000.

[29] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Multihop performance," *IEEE Trans. Mobile Comp.*, vol. 2, pp. 337–348, Oct. 2003.

[30] M. Zorzi and R. R. Rao, “Geographic random forwarding (GeRaF) for ad hoc and sensor networks: Energy and latency performance,” *IEEE Trans. Mobile Comp.*, vol. 2, pp. 349–365, Oct. 2003.