INTRODUCTION

Diversity has been an effective technique in combating the wireless channel. Specifically, spatial diversity can be achieved with multiple-input multiple-output (MIMO) techniques by employing multiple antennas at the transmitter and/or receiver side. MIMO techniques have been widely acknowledged and adopted in various wireless standards. However, in future cellular networks or wireless sensor networks, due to the size and hardware implementation limitations, a compact mobile terminal or wireless sensor node may not be able to support multiple antennas. In order to overcome such a limitation, a new form of diversity, called user cooperation diversity or distributed spatial diversity, has been proposed recently [1]. Figure 1 shows a cooperative wireless network. In such a network, mobile stations (MSs) collaborate with each other in communicating with a remote base station (BS). When MS1 sends messages to the BS, MS2 and MS3 can overhear the transmission. Therefore, they can help MS1 forward messages to the BS. Obviously, through cooperative transmission each user’s signals are also forwarded by other users through different paths. Spatial diversity can then be achieved. Cooperative communications can also be applied to achieve extended coverage with reduced transmission power. Due to these benefits, the concepts have been proposed in some new wireless standards, such as WiMAX 802.16m.

The concept of cooperative communications can actually be traced back to the early work of Cover and Gamal on achievable capacity of a relay network in 1979. It was rediscovered recently in relation to great potential applications in cellular and wireless sensor networks [1, 2]. The distributed nature of wireless networks provides a unique opportunity for cooperation and distributed signal processing. Design of efficient cooperative protocols and distributed signal processing techniques has been an important issue in implementing cooperative communications in wireless networks. Therefore, recent research on cooperative wireless networks has focused on designing relaying protocols, signaling, and distributed coding. In particular, design of efficient relaying protocols and distributed coding schemes has attracted significant attention, and a number of novel relay protocols [1–3] and distributed coding schemes [4–15] have been developed in the past several years. Capacity-approaching performance has been achieved by some elegantly designed distributed coding schemes.

Given that the research in this area started recently, many open issues in the design and implementation of cooperative communications still have not been fully addressed. The aim of this article is to present a tutorial survey of recent development in distributed coding design in cooperative wireless networks.

COOPERATIVE COMMUNICATIONS AND RELAY NETWORKS

When we talk about cooperative networks and relay networks, they have the following distinctions. In cooperative networks, each node
An example of a cooperative communication system.

Figure 1. An example of a cooperative communication system.

Acting as both a source and a relay node. That is, each node not only transmits its own information but also helps other nodes to transmit signals. In relay networks, the relay nodes are explicitly built nodes only for the purpose of relaying and forwarding information. They do not have their own information to transmit. Despite these differences, as far as signal processing at relays is concerned, they are almost the same. Therefore, in this article we use these two terms interchangeably. In this section, we first describe a relay network model. As an example, let us consider a relay network shown in Fig. 1, consisting of one source MS1, two relays MS2 and MS3, and one destination BS. Due to the practical constraints in hardware implementation, it is usually assumed that each node cannot transmit and receive at the same time. The overall transmission can be divided into two phases. The source node first broadcasts its messages to both relays and destination. Upon receiving signals from the source, each relay processes the received signals and then forwards them to the destination. There are various strategies that can be applied at relays to process the received signals. In the simplest case, relays can just simply amplify the received signal. Relays can also reconstruct the original messages from the received signals with or without decoding operations. In a more complex protocol, relay can calculate and transmit additional information related to source messages, such as incremental redundancy, in order to improve the system performance. We will discuss various relaying protocols in the next section.

When the relays transmit in the second phase, the source could either transmit at the same time or remain silent. Furthermore, the source and relays can transmit through orthogonal channels or nonorthogonal channels. In orthogonal transmissions, all relays transmit to the destination through orthogonal channels, such that the transmitted signals from each relay can be separated at the destination, without any interference from other relays. For non-orthogonal transmissions, relays transmit to the destination at the same time and same frequency. Therefore, the received signal at the destination is the superposition of signals transmitted from the source and relays. A silent source in the second phase and orthogonal transmissions from relays simplify the receiver structures as no multiple access interference suppression is needed. However, this will reduce the channel capacity. It is up to a practical system designer to determine which transmission mode should be used in the relay networks.

At the destination, depending on the transmission mode and relay protocols, various processing methods will be used to recover the source messages based on the received packets transmitted from the source and relays, such as linear combinations and iterative decoding. Since signals received at the destination arrive via different paths from the source and various relays, spatial diversity can be achieved at the destination.

**Relay Protocols**

Based on operations at the relays, there are several commonly used relay protocols. In this section, we give an overview of these protocols. In the next section, we discuss how these protocols are integrated in various distributed coding schemes.

**Amplify and Forward**

Amplify and Forward (AAF) is one of the simplest relay protocols [2]. In AAF, upon receiving signals from the source, each relay just simply forwards to the destination a scaled version of the received signals, including both information and noise. By properly combining received signals from the source and relays, the destination node makes a final decision. Since the destination receives multiple copies of signals transmitted from the source and relays through multiple independent paths, spatial diversity can always be achieved by the AAF protocol at high signal-to-noise ratios (SNRs). Obviously, the major drawback of AAF protocols is noise amplification at the relays. The outage probability and end-to-end bit error rate (BER) performance of AAF have been widely investigated [2]. Given a fixed transmission power at the source, relay, and destination, it is shown that the performance of AAF depends on the position of the relay relative to the source and destination. When the relay is positioned midway between the source and destination, AAF achieves its optimum performance and worsens as the relay moves closer or further to the source [15].

**Demodulation and Forward**

To eliminate the effect of noise amplification, several relay protocols have been proposed. Demodulation and Forward (DemAF) is one of the simple solutions among them. In DemAF, the relay simply demodulates the received signals, with no decoding, and remodulates to reconstruct the symbols transmitted by the source. This process can simply remove the noise components residing in the received signals at relay. In [3] the performance of DemAF has been analyzed when taking into account the
corresponding soft information instead of making in DAF is to calculate and forward the An alternative approach to avoiding error propagation is to switch to the DAF protocol. Simulation results show considerable improvement compared to the AAF protocol. Otherwise, the relay uses an adaptive algorithm to decode the received signals and re-encodes them before forwarding to the destination.

When the channel quality in the link between the source and relay is good, the process of decoding and re-encoding provides more powerful error correcting capabilities than DemAF. Thus, the method can considerably outperform both AAF and DemAF. However, when the link from the source to the relay suffers from deep fading, decoding errors may occur at the relay. In this case, if the relay re-encodes these incorrect bits, error propagation will occur and lead to even worse performance.

**Adaptive Relay Protocol**

So far, we have seen that different protocols mentioned above all have their advantages as well as disadvantages. Intuitively, we would like to ask one question: are there any possible protocols, which can not only effectively mitigate the noise amplification, but also avoid error propagation? Adaptive Relay Protocol (ARP) is one of the protocols developed to meet this need [4]. ARP has advantages of both AAF and DAF and minimizes their negative effects at the same time. In ARP, each relay adaptively selects the AAF or DAF protocol based on whether its decoding result is correct or not. All the relays that fail to decode correctly use the AAF protocol to amplify the received signals and forward them to the destination. On the other hand, all the relays that can successfully decode the received signals use the DAF protocol. The signals received at the destination, forwarded from all relays by using either AAF or DAF protocol, are combined into one signal to recover the source information. It has been shown in [4] that ARP considerably outperforms the AAF scheme and simultaneously avoids error propagation due to the imperfect decoding at relays in a DAF protocol. Thus, it outperforms both AAF and DAF protocols. The performance gain grows as the number of relays increases, and it approaches the perfect DAF scheme at high SNRs. Recently, some threshold-based ARP protocols have also been developed [5]. In such protocols the relay adaptively selects the relay protocols by comparing the received SNR to a threshold. If the SNR is lower than the threshold, the relay uses the AAF protocol. Otherwise, the relay switches to the DAF protocol. Simulation results also show considerable improvement compared to the DAF protocol.

**Soft Information Relaying**

An alternative approach to avoiding error propagation in DAF is to calculate and forward the corresponding soft information instead of making a decision on the transmitted information symbols at the relay [6, 7]. Soft information relaying provides additional information to the destination decoder to make decisions, instead of making premature decisions at the relay decoder. It has been shown that Soft Information Relaying (SIR) considerably outperforms AAF and DAF protocols. Soft information is an analog signal. To transmit such analog signals, compression and/or quantization should be performed at the relays to covert analog soft information into digital signals.

**Distributed Coding Structure**

The performance of relayed transmission can be improved if the source and relays cooperate to perform joint signal design and coding. The major difference between distributed coding and conventional channel coding schemes is that in distributed coding, the overall code-word is constructed in a distributed manner. That is, different parts of the codeword in distributed coding are transmitted by different nodes through independent wireless links. This creates additional degrees of freedom, but also poses new challenges in code construction. Although we can directly apply the concepts of conventional channel coding to construct distributed coding in wireless relay networks, some practical issues in designing these distributed coding schemes have to be taken into account, such as decoding errors at relays, channel variations in different parts of the network, rate and power allocations at the source and relays. In this section we present an overview of various distributed coding structures that have been successively developed over the past several years for wireless relay networks.

**Distributed Space-Time Coding**

Let us consider a wireless relay network in which the source and relays cooperatively communicate with a common destination. This cooperative transmission among the source and relays forms a virtual antenna array. Therefore, conventional space-time coding schemes can be applied to relay networks for achieving the cooperative diversity and coding gain. Two types of distributed space-time coding (D-STC) schemes have been developed, including distributed space-time block codes (DSTBCs) and distributed space-time trellis codes (DSTTCs).

Several DSTBC schemes have been proposed. A simple DSTBC scheme was proposed by Laneman based on orthogonal STBCs. In such a scheme different relays transmit different columns of the STBC code matrix, and at the destination a DSTBC codeword is formed. Since the orthogonal STBCs do not always exist for any number of antennas, such a DSTBC scheme has certain requirements on the number of active relays. The design of such a DSTBC becomes especially difficult for a large network with a large number of relays. To solve this problem, several solutions have been proposed. In [8] Yiu et al. proposed a DSTBC scheme, which selects a subset of nodes for transmission, and each active node transmits a linear transformation of DSTBC codewords. The transformation is
In DSTTC the convolutional codes are used as the constituent codes at the source and relay nodes. In order to further improve the system performance, some Distributed Low Density Parity Check (D-LDPC) coding schemes have been developed recently [10]. In D-LDPC the constituent codes at the source and relay are LDPC codes. An LDPC code with a predetermined code rate is first generated. The whole codeword consists of three parts. The first one is transmitted by the source in the first time slot. The codeword is chosen from an LDPC codebook, $C_{SD1}$. In the second time slot, two parts are chosen from two other LDPC codebooks, $C_{SR1}$ and $C_{SR2}$. They are transmitted by the relay and source in the second time slot. The bits chosen from $C_{SR1}$ and $C_{SR2}$ are transmitted by the source and relay to form another LDPC code, $C_{SD2}$. As a result, the design of D-LDPC codes requires joint optimization of the code profiles of $C_{SR1}$ and $C_{SD1}$ [10]. The density evolution (DE) has been widely used for the optimization of LDPC codes. DE can accurately track the evolution of the probability densities in a belief propagation decoding algorithm. For the conventional LDPC codes, it has been shown that good check node distributions (CNDs) are concentrated. That is, all parity check nodes should have nearly equal degrees. This property can be used to simplify DE implementation by selecting several concentrated CNDs and searching the best variable node distributions for each CND. Unfortunately, for D-LDPC such an assumption is not valid, because the rates of $C_{SR1}$ and $C_{SR2}$ are very different [10]. This creates a significant challenge for DE implementation. Some simplified DE algorithms have been developed in [10] by using Gaussian approximation of the DE distribution and putting some constraints on the CNDs of $C_{SR2}$ and $C_{SD1}$ to reduce the search space. The search for good code profiles can be made using linear programming, and near optimum codes can be found by selecting the code with the optimum convergence threshold. It has been shown that through proper code design, an D-LDPC scheme over wireless relay channels can perform very close to the theoretical limit [10].

**Distributed Turbo Coding**

Similarly, by following a turbo coding structure, another capacity approaching structure, referred to as distributed turbo coding (DTC), has been proposed [11, 12]. Figure 3 shows the block diagram of a DTC system. In a DTC scheme, the relay employs the DAF protocol. The source broadcasts the coded signals to both the destination and relay. The relay calculates another codeword $X_r$ and generates another codeword $X_s$ to the destination. The total received signals at the destination in these two time slots form a DSTTC codeword $X$, given by

$$X = \begin{bmatrix} X_{s1} & X_{s2} \\ 0 & X_{r1} \end{bmatrix}$$

We should note that various relay protocols can be used to construct DSTTCs, such as AAF, DemAF, or DAF. When DAF or DemAF is used in DSTTCs, detection or decoding errors have to be taken into account in the code design. Most papers usually assume that the relay can decode correctly, and thus the construction of DSTTC can be done in a similar way as in conventional STTC schemes. In order to achieve optimum performance in DSTTCs, encoders $A_1$, $B$, and $A_2$ should be jointly optimized. It has been shown that the code construction rules of DSTTC still follow the rank and determinant design criteria of conventional STTC [9], but the codeword difference matrix between two codewords $X$ and $\hat{X}$ here is an extended matrix, given by

$$B(X, \hat{X}) = X - \hat{X} = \begin{bmatrix} X_{s1} - \hat{X}_{s1} & X_{s2} - \hat{X}_{s2} \\ 0 & X_{r1} - \hat{X}_{r1} \end{bmatrix}$$

In [9] a practical DSTTC scheme, which takes into consideration detection errors at relays in the design of DSTTCs, has been proposed. In this scheme an equivalent link, representing the source-relay-destination path, has been proposed by using the equivalent SNR of the link model to take into account the detection errors at the relay [3]. It has been shown that the optimum code design still follows the rank and determinant criteria. However, the codeword difference matrix has to be modified to take into consideration the detection errors at relays [9].

**Distributed Low Density Parity Check Code**

Distributed space-time trellis coding scheme.
information and relay. The relay decodes the received signals and interleaves them prior to re-encoding. The signals received at the destination consist of a coded signal transmitted from the source and coded interleaved information transmitted from the relay. These two signals form a distributed turbo code. It has been shown that such a coding strategy performs close to the theoretical outage probability bound of a relay channel. However, in such a coding scheme, it is usually assumed that the relay performs error-free decoding, which we refer to as the perfect DTC. Design of DTC schemes when imperfect decoding occurs at the relay has become a practical and important issue. Similar to the methods for overcoming error propagation in relay protocols, there are also two possible ways to design the DTC to avoid decoding errors at relays.

The first approach consists of applying ARP to DTC, which results in a DTC with ARP (DTC-ARP) scheme [14]. Figure 4 shows its basic principle. Similar to the ARP, in the DTC-ARP all relays are separated into two relay groups, referred to as AAF and DAF relay groups. All the relays that decode correctly are included in the DAF group, and the remaining relays, which fail to decode correctly, are included in the AAF group. All the relays in the AAF relay group amplify the received signals and forward them to the destination, while all the relays in the DAF relay group decode the received signals, interleave the decoded symbols, re-encode, and forward them to the destination. All signals forwarded from the AAF relay group are combined at the destination into one signal, and those from the DAF relay group are combined into another signal. The overall destination codeword consists of the combined coded information symbols transmitted from the AAF relay group and combined coded interleaved information symbols transmitted from the DAF relay group. These two signals form a distributed turbo codeword. An iterative decoding is then performed between these two signals. Simulation results have shown that the DTC-ARP performs very close to the perfect DTC at high SNRs.

Alternatively, we can also apply the SIR protocol to the DTC. However, the relays only have soft estimates of information bits when relays are unable to decode correctly. Therefore, the major challenge here is to calculate soft information for the parity symbols of interleaved information symbols at relays. Although we cannot directly encode soft information, we can actually first calculate the a posteriori probabilities (APPs) of information symbols and use them to calculate the APPs of parity symbols for the interleaved information symbols. In [7] a probability inference method has been developed that traces the trellis of the code to calculate the probability of each symbol at time $k$ based on the probabilities of all symbols at time $k-1$. After obtaining the probabilities of the parity symbols of the interleaved information symbols, the corresponding soft estimates can then be calculated as the mean of modulated symbols. Figure 5 shows the block diagram of such a distributed coding scheme with soft information relaying.

Figure 6 compares the FER performance of various DTC schemes. To make a fair comparison with the DTC-SIR, in the DTC-ARP scheme, when the relays decode the received signals correctly, they forward only the parity symbols of the interleaved information symbols; when the relays fail to decode correctly, they forward the amplified version of the information symbols. Let $\gamma_{sr}$, $\gamma_{rd}$, and $\gamma_{sr}$ represent the SNR in the link from the source to the relay, from the source to the destination, and from the relay to the destination. We assume that $\gamma_{sr} = \gamma_{rd} = \gamma$, $\gamma_{rd} = \gamma_{gap}$, and $SNR_{Gap} = 10\log(\gamma_{gap})$. From the figure, we can see that the performance of DTC-SIR is about 2–3 dB better than that of DTC-ARP. As $SNR_{Gap}$ increases, the performance of the DTC-SIR very closely approaches the perfect DTC scheme, for which we assume the relay can always decode the received signals correctly.

**DISTRIBUTED NETWORK AND CHANNEL CODING**

The distributed coding schemes discussed above are developed for the unicast relay networks in which the messages are sent from a single source to a single destination through multihop relays. In some application scenarios, such as uplinks in cellular networks, multiple MSs communicate with a common BS through multihop relay stations. Network coding has been shown to be an effective technique to increase the spectral efficiency of such a network. Recently, the concept of network coding has been applied to such relay...
networks, where multiple source nodes communicate with a common destination through multi-hop relays. This scheme is referred to as adaptive network coded cooperation (ANCC) [13]. In ANCC each source node broadcasts its information symbols to both relays and destination. Each relay then decodes the received information message from all source nodes. Among the successfully decoded information messages, each relay randomly selects a small number of messages and performs a binary summation of these selected messages to generate a parity check message. Through the above operations at the relays, a graph code, such as LDPC or low density generator matrix (LDGM) network code, is finally formed at the destination. A belief propagation decoding algorithm can then be used to decode the graph codes formed at the destination. Results have shown that such a distributed network channel coding scheme significantly outperforms repetition-based schemes and performs on par with D-STC schemes.

**CONCLUSIONS AND OPEN PROBLEMS**

In this article we have given an overview of recent research achievement in distributed coding technology. Through proper design, distributed coding can achieve both spatial diversity and a significant coding gain. Some efficient distributed coding schemes have been proposed in the past several years, but there are still many issues in both theory and practical implementation that have not been addressed.

Most existing distributed coding schemes are developed based on conventional channel coding schemes, such as STC, turbo coding, and LDPC coding. Furthermore, most distributed coding schemes rely on some assumptions, such as error free decoding at relays. Some initial work on modeling detection errors in DemAF has been done by Wang et al. in [3]. This model has been applied to construct DSTTCs in [9]. It has been shown that detection errors do have some effects in constructing practical distributed coding. However, there is still no accurate mathematical representation to model the decoding errors. As a result, no paper has actually formulated the accurate design criteria for distributed coding with DAF protocols when decoding errors occur at relays.

Additionally, most existing distributed coding schemes mainly concentrate on fixed code rates and power allocations. Design of a distributed coding scheme that can adaptively allocate the rate and power and distribute the code bits among the source and relays, as well as adaptively select the relay protocols, has the potential to increase overall network throughput, reduce network power consumption, and improve reliability. Furthermore, not much has been done on the design of distributed coding schemes for MIMO relay networks. MIMO relay networks are much more complex than a single MIMO system or a single-input single-output relay network. Multiple antennas at each node provide an additional dimension for designing distributed coding. Many open issues reside in this area, such as joint precoding design at source and relays for spatial multiplexing transmission, interference cancellation at relays and destination, design of optimum vector relay protocols, and coding structures and code design criteria for MIMO relay networks.

There are also other practical issues that should be taken into account in implementing cooperative communications in wireless networks, such as synchronization among relays, signaling design, channel estimations, user coordination, and resource management. Some initial work has been done to solve some individual problems. For example, some asynchronous relaying schemes have been proposed to avoid the synchronization problems; differential modulation schemes were developed to overcome channel estimation issues; partner selection schemes and adaptive resource (power/frequency/time slots) allocations have been proposed for efficient user coordination and efficient resource management. However, different schemes rely on different system assumptions and were developed under different system models. It is important to develop a unified system model and approach that can solve all these practical issues in practice. In practical cooperative...
systems, there are also some security issues which should be considered, like attacks from malicious relays, denial of service from selfish relays for the sake of saving its energy, etc. How to ensure a secure transmission is also an important issue in cooperative wireless networks.

REFERENCES


BIOGRAPHY

YONGHUI LI (M’04, SM’09) (lyh@ee.usyd.edu.au) received his Ph.D. degree in electronic engineering in November 2002 from Beijing University of Aeronautics and Astronautics. From 1999 to 2003 he was affiliated with Linkair Communication Inc, where he held a position of project manager with responsibility for the design of physical layer solutions for LAS-CDMA system. Since 2003 he has been with the Telecommunication Laboratory, University of Sydney, Australia. He is now a senior lecturer in the School of Electrical and Information Engineering, University of Sydney. He currently also holds the Australian Queen Elizabeth II fellowship. His current research interests are in the area of wireless communications, with a particular focus on MIMO, cooperative communications, coding techniques, and cognitive radios. He holds a number of patents granted and pending in these fields. He was an Associate Editor of EURASIP Journal on Wireless Communications and Networking, 2006–2008, and an Editor of the Journal of Networks. He also served as Editor for the special issue on Advances in Error Control Coding Techniques in EURASIP Journal on Wireless Communications and Networking. He has also been involved in the technical committees of several international conferences, such as ICC, GLOBECOM, VTC, and PIMRC.