Cooperative Diversity through Reliability Filling

John M. Shea, Tan F.Wong and Arun Avudainayagam
Wireless Information Networking Group (WING)
Department of Electrical and Computer Engineering
University of Florida

May 21, 2003

Abstract

Recently, cooperative diversity techniques have been investigated for use in scenarios that do not permit the use physical antenna arrays. Most of these techniques have been based on some form of repetition coding. In this paper, we present a technique called Reliability Filling that achieves cooperation diversity using error control codes and soft input-soft output decoders. Unlike earlier techniques, the most important aspect of our technique is that diversity benefits are exploited during iterative decoding instead of before decoding. Also, our scheme does not require full decoding at any of the cooperating nodes. It will also be shown that the combining overhead for our schemes is a fraction of optimal combining overhead. A practical iterative technique to achieve reliability filling is also presented.

1 Introduction

The idea of users cooperating to achieve spatial diversity has received a lot of focus from researchers in recent years [1, 2, 3, 4, 5]. Diversity achieved when users in a network collaborate and improve each others performance has been termed cooperative diversity (Multiuser diversity). Cooperative diversity techniques are network based alternatives to improve performance in scenarios that do not allow the use of physical antenna arrays. For example, the small size of modern radios prohibit the use of antenna arrays in an ad hoc network or in the downlink of a cellular system. Cooperative techniques exploit the broadcast nature of the channel and the inherent antenna array that is present in any wireless network. At least a few nodes in a network are well separated and thus are very likely to have independent channels to the transmitter. These nodes are capable of receiving any ongoing transmissions and can act like elements of an antenna array. This is referred to as a distributed array [6],[5]. since the elements are not physically connected to each other. For the rest of the paper, the nodes other than the intended destination that belong to the distributed array will be referred to as relay nodes.

The optimal combiner for independently received signals is a maximal ratio combiner. This would be the simplest cooperation strategy that the nodes can employ. This is also optimal in terms of error/outage performance and provides full diversity in the number of nodes that cooperate. However, the lack of physical connections in distributed arrays makes maximum ratio combining(MRC) a bandwidth expensive procedure. For example, for an additive white gaussian noise (AWGN) channel, performing MRC would involve transmitting the soft
Thus efficient techniques are needed that have a small overhead but that can still provide all the benefits promised by cooperative diversity. In this paper, we propose a technique called reliability filling that achieves full diversity and yet requires only a fraction of the overhead required for MRC. The paper is organized as follows. The next section describes existing approaches to cooperative diversity and summarizes past research on this topic. A summary of collaborative decoding, our previous work on cooperative diversity, is also given in this section. A generalization of this technique called reliability filling is introduced in section III. Simulation results are given in section IV and the paper is concluded in section VI. The system model and a summary of our previous work addressing a few drawbacks of existing cooperative diversity schemes is given in Section III. Reliability filling is an idealized technique and a practical implementation to achieve reliability filling is given in section IV. In Section V, we present simulation results and conclude the paper in section VI.

2 Background

In this section we present a brief history of cooperative techniques and highlight some important contributions and key ideas in cooperative diversity.

2.1 Related research

Though cooperative diversity has grabbed the attention of researchers in recent years, the idea of user cooperation has existed for a long time. The concept of the relay channel that was proposed and studied in the early 70’s can be thought of as the incipient stages of cooperative techniques. In this setting, an intermediate node called the relay received information from the source, processed the information and re-transmitted some “doping” information to the destination in order to help decoding at the destination.

In 1979, Cover and Gamal put forward three different approaches to achieve user cooperation in a relay channel [7]. In the first approach the relay passively aided the communication between a source and destination by not transmitting and hence reducing interference to the original communication. This technique was referred as facilitation. In the second approach called cooperation, the relay decoded the transmission from the source and retransmitted some “doping” information to the destination. This doping information consisted of the bin index obtained by a random binning of the source message. In the next time instant the original source retransmitted a new message and a bin index for the previous message and the destination used the bin indices obtained from the source and the relay to improve reception (check the validity of this sentence). This scheme however, relies on full decoding at the relay and thus is limited by the rate between the source and the relay. To overcome this difficulty, the authors proposed an observation scheme. In this scheme, the relay just forwarded the observed symbol values to the destination. MRC on an AWGN channel may be regarded as an instance of this scheme.

Sedanaris et. al. [8, 9, 1, 10] studied the idea of user cooperation in the setting of a CDMA wireless system. They proved that with knowledge of channel phase at the transmitter, user cooperation could increase the sum capacity of the users. Laneman [11] also showed that cooperative techniques can only increase the sum rate over non-cooperative transmission only if channel knowledge is available to the transmitter. We, however, are only interested in obtaining diversity advantages through cooperation and do not focus on techniques to improve the sum capacity. Reliability filling, a technique we propose for achieving cooperative diversity,
does not require channel knowledge at the transmitter and thus is not capable of increasing the sum capacity. We shall however show that this technique is capable of providing full diversity benefits promised by cooperative techniques.

Most of the other research [2, 3, 12] have focussed on achieving diversity through opportunistic transmission schemes. The cooperating nodes are just “dumb” relays and the schemes just provide heuristics on when and how the relays should be used or if the relays should be used at all. We summarize sum of the schemes that are relevant to our work below. In [2, 11], Laneman et. al. proposed two different schemes to achieve cooperative diversity. In the first one, called decode-and-forward, the relays first decode the source message and then forward the decoded information bits to the destination. This could be viewed as an instance of repetition coding. This is an instance of Cover’s cooperation scheme with the refinement/doping information being just the information bits. In the second scheme called amplify-and-forward, the relays just amplify their received symbols subject to a power constraint and then forward it to the destination. This could be viewed as a variant of repetition coding on the codeword. Cover’s observation scheme falls in this category if the amplification factor is set to unity i.e., in this case, the relays just forward their received symbols to the destination. In [2, 11], the authors proved two key results about these schemes.

- Though properly designed decode-and-forward schemes can offer full capacity benefits offered by cooperative transmission, full diversity advantages (in the number of collaborating nodes) cannot be offered by these schemes.
- Amplify-and-forward schemes are capable of achieving full diversity advantages promised by cooperative schemes.

The reason for the first observation could be traced back to the fact that decode-and-forward schemes rely on full decoding at the relay and hence are limited by the channel between the source and the relay.

The biggest drawbacks of these schemes are that they are not very amenable to practical implementation. Since, all the above schemes are based on some type of repetition, the communication overhead is very high. The schemes are also not easily scalable to large cooperating groups. Some of these schemes also require some feedback from the destination to the transmitter to perform scheduling for the opportunistic transmission from the relays.

Thus, practical schemes with small communication overhead are required that can achieve the advantages promised by cooperative techniques. These schemes should also have easily implementable scheduling techniques at the MAC layer. The next subsection briefly summarizes our work in the area of cooperative diversity.

2.2 Collaborative Decoding

A suboptimal approach to achieve cooperative diversity was presented by Wong et. al in [5] [6]. In [13], Shea proposed a similar scheme for use with hybrid ARQ. In [14], we investigated various approaches for collaboration with a group of two or more nodes.

Our schemes exploit the fact that most wireless systems use powerful error correction codes and soft-input soft-output (SISO) decoders. In our schemes, each receiver uses a SISO maximum a posteriori (MAP) decoders. An a priori probability and received channel values are the typical inputs to such decoders. At the output, the decoders produce a posteriori probabilities (APPs). If the decoders operate in the log domain (log-MAP decoders), the output consists of log-likelihood ratios (LLRs), $L(X_i|Y=y)$, of the APPs and is referred to as soft
information(output).

\[ L(X_i|Y = y) = \log \frac{\Pr(X_i = +1|Y = y)}{\Pr(X_i = -1|Y = y)} \] (1)

The magnitude of the soft output is called the reliability of the decision and is a measure of the correctness of a decision.

The basic principle of our techniques is as follows. Each node first decodes the transmission from the source. Based on the reliability of the decoded bits, each node exchanges soft information of a certain set of bits with the other nodes. We summarize one of our schemes, the most reliable bit (MRB) exchange scheme, in a little more detail. In the MRB scheme, the nodes decode the source message and then rank the decoded bits according to the reliability. Each node broadcasts the soft output for a fixed percentage of its most reliable bits to the other nodes. The other nodes use this soft information as \textit{a priori} information in their SISO decoders. We called the procedure of soft information exchange among the nodes as \textit{reliability exchange}. This process is then repeated for a fixed number of iterations. The process of iterating between reliability exchange and SISO decoding is referred to as \textit{collaborative decoding} since the nodes assist each other to improve performance. The results in [5, 6, 14] suggest that this technique of exchanging soft information is an efficient way of achieving cooperative diversity.

Note that these techniques could be considered to lie in the realm of the \textit{decode-and-forward} schemes (Cover and Gamal’s cooperation schemes) with the “doping” information consisting of the soft information. However, there is one big difference between our schemes and the \textit{decode-and-forward} schemes mentioned earlier. Our SISO decoders use bit-by-bit MAP decoding like the BCJR [15] algorithm and hence full decoding is not needed to extract useful “doping” information for a small subset of bits. The original \textit{decode-and-forward} schemes relied on full decoding at the relays and hence were limited by this requirement. Thus, our reliability exchange schemes were an improvement over the \textit{decode-and-forward} schemes. It was also shown in [14] that the overhead of these schemes were only a fraction of that of MRC and that simple MAC layer algorithms like sequential access could be used in a practical implementation. These schemes were also easily scalable to small cluster sizes of around ten. Thus, the reliability exchange schemes addressed some of the disadvantages of the original cooperative diversity schemes.

In the next section we describe some drawbacks of the reliability exchange scheme and provide a generalization of the existing schemes called reliability filling. We will show that reliability filling overcomes these disadvantages while still maintaining low overheads and ease of practical implementation. It should also be mentioned that achieving cooperative diversity through coding has also been studied in [4, 16] in the context of an cellular uplink scenario. Their scheme also relies on full decoding at the relay and it is shown in [17] that full diversity is achieved only when the relay decode correctly. The main difference between their ideas and ours is that we use the diversity advantage during iterative decoding whereas the authors of [4] take the traditional approach of exploiting diversity before decoding i.e., MRC is done before decoding begins.

## 3 Reliability Filling

### 3.1 System Model

The system topology that we consider is shown in Figure 1. A distant transmitter broadcasts a packet to a cluster of receiving nodes. Typical scenarios could be military applications in which a battleship broadcasts a message to a platoon of soldiers on the mainland or commercial
applications wherein a base station communicates with a cluster of mobile users. We assume that ARQ is not possible because of the power limitations of the mobile and the distance to the transmitter. Thus, traditional techniques such as code combining [18] or iterative packet combining [19] are not feasible.

The message at the source is packetized and encoded with a code that permits SISO decoding. For the results shown in this paper we use a convolutional code and the max-log-MAP implementation of the BCJR algorithm in the decoder. The encoded codeword is then broadcast to a cluster of receiving nodes over an imperfect channel. We assume that the channels within the cluster to be perfect owing to the proximity of the nodes. This assumption keeps our results general without being tied down to a specific modulation and coding scheme that should be employed in the cluster. Upon receiving the codeword, the nodes attempt to decode the message. If any node in the cluster is successful in decoding the message successfully, we consider the message to be successfully received (the node that decoded the packet successfully can broadcast the information bits to all the nodes in the cluster). This can be thought of as an instance of the decode-and-forward scheme where the node that was successful in decoding relays the information bits to the destination over an error-free channel. A packet error is assumed only when none of the nodes in the cluster are successful in decoding the message. Such an error probability will be referred to as the network packet error probability. Collaborative decoding starts whenever a network packet error is encountered. Collaborative decoding continues until the packet is decoded correctly by at least one node in the network or until a fixed number of iterations elapses. All our previous results in [5, 6, 14] considered the scenario where the initial transmission was over independent non-fading AWGN channels.

When we studied the case of quasi-static (block fading) channels, we found that the performance of reliability exchange schemes was considerably worse than optimal combining (MRC). This is shown in Figure (ref). Even when the soft information is exchanged for all the information bits, the performance is around 2.5 dB worse than MRC. We also found that better performance was obtained if we exchanged the received symbol values instead of soft information. The drawback of exchanging received values is that the overhead increases since information is exchanged for coded bits and not information bits. But in order to obtain full diversity advantages, it seemed necessary to exchange information for the coded bits. Exchanging received symbols is an instance of the amplify-and-forward schemes. Amplify-and-forward schemes are guaranteed to achieve full diversity advantages [11]. MRC which is an instance of this category also exchanges received symbol values. Thus, our observations based on simulations concur with the analytical results in [11].

However, MRC combines this information in an inefficient manner in terms of performance. Because of the use of powerful error correction codes, there are certain bits (trellis sections) about which reliable decisions can be made without the exchange of information.
There are other trellis sections which are a little unreliable but which only need information from a few other nodes to make reliable decisions and there are very unreliable trellis sections that need information from all other nodes. However, MRC combines the same amount of information for all trellis sections, irrespective of the reliability of the original decisions.

Another reason for the failure of our reliability exchange schemes in blockfading channels is that each node in our schemes shares information for the same percentage of bits. In fading channels, some nodes have better channels to the original transmitter and hence have made a greater number of correct bit decisions. Such nodes should share more information with other nodes when compared to the relays with bad channels. These observations led to three key points that should be kept in mind while designing cooperative protocols.

1. In order to obtain full diversity advantages it is necessary to exchange information closest to the RF front end i.e., the received symbol values (soft demodulator outputs).

2. More information needs to be combined for unreliable trellis sections whereas more reliable sections need less information.

3. Nodes with good channels should share more information than nodes with bad channels.

Reliability filling is a technique based on water-filling in the reliability domain that takes into account the above observations.

### 3.2 Water-filling in the Reliability Domain

We first introduce an idealized technique that is similar to MRC, but in which the number of coded symbols combined per trellis section is reduced based on the reliabilities of the decoded bit decisions. Assume that the decoding process is controlled by genie that knows the reliabilities $|L(X_i|y_j)|$ of the information bits at all elements of the distributed array. For each trellis section, the genie picks the number of nodes from which coded symbols should be combined based on the reliability information. So even though reliabilities of the information bits are used to select the nodes for combining, the coded symbols are the quantities being combined as in the amplify-and-forward schemes (MRC).

The genie selects bits from various nodes for combining based on water-filling in the reliability domain. This procedure will be referred to as reliability filling. The selection procedure works as follows. Let

$$S_i = \{S \subset \{1, 2, \ldots, M\} : \sum_{j \in S} |L(X_i|y_j)| \geq T\}$$

Thus, $S_i$ is the set of all possible combinations of nodes in the cluster such that the sum of the reliabilities of bit $i$ at those nodes exceeds a threshold $T$. Let

$$N_i = \min_{S \subseteq S_i} |S'|$$

$N_i$ is the set of minimum number of nodes required such that that the sum of the reliabilities of bit $i$ at those nodes exceeds a threshold $T$. Then the set of nodes $C_i$ for which information will be combined is given by

$$C_i = \begin{cases} \arg\max_{S \subseteq S_i : |S| = N_i} \left\{ \sum_{j \in S} |L(X_i|y_j)| \right\}, & \text{if } S_i \neq \emptyset \\ \{1, 2, \ldots, M\}, & \text{if } S_i = \emptyset \end{cases}$$


Thus, when $S_i = \emptyset$, coded symbols are combined from all nodes in the cluster. When $S_i \neq \emptyset$, coded symbols are combined from the smallest number of nodes such that sum of the reliabilities from those nodes for bit $i$ is maximized. Note that for different trellis sections, a different number of nodes will be involved in the combining process.

Thus, for bits (trellis sections) with low reliabilities, information from more nodes are combined so that the sum of the reliabilities of the bits combined is greater than the threshold. For bits with high reliabilities, information from a few nodes only are combined. For example, if a bit is already decoded with a reliability greater than the threshold, that trellis section does not receive any information from other nodes. It is important to note that even though the reliabilities are used to select the bits to be combined, the received coded symbols are combined as in the amplify-and-forward protocols. The combining overhead is defined as the number of bits broadcast by a relay for successful decoding. This refers to the extra information that should be passed around in the network for successful decoding of the packet. It will shown in section V that reliability-filling offers the capability to tradeoff overhead for performance. It will also be shown that reliability filling can achieve the same performance as MRC and have only around $30 \%-40\%$ of the overhead. The scheme described above however cannot be implemented practically since a genie with knowledge of the reliabilities of all the bits in the system picks out the receivers that should participate in combining for a particular trellis section. In the next section we propose a practical method to implement reliability filling.

4 Proportional Transmission: A Practical Approach to Reliability Filling

The three aspects of cooperative combining schemes mentioned in section III A should be accounted for in practical scheme. Any practical scheme would require a minimum amount of information exchange within the relay to decide the set of trellis sections for which soft demodulator outputs would be broadcast to other nodes. The information allows each node to choose the threshold $T$. This information also constitutes overhead and an efficient protocol should minimize this coordination information.

We use multiple iterations of information exchange to approximate the principles of reliability-filling. In each iteration, information is combined for non-overlapping trellis sections. That is, if a trellis section receives information in one iteration, it cannot receive information in another. During the $i$th iteration, all of the receivers will have a common set of trellis sections for which information has not been transmitted in the previous iterations. Let $B_i$ denote the number of these sections. Let us first consider the coordination protocol that enables each node to determine the threshold $T$. In the $i$th iteration, receiver $j$ broadcasts the mean, $\mu_{i,j}$, of the reliabilities of the $B_i$ bits. The reliabilities have a distribution that is approximately Gaussian and that approximately satisfies the symmetry condition (cf. [20, 21]). Thus,

$$\sigma^2_{i,j} \approx 2\mu_{i,j}$$  \hspace{1cm} (5)

A threshold $T_i$ can be chosen such that the expected value of the total information transmitted in iteration $i$ meets some constraint. The constraint is specified as the proportion $p_i$ of the total number of trellis sections remaining at all receivers at the beginning of iteration $i$. Since the trellis sections which have not received information are the same for each node, the number of trellis sections that have not received any additional information at the beginning of iteration $i$ is given by $MB_i$. This general approach will be referred to as proportional transmission. The
proportional transmission vector for $J$ iterations is given by $\mathbf{P} = [p_1, p_2, \ldots, p_J]$. Then each receiver uses the means transmitted during coordination to determine a common threshold $T_i$ as follows,

$$\sum_{j=1}^{M} Q(\frac{T_i - \mu_{i,j}}{\sqrt{2}\sigma_{i,j}}) = MP_i$$  \hspace{1cm} (6)

Note that the design parameter is the proportional transmission vector. It will shown in the next section that performance can be traded off for overhead by changing the elements of $\mathbf{P}$. Once the threshold is chosen, node $j$ will broadcast the coded symbols for the trellis sections for which $|L(X_i|y_j)| > T_i$. Every other node in the collaborating cluster will then make use of this information.

Note that a small value of $p_i$ will result in a high value of the threshold $T_i$. Thus for a given $p_i$, a node that has a good channel to the transmitter is likely to have more trellis sections with reliabilities greater than the threshold than a node that experiences severe fading. Thus, in each iteration a node with “good” received statistics is likely to transmit more information than a node with a bad channel. Thus property P3 of cooperative protocols (see section III B) is satisfied. We impose the following restrictions in order to satisfy property P2, i.e., fewer copies of received symbols must be combined for reliable trellis sections and more copies must be combined for unreliable ones.

1. Trellis sections selected in one iteration are never selected in later iterations

2. The proportional transmission percentage increases in each iteration i.e.,

$$p_i > p_j \hspace{0.5cm} \forall \ i > j$$

Thus, initially $p_i$ is small ($T_i$ is large) and hence only a few very reliable bits have reliabilities greater than the threshold. Since only a few bits at different nodes satisfy the reliability constraint, only a few copies, if any, are combined for very reliable bits. In later iterations, $p_i$ increases and hence $T_i$ decreases. Thus, it is likely that a bit will satisfy the reliability constraint at more nodes and hence more copies will be combined for these less reliable trellis sections. Note that a trellis section that satisfies the reliability constraint in an earlier iteration also satisfies the reliability constraint in later iterations. However, restriction R1 above ensures that combining in multiple iterations for the same bit is avoided.

We also impose the restriction that the last element of the $\mathbf{P}$ should be 100%. This is done in order to decrease the overhead involved in the scheme. When a node selects and broadcasts information for a subset of bits, the other nodes do not know the bits that have been selected. Thus each transmission should be preceded by set of bit indices that indicates the set of bits for which coded symbols are about to be broadcast. If the block size is $B$ bits then we assume that $\log_2 B$ bits are required to represent each bit index. Thus, when the block size becomes large, the transmission of bit indices could become a potential bottleneck and could overshadow the advantages of proportional transmission. In the last iteration, a lot of bits are selected for transmission and hence the overhead due to bit indices will become very large. By requiring that every receiver will transmit its information for all remaining trellis sections in the last iteration, we avoid transmitting the indices. Since the trellis sections that have not received any information until the last iteration is known to all the receivers, the indices need not be broadcast.

If at the end of any iteration, any node in the cluster recovers the message successfully, the transmission is assumed to be successful and collaborative decoding is terminated. Overhead
savings when compared to MRC can be achieved if the proportional transmission vector is carefully chosen so that proportional transmission does not continue to the last iteration for every packet. Thus, broadcasting the received information for the remaining large number of trellis sections is avoided thereby reducing overhead. Large initial values of the proportional transmission vectors could achieve a low probability of continuing to the last iteration however the overhead involved in transmitting bit indices could overcome the savings achieved by not avoiding the last iteration. Another disadvantage of having large initial values for $P$ is that the threshold is large in the initial iterations. Thus, fewer nodes will be involved in the combining process for a larger number of bits. Thus, performance degradation when compared to optimal combining can be expected (similar diversity advantages are obtained but there is a loss in combining gain). Hence, the proportional transmission vector should be chosen carefully to get good performance in terms of diversity/combining gains as well as overhead. In the next section simulation results are presented to substantiate the above claims about proportional transmission.

5 Simulation Results

For all the results shown in this paper, a rate $\frac{1}{2}$, memory 3 convolutional code with generator polynomials $1 + D^2$ and $1 + D + D^2$ is used at the distant transmitter to encode the message sequence. The information at the transmitter is segmented into 900 bit fragments before feeding it into the encoder. For all the results we restrict the maximum number of iterations of proportional transmission to 3.

References


