RESEARCH ARTICLE

Efficient multihop transmission scheme for error-free relay forwarding in cooperative networks

Ashish James1*, A. S. Madhukumar1 and Fumiyuki Adachi2

1 CeMNeT, School of Computer Engineering, Nanyang Technological University, Singapore 639798
2 Graduate School of Engineering, Tohoku University, Sendai, 980-8579 Japan

ABSTRACT

Multihop cooperative communication is emerging as a key concept to extend the coverage area of the network and potentially increase the capacity. The spectral efficiency of such networks can be improved by adapting the transmission to time-varying channel conditions, referred to as incremental relaying. Although such incremental relaying concepts are progressively being studied, many challenges, such as erroneous transmissions by intermediate nodes and end-to-end delay of the network, limit its practical use due to lack of an efficient implementation. This paper proposes an efficient multihop incremental relaying technique. In this method, erroneous relay forwarding is mitigated, and the overhead for coordination among nodes is reduced by exploiting the implicit feedback channel available due to the broadcast nature of wireless transmissions. The proposed scheme fully leverages the benefit of overhearing and eliminates the additional feedback slots required for validation. Further, it ensures reliable forwarding of information, which optimizes the throughput of multihop networks. Thorough analysis of the proposed scheme is performed under different deployment environments, and the theoretical analyses presented in this paper are supported with results from extensive simulation studies. Copyright © 2013 John Wiley & Sons, Ltd.

KEYWORDS

cooperative communication; incremental relaying; RCPC; multihop

*Correspondence
Ashish James, CeMNeT, School of Computer Engineering, Nanyang Technological University, Singapore 639798.
E-mail: ajames@ntu.edu.sg

1. INTRODUCTION

As future wireless communication systems are expected to deliver transmission rate in the order of several hundreds of Mbps or even Gbps, research efforts have been driven toward achieving higher spectral utilization. The use of multiple antennas at both the transmitter and receiver had been at the forefront of such drive, which results in the formation of multiple-input multiple-output (MIMO) systems. However, various limitations put constraints on the implementability of MIMO architecture on mobile devices. The small physical dimension of mobile devices does not provide adequate spatial separation between multiple antennas to ensure uncorrelated fading and hence, the spatial diversity. In addition, a mobile device with limited battery capacity might not be able to accommodate the power drainage from multiple RF chains. Such practical limitations have prompted a recent flock of research attention toward another space–time technique known as cooperative communication.

In cooperative communication, a node (mobile device) within a network will assist every other node (within the same network) in their information delivery. Such techniques utilize the broadcast (BC) nature of wireless signals by overhearing the source signal for the intended destination at the neighboring nodes. These nodes known as relays process the overheard signals and transmit toward the destination. By introducing the possibility of such cooperation, each device can be viewed as having an array of virtual antennas at its disposal. The virtual antenna array provides spatial diversity by transmitting over statistically independent relay paths. This advantage could in turn be exploited to enhance network performances such as reduced link uncertainties caused by fading and extended network coverage [1,2]. In its most simplified form, cooperative networks are composed of three types of nodes, namely, source $S$, relay $R$, and destination $D$, with transmission of every packet being carried out by source and relay over two slots.
Recently, owing to its significant advantages, cooperative communication has emerged as a strong candidate for underlying technology in future wireless applications, including 4G cellular networks, wireless sensor networks (IEEE 802.15.4), and fixed broadband wireless systems (WiMax, IEEE 802.16j). The reliability of cooperative networks can be alleviated by the use of forward error correction techniques [3]. Even with such great attributes, the overall performance of cooperative networks is dependent on detection reliability at the relay nodes. Hence, the performance of such networks can be further enhanced by restricting the relaying to those cases where it is beneficial [4]. Such a restricted cooperative relaying scheme is proposed in [5]. The idea is to allow a relay node to choose between transmitting re-encoded information bits (if decoding is successful) and soft reliability information (if decoding fails). These basic cooperative transmission schemes suffer from one common shortcoming. The transmission rate of these schemes is rigidly fixed and results in non-optimal bandwidth efficiency. Selection of a transmission rate that is too high might lead to unstable performance as packet delivery will be overly sensitive to channel fluctuations. On the other hand, selection of a transmission rate that is too low might lead to stable performance but low bandwidth efficiency during good channel conditions.

Coded cooperative schemes with adaptive code rate have been proposed in this context. Such schemes are generally leveraging on automatic repeat request (ARQ) transmission mechanism to deliver the variable code rate. One straightforward implementation of ARQ into a cooperative transmission scheme is to permit a relay to decide whether it should proceed to retransmit the received information. The relay enters into cooperation mode only upon successful decoding. In this way, the effective bandwidth utilization is either one symbol per channel use during non-cooperation or one half symbol per channel use when relay is cooperating. Bandwidth efficiency can be further enhanced by increasing code rate variability or essentially, the number of code rates that are supported by the system. This can be achieved by combining the concept of ARQ with incremental redundancy. In incremental redundancy scheme, the coding algorithm used is one that allows partial decoding. This means that, though such algorithms produce large number of bits during the encoding operation, successful decoding is possible even with a subset of those bits under reliable channel conditions. Examples of coding algorithms that possess such properties are rate-compatible punctured convolutional (RCPC) codes and punctured low density parity check codes [6,7]. With such codes, the initial transmission is at the highest code rate defined for the system, which consists of the minimum number of bits that are sufficient for successful decoding in a reliable channel. The code rate variability can be applied either only to the information-designated (source) node or to both the relay and the information-designated node. By permitting code rate variability only at the source node results in a simpler option.

In incremental redundancy-based cooperative relaying systems, decoding failure at the destination triggers the generation and transmission of ARQ packets by either the source or relay nodes [8,9]. The relay node takes over the transmissions by transmitting the additional coded bits upon reception of an ARQ packet. More coded bits at the destination decrease the effective code rate, which in turn improves the system performance. On one hand, such scheme offers low processing complexity at the relay (as a relay is only required to decode only once), and there is minimal coordination among relay and source nodes. On the other hand, this option is suitable only when decoding at the relay is reliable and otherwise results in error propagation within the system. Most of the present schemes assume perfect decoding to maintain the diversity of the system [1,2]. Imperfect decoding however degrades the system performance, where the severity of this degradation is dependent on decoding reliability at the relay nodes.

Several schemes have been devised to tackle this problem but at the cost of increased system complexity and compromising on the bandwidth efficiency [10,11]. In one of these methods, the relays compute the log-likelihood ratio (LLR) of the received bits and forward it to the destination only when the LLR of a bit is above a predetermined threshold. In an alternative method, a relay node will respond to each ARQ packet by broadcasting a ready-to-relay (RTR) packet upon successful decoding (ensured by checking the received information for errors) and remains silent otherwise. Source continues to transmit bits in an incremental manner until it is taken over by the relay. In addition to the increased complexity, such schemes incur significant overhead for coordination purposes as a slot for RTR transmission needs to be reserved at the end of every ARQ slot.

An attempt to reduce such coordination overheads has been studied in [12,13]. In that attempt, the BC nature of wireless transmission is exploited as an implicit feedback channel between the relay and the source nodes. The scheme proposed offers two notable advantages. First, it allows the source node to validate the accuracy of decoding at the relay without requiring an additional feedback slot. Second, it eliminates the need for validation checking at the relay and hence lowers the computational complexity at the relay. The scheme presented in [12,13] has been analyzed in the context of two hops (or one relay stage) network. A cooperative system with two hops transmission has been the subject of intense research, and marriage between theory and practice is increasingly apparent on the horizon. However, the overall transmission range of such networks is constrained by the limited transmission power of individual wireless nodes. The energy expenditure and transmission delay for such networks also increase with the distance between nodes. Such limitations have prompted the use of multihop transmission as a potential technique for future wireless networks to provide high data rate with wider coverage and efficient spectrum utilization. IEEE 802.16j standard has been proposed by the standardization bodies for supporting multihop relay communications [14].
Multihop transmission breaks the path between source and destination nodes into several hops via various intermediate relay nodes. In such networks, messages are relayed over sequential point-to-point links where the intermediate nodes relay the received information toward the final destination. By having much less path loss, multihop transmission reduces the interference levels, terminal radiation, and power consumption of nodes in the network [15,16]. The performance of a cooperative relay network can be enhanced further by allowing multihop transmission. A cooperative system with multihop transmission has only been receiving much attention lately, and its practical adoption could still be years away. The task of ensuring accurate relay forwarding becomes more complex with increasing number of nodes in such multihop systems. The potential delay and the overheads for coordination purposes also increase with the number of nodes in the network. In this context, a system of cooperative relay network with multihop transmission is proposed in this paper, and the method is referred as multi-node incremental code redundancy (MNICRIF). The objective of this paper is to investigate the fundamental performance and trade-offs associated with such a scheme and to determine the suitability of implicit feedback channel in multihop transmission.

The rest of this paper is organized as follows. The system model employed for analyzing the proposed scheme is discussed in Section 2. Detailed discussion on the proposed scheme, MNICRIF is presented in Section 3, and its performance is analyzed in Section 4. Numerical results demonstrating the performance and trade-offs associated with using an implicit feedback channel in multihop systems are presented in Section 5. Finally, concluding remarks and future extensions of the present work are presented in Section 6.

2. SYSTEM MODEL

In this paper, a wireless network as depicted in Figure 1 is considered, where a continuous stream of information bits generated by source $S$ is communicated to the destination $D$ by cooperating with the $m \times n$ intermediate relays ($R_{ij}$). All nodes are equipped with a single omnidirectional antenna and are constrained by half-duplex transmission. The present paper uses RCPC codes with memory $M$ and puncturing period $P$ (the number of information bits used in a single puncturing operation) as the incremental redundancy-based forward error correction code. The effective code rate of the proposed system can be computed by considering the number of redundant bits required for successful decoding. For example, let $L_{\text{min}} (L_{\text{max}})$ be the minimum (maximum) number of additional coded bits that are transmitted. This translates to maximum (minimum) coding rate supported by the system as

$$\frac{P}{(P + L_{\min})} \left( \frac{P}{(P + L_{\max})} \right).$$

2.1. Transceiver & channel model

The continuous stream of information generated by $S$ is first grouped into blocks of size $M \times 1$. At any instant $j$, the information block fed to the RCPC encoder can be represented as $u(j) = [u(jM) \cdots u(jM + M - 1)]^T$. Without loss of generality, $j$ will be omitted from subsequent discussions. The RCPC encoder first divides each data block of $M$ symbols into $\frac{M}{P}$ smaller sub-blocks, $u_p = [u(pP) \cdots u(pP + P - 1)]^T$, each of size $P$, where $p$ is used to index each sub-block. Each $P \times 1$ block is used by the RCPC encoder to produce $(P + L_{\max})$ bits. This means that for every $M$ input bits, $\frac{M}{P}(P + L_{\max})$ bits are produced. The output of the RCPC encoder is given as

$$E = \begin{bmatrix}
e_0(0) & \cdots & e_0 \left( \frac{M}{P} - 1 \right) \\
\vdots & \ddots & \vdots \\
e_{p + L_{\max} - 1}(0) & \cdots & e_{p + L_{\max} - 1} \left( \frac{M}{P} - 1 \right) \\
\end{bmatrix} = \begin{bmatrix}
e_0 & \cdots & e \left( \frac{M}{P} - 1 \right)
\end{bmatrix}^T,$$

(1)

Figure 1. System configuration with source node ($S$), destination node ($D$), and $m \times n$ intermediate relay nodes ($R_{ij}$).
At this point, the code rate corresponding to the initial output of the RCPC encoder is $\frac{P}{P+L_{\text{max}}}$. Encoded sequences with different rates can be obtained by multiplying the encoded output with rate-compatible puncturing matrices. The set of puncturing matrices considered in this paper is one that produces equally good codes during the entire transmission process, generated according to [6]. The maximum number of slots $(F)$ is dependent on the minimum code rate supported by the system $(\frac{P}{P+L_{\text{max}}})$. $A_0$, $A_1$ to $A_F$ are considered as the puncturing matrices during slots $0$ to $F$, respectively, with $S$ transmitting during slot $0$. The respective size of the puncturing matrices used for the corresponding slots are defined as $\|A_0\| = (P + L_0) \times (P + L_F)$, $\|A_1\| = (L_1 - L_0) \times (P + L_F)$, $\ldots$, $\|A_F\| = (L_F - L_{F-1}) \times (P + L_F)$, where $L_0, L_1$ to $L_F$ corresponds to the coded bits received during slots $0$ to $F$, respectively. The encoded vectors to be transmitted during each slot are obtained by performing the puncturing operation that consists of pre-multiplying the respective matrices with the initial encoded matrix of primary rate $\frac{P}{P+L_{\text{max}}}$ (hereinafter referred to as the primary code matrix). The outputs of the puncturing operation for different slots can be given as

<table>
<thead>
<tr>
<th>Slot</th>
<th>$E_0$</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$E_F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$A_0 E = \left[ e_{0,0} \cdots e_{0,P} \right]^T$</td>
<td>$A_1 E = \left[ e_{1,0} \cdots e_{1,L-1} \right]^T$</td>
<td>$A_{F-1} E = \left[ e_{F,0} \cdots e_{F,L-1} \right]^T$</td>
<td>$A_F E = \left[ e_{F,0} \cdots e_{F,L-1} \right]^T$</td>
</tr>
</tbody>
</table>

Transmission of the punctured matrices are carried out row-by-row, and $e_{i,p}$ corresponds to the $p$th vector transmitted during slot-$t$ ($t \in [0, \ldots, F]$), $p \in [0, P + L_0 - 1]$ for the initial encoding, $p \in [0, L_1 - L_0 - 1]$ for the intermediate slots, and $p \in [0, L_F - L_{F-1} - 1]$ for the final encoding.

All nodes are assumed to operate in BC mode, and the coded bits are transmitted in several slots till the decoding is successful at $D$. The transmission is initiated by $S$, and all the intermediate nodes within its transmission range receive the transmitted information. The corresponding received signal for the $p$th encoded vector transmitted during slot-$t$ over a Rayleigh fading channel can be expressed as

$$y_q(t, p) = h_{vq}(t, p)x_v(t, p) + n_q(t, p)$$

where $v \in \{S, R_{ij}\}$ and $q \in \{R_{ij}, D\}$, $x_v$ is the transmitted codeword matrix, $h_{vq}$ is the channel transfer matrix between transmitting node $v$ and receiving node $q$, which is selected from a complex Gaussian distribution corresponding to Rayleigh fading amplitudes, $n_q$ is the noise matrix, which corresponds to additive white Gaussian noise with zero mean and variance $N_0$, and $y_q$ is the received signal matrix. Each node is assumed to transmit under equal power constraint, and hence $E[|x_v|^2] = \epsilon \forall v$, where $\epsilon$ is the average energy per symbol during transmission. The distribution of the channel matrix $h_{vq}$ has a variance of $\rho_{vq}$, which captures the pathloss between nodes $v$ and $q$ as well as shadow fading. The variance $\rho_{vq}$ is dependent on the pathloss model and is proportionally related to the distance $(d_{vq})$ between nodes $v$ and $q$ as $\rho_{vq} \propto 1/d_{vq}^\alpha$, where $\alpha$ is the pathloss exponent. With sufficient inter-leaving, samples of $h_{vq}$ are considered to be independent. Equalization is performed on the received signal in (3) using conventional channel equalization methods such as zero forcing. The equalizer output at node $q$ is denoted as $z_q(t, p)$. After equalization, de-puncturing operation is performed by stacking the received equalized vectors one after another and pre-multiplying the stacked vectors with the de-puncturing matrix. De-puncturing matrix for slot-$t$ is obtained by concatenating all the received puncturing matrices till slot-$t$. The output of the de-puncturing operation at node $q$ is given as

$$\hat{E}_t(q) = \left[ A_0 E_1^T \cdots A_F^T \right] \begin{bmatrix} z_q(0, 0) \\ \vdots \\ z_q(t, L_t - L_t - 1) \end{bmatrix}$$

After obtaining the de-punctured matrix, Viterbi decoding is performed. For simplicity of exposition, binary phase shift keying (BPSK) modulation with coherent detection is used. The LLR of the equalized signal is used to recover the transmitted information bits, and the received LLR values for node $q$ with BPSK modulation can be derived as $\text{LLR}(q) = 2E_q(q)\gamma_{vq}$, where $\gamma_{vq}$ is the received signal-to-noise ratio (SNR) at node $q$ for transmissions by node $v$. The decoded vector at node $q$ is denoted as $\hat{u}_q = \text{Viterbi}(\text{LLR}(q))$.

### 2.2. Cooperative transmission protocol

The practical potential of cooperative communication was initially explored in detail in [1–4]. Given the high practical possibility, the promise offered by cooperative communication has attracted significant attention among the wireless research community. In a short span of few years’ time, vast amount of works have been reported in the literature. Significant portion of the works are dedicated toward development of theoretical framework from physical layer and information-theoretic perspective. However, medium access control (MAC) layer protocols in most of today’s wireless standards are designed way before the emergence of cooperative paradigm, and tailoring existing protocols to deliver the promised potential of cooperative communication appears to be nontrivial. The challenges include non-flexible structure of current protocols together with legacy, security, and scalability issues. One recent attempt is found in [17] where cooperative communication is achieved by tweaking the most recent IEEE802.11 standard. Even though such an approach is not fully capable of exploiting the complete benefits of cooperative communication, some of its concept has been applied in the
present paper. A brief description on the MAC implementation in [17] is provided next, followed by details of the transmission protocol for the proposed MNICRIF.

In [17], it is assumed that $S$ keeps a list of potential relays. When $S$ has data to be transmitted to $D$, it will initiate the transmission process by sending packets such as the request to send (RTS) packet. Together with each RTS packet, $S$ appends a request addressed to $R$, the node that can best assist $S$ in relaying its message to $D$. In addition, it is assumed that $S$ predetermines the transmission configuration such as modulation, code rate, and so on and also appends them as part of the initialization request packet. Upon identification of such a message, the appointed $R$ looks out for acknowledgement from $D$ as an indication for $S$ to go ahead with the transmission. If $R$ is capable of delivering the assistance as requested by $S$, it broadcasts an RTR packet to both $S$ and $D$. Following the initialization process, the simple cooperative relay transmission proceeds with transmission from $S$ followed by $R$.

On the basis of the simple MAC protocol mentioned earlier, several extensions or modifications can be introduced to allow more complex multihop transmissions such as the one proposed in the present paper. Note that the following description does not constitute a foolproof MAC protocol. Rather, it is an illustration of how modifications can be made to existing protocols that can potentially support more complex cooperative schemes. In this paper, it is assumed that $S$ performs some kind of advanced optimization analysis to determine the optimal number of hops required for its information delivery to reach $D$, level of participation from relays in each hop, and transmission power that meets the power restriction for each node. To the best of the authors’ knowledge, most incremental redundancy-based cooperative schemes assume the presence of link between $S$ and $D$. In this work, such assumption is relaxed. Similar to [17], $S$ initiates the transmission process by sending an RTS packet to which $D$ (if within the transmission range or the last node in the route) responds with a clear to send (CTS) packet. Routing and MAC for transmission are not considered within the scope of the present work. After the RTS/CTS handshake, $S$ initially encodes the information bits with a cyclic redundancy check code for error detection. The resultant code is then encoded with RCPC codes and transmitted in an incremental fashion with code rates $r_{c0}, r_{c1}$ to $r_{cF}$ ($r_{c0} > r_{c1} > \cdots > r_{cF}$), where $F$ denotes the maximum number of slots or the maximum relay stages in which the additional coded bits are transmitted. The initial code rate is determined on the basis of the capacity across the first hop and has been agreed during the initialization phase. After the first slot, among the $m$ intermediate relay nodes in the next hop, only the best relay node is selected for transmission. The best relay is selected with proper coordination among the nodes based on their forward links (determined by the signal quality of the received CTS packet from $D$) and on the quality of information at the relay (determined by the number of errors detected at the relay).

Considering the transmitting node in the present hop as $S$ and all the nodes in the subsequent hops as relay nodes, the received messages from $S$ are retained in the buffer by all the nodes within its range for future combining with additional coded bits. The best relay in the subsequent hop transmits the next set of coded bits in the next slot. As all nodes transmit in BC mode throughout, $S$ in the previous hop monitors the relay transmissions and determines the quality of relayed information by comparing the received bits with the original set of information bits. On the basis of this comparison, $S$ takes over the transmission of additional coded bits if relay transmissions are found erroneous. CSMA/CA-based medium access technique is utilized, giving higher priority to transmissions by $S$ in the previous hop (it has a lower back-off time compared with other nodes in the subsequent hop). The transmission thus proceeds from hop-to-hop, and the number of coded bits transmitted by a relay node in each hop is assumed to be determined during the initialization phase. This assumption does not affect the system performance as hop-by-hop reliability of the system is guaranteed by the implicit feedback channel. This process continues until the message reaches $D$. More details of the proposed method are discussed in the next section.

### 3. MULTI-NODE INCREMENTAL CODE REDUNDANCY WITH IMPLICIT FEEDBACK

The objective of reliable incremental redundancy-based cooperative schemes is to confine relay transmissions to those cases where relaying is beneficial. This also increases the throughput of the system. However, the overheads required to ensure reliable relay transmissions limit the spectral efficiency of such networks. This overhead increases with the number of nodes in the network. Also, the explicit information about when the relay switches from listening to transmission mode is not known a priori, which further increases the complexity of the system. The proposed MNICRIF exploits the BC nature of wireless transmissions to reduce these overheads in multihop networks. It also ensures hop-by-hop reliable transmission in such networks. The proposed MNICRIF scheme is explained with the help of the flowchart in Figure 2. All the nodes are assumed to be connected with each other, and hence, a direct link is assumed to be present between $S$ and $D$, even though such an assumption can be relaxed for the proposed scheme. Nonetheless, direct link between $S$ and $D$ is made to exist by configuring the forgetting factor (a status flag used to determine whether source transmissions are received at the destination), $FF = 0$.

During the initialization phase, the optimal number of slots ($Q_1$) required for successful delivery of information to $D$ at each relay stage $i$ (where $i = 0$ and $i = F$ denotes $S$ and the last hop in the route, respectively) is predetermined and made available to all nodes in the network. The transmission is initiated by $S$ and commences...
at a high code rate yielding a higher throughput in a good channel. In this paper, non-reception of acknowledgments (ACKs) is considered as erroneous decoding at \( D \). During such instances, the transmitter in the relay stage \( i \) performs the puncturing and de-puncturing operations as explained in Section 2 and transmits the coded bits in an incremental manner for the permissible number of slots, \( Q_i \). Once all the redundant bits for relay stage \( i \) have been transmitted, the best relay node (denoted as \( N_{RX} \)) in the next hop takes over by broadcasting the next set of additional coded bits in small chunks at a time. The intermediate relay nodes are considered to possess minimal processing complexity, and hence, error detection mechanism is not utilized at these nodes. Such schemes where relay decode the received information suffer from potential error propagation. This is mainly induced by erroneous decoding, and transmission of this information leads to severe degradation of the system performance. The proposed scheme mitigates this error propagation by exploiting the BC nature of wireless transmissions. Due to the
BC nature, the transmissions by the relay node in the next hop (denoted as $N_{TX}$) can be overheard by the source node in the previous hop. This node serves as an implicit feedback monitoring node (denoted as $N_{IFM}$) for the relay stage $i$. The quality of transmission by $N_{TX}$ is determined at $N_{IFM}$ by comparing the relayed bits with the original bits. The IFM node ($N_{IFM}$) takes over the role of $N_{TX}$ in transmitting the additional set of coded bits during erroneous transmissions. Otherwise, the transmissions by $N_{TX}$ continues till either, successful decoding occurs at $D$ or for the permissible number of slots for relay stage $i$. The IFM node needs to monitor the transmissions only during the first slots of each relay stage and can enter into sleep mode for other slots once the reliability of relayed information is acknowledged. Although these extra observations of the transmitted signals do not introduce much overhead (except, possibly for the additional receive hardware), most of the transmission schemes tend to ignore or discard them. The previous process is then repeated hop-by-hop till either the coded bits at the maximum rate have been transmitted or successful decoding at $D$ occurs.

The proposed scheme can be clearly explained with an example. Figure 3 illustrates such a simplified network with three relay nodes and four transmission stages. For analysis purpose, the forgetting factor $F_r$ is set as 1 (there is no direct link between $S$ and $D$). Figure 3(a) shows the transmission during initial stage zero. At this stage, $S$ transmits the first set of encoded bits, and all three relay nodes ($R_1$, $R_2$, and $R_3$) are actively receiving. Note that at this stage, $D$ is not yet listening and can enter into sleep mode. Of the three relay nodes, only $R_1$ will perform decoding. Figure 3(b) shows the transmission during stage one. At this stage, $R_1$ transmits the information that it has decoded, whereas $R_2$, $R_3$, and $D$ are actively receiving. $R_1$ begins by transmitting a small number of bits that are exploited as implicit feedback by the IFM node, $S$. If $S$ detects that the decoding at $R_1$ is incorrect, $S$ will take over the transmission of remaining bits for stage one. Assuming that integrity check passes (i.e., $R_1$ is transmitting the correct coded bits), $R_1$ continues with the transmission of redundant bits. After all the bits for stage one are transmitted, $R_2$ and $D$ perform decoding. If $D$ is unable to decode successfully, transmission proceeds to stage two as shown in Figure 3(c). Similarly, $R_2$ begins by transmitting small number of bits, and this time, $R_1$ will be the IFM node. Now, assume that decoding at $R_2$ is erroneous. Transmission is taken over by $R_1$, whereas $R_3$ and $D$ continue receiving. Once all the bits for this stage are transmitted, $R_3$ and $D$ will perform decoding. If decoding is successful, transmission for the current packet is terminated, and the process continues with the next packet.

From the previous description, it can be inferred that MNICRIF guarantees hop-by-hop reliability by the integration of the implicit feedback channel to multihop networks. This ensures the availability of the correct information at $D$ irrespective of relay message quality. The decision of selecting the transmitter at each hop is dependent on the instantaneous channel conditions, which affects the quality of decoding at each hop (relay stage). The inherent implicit feedback channel available due to BC transmissions thereby reduces the overhead due to frequent feedback in time-varying channels.

4. PERFORMANCE ANALYSIS OF MNICRIF

It is highly likely for the decoding to be imperfect at some hops as delivery of information is carried out over individual wireless channels from $S$ to $D$. Decoded information stream with erroneous bits leads to imperfect reconstruction of the primary code matrix. This would then lead to error propagation within a multihop system. To the best of authors’ knowledge, such error propagation in multihop cooperative systems has not been adequately addressed in the available literature, and the present method to retard such propagation is first of its kind. A detailed analysis of the proposed scheme with its error rate probability, the number of hops required for a particular quality of service (QoS), and throughput of the system are presented in this section.

Figure 3. Example of multi-node incremental code redundancy with implicit feedback based network with three relay nodes and four transmission stages: (a) Stage-0, (b) Stage-1, (c) Stage-2, (d) Stage-3, and (e) Stage-3.
4.1. Performance bounds of MNICRIF

In the proposed scheme as illustrated in Figure 4, the coded bits are received from different relay stages based on the relay decoding quality to mitigate error propagation within the system. The present analysis assumes the channels between nodes as uncorrelated and Rayleigh distributed as explained in Section 2. Hence, the coded bits are received over independent channels from the different relay stages and experience independent fading paths. In the proposed scheme, the transmission is initiated (stage 0) by the source node at the specified initial code rate and proceeds as explained in the former section. The number of bits transmitted at each stage is predetermined, and the initial transmission at each stage serves as the implicit feedback, which is exploited to determine the relay detection quality. The transmission proceeds till either an ACK is received from D, or the permissible number of coded bits are transmitted at each stage is predetermined, and the initial transmission is initiated (stage 0) by the source node at the specified initial code rate and proceeds as explained in Section 2. Hence, the coded bits are received from different relay stages based on the ratio of coded bits received from each stage. Let the instantaneous SNR for the i th stage (0 \leq i \leq F) be represented as \( \gamma_i \), where \( \gamma_i = |h_{R_i D}|^2/\sigma_i^2 N_0 \) and \( \gamma_0 = |h_{S D}|^2/\sigma_0^2 N_0 \) (if a direct link exists between S and D). With \( (P + L_0), (L_1 - L_0) \) to \( (L_i - L_{i-1}) \) bits being transmitted as previously discussed during stages 0, 1 to i, the effective instantaneous SNR at \( D \) for the i th relay stage is given as

\[
\gamma = \frac{P + L_0}{P + L_i} \gamma_0 + \frac{L_1 - L_0}{P + L_i} \gamma_1 + \cdots + \frac{L_i - L_{i-1}}{P + L_i} \gamma_i
\]

Equation (5) can be simplified as \( \gamma = \beta_0 \gamma_0 + \beta_1 \gamma_1 + \cdots + \beta_i \gamma_i \) where \( \beta_i \) represents the fraction of the bits received from relay stage \( i \). With Viterbi decoding at \( D \), bit error rate (BER) is defined by its union upper bound as [18]

\[
P_b \leq \frac{1}{P} \sum_{d=d_{\text{free}}}^{\infty} c_d P_d \tag{6}
\]

where \( P_d \) is the probability that a wrong path at distance \( d \) is selected. \( c_d \) is code dependent and equals the total number of encoded values with weight \( d \) paths and \( d_{\text{free}} \) is the free distance of the code. From (6), it is evident that the performance bounds for the system can be computed by evaluating \( P_d \) and \( c_d \) values. The RCPC codes presented in [6] is employed in this paper, and hence, the corresponding set of \( c_d \) values can be applied. For an RCPC system experiencing Rayleigh fading channel with infinite interleaving/deinterleaving (i.e., an ideal memoryless channel), \( P_d \) for coherent detection evaluates to [19, p. 859]

\[
P_d = q^d \sum_{k=0}^{d-1} \binom{d-1+k}{k} (1-q)^k \tag{7}
\]

where \( q \) is the BER of the uncoded system. The instantaneous BER for BPSK transmission for relay stage \( i \) can be derived as \( q(\gamma_i) = Q(\sqrt{2\gamma_i}) \). The average BER (\( q_i \)) is obtained by integrating \( q(\gamma_i) \) with respect to the PDF of the Rayleigh faded signal as [19]

\[
q_i = \frac{1}{2} \left( 1 - \sqrt{\frac{\gamma_i}{1 + \gamma_i}} \right) \tag{8}
\]

As the coded bits are received at \( D \) over independent fading channels, the overall BER of the uncoded system is the weighted sum of the received SNRs across all the relay stages. The weights are dependent on the ratio of coded

![Figure 4. Incremental redundancy-based transmission in a multihop network at different slots.](image-url)
bits received from each relay stage, and the average BER can be given as

\[ q = \sum_{i=0}^{F} \hat{\beta}_i q_i \]  

(9)

where \( F \) is the maximum number of relay stages. Finally, union upper bound of the BER is obtained by substituting the respective \( P_{D} \) obtained from (7) and (9) into (6).

### 4.2. Optimum number of hops for a predefined QoS

In multihop systems, the performance objectives are often in contention with each other. Even though multihop systems achieve a high rate of transmission in conjunction with extended coverage area, the system suffers from lower end-to-end throughput and delay. To obtain favorable throughput and delay performance, the system may choose to have less number of hops. The number of hops should be optimized with respect to the QoS requirement of the system and should be adapted on the basis of the SNRs along the links connecting the nodes in the network.

The formulation of the optimum number of hops problem is based on the following facts. For a specified QoS, the desired BER level, \( P_{D} \), at \( D \) is predefined for the system. On the basis of (6) and (7), the average required BER performance of a multihop system can be computed and is a function of \( \beta_i \) and \( q_i \). Considering the nodes as static, the average SNRs \( \tilde{\gamma}_i \) across the relay stage \( i \) and \( D \) can be computed, which in turn can be employed to evaluate the corresponding BER \( q_i \). Then, the weights or fraction of bits from each relay stage should be optimized on the basis of the required average BER \( q \). The optimization problem can then be defined as

\[ F_{opt} = \arg \min_{F} q(\hat{q}_i, \beta_i) \]

s. t. \( \sum_{i=0}^{F} \beta_i = 1 \)

(10)

where the average BER function is minimized on the basis of the number of relay stages. The constraint \( \sum_{i=0}^{F} \beta_i = 1 \)

reflects the fact that the aggregate of the fraction of coded bits received from different relay stages is bounded to unity. Here, \( i \) is the index of the relay stage, and \( F \) is the maximum relay stage. It is clear that the problem in (10) is quite general with a linear constraint, and the optimal solution can be computed on the basis of the water filling algorithm. For each relay stage, a binary search technique, which has a computational complexity of \( O(F \log F) \), is performed for varying fraction of bits (which satisfies the optimality criteria) to find the optimal fraction, which results in the minimum BER. Then, a binary search is performed across the BER values for the relay stages obtained with this optimal fraction of bits to determine the optimal number of hops. The iterative search to determine the number of relay stages is performed with respect to the QoS requirement of the system [20].

### 4.3. Throughput of the network

The achievable rate of the proposed scheme is determined on the basis of the coded bits, which are received and processed by the receiver from the different intermediate relay stages. In incremental redundancy, the receiver buffers the past received signals and hence accumulates the mutual information. As the mutual information is accumulated over the various relay stages, \( \Pr(i_j < C_R) \leq \Pr(i_i < C_R) \) for \( j \leq i \), where \( i \) and \( j \) characterize the different relay stages, \( C_R \) is the threshold rate (based on the capacity of the channel) for reliable decoding at \( D \), and \( I_x \) represents the accumulated mutual information at stage \( x \) [21]. The corresponding probability can be evaluated as

\[ \Pr(i_i < C_R) = \Pr\left( \sum_{k=0}^{i} \log_2 (1 + \gamma_k) < C_R \right) \]

(11)

where \( \gamma_k \) is the received SNR at stage \( k \). Then, the probability of successful decoding at the \( i \)th stage is given as

\[ \Pr(i_i > C_R | I_0 < C_R, I_1 < C_R, \ldots, I_{i-1} < C_R) = \Pr(i_{i-1} < C_R) - \Pr(i_i < C_R) \]

(12)

Considering successful decoding to occur at the \( i \)th stage transmission, the effective code rate of the proposed incremental redundancy-based system is \( P/(P + L_i) \), where \( L_i \) corresponds to the number of redundant bits received in the \( i \)th stage. The throughput of the proposed scheme with a maximum of \( F \) stage relay transmission can be computed as

\[ \eta = \frac{P}{P + \sum_{i=0}^{F} L_i \Pr(i_i > C_R | I_0 < C_R, I_1 < C_R, \ldots, I_{i-1} < C_R)} \]

(13)

By incrementally transmitting the redundant information, the maximum throughput for the proposed scheme is achieved when the initial transmission itself is successful. It is further evident that the throughput reduces with the number of relay stages. The proposed scheme by mitigating erroneous transmissions offers better reliability, and thereby, a throughput advantage can be obtained.
5. NUMERICAL RESULTS

To verify the theoretical analysis presented in the former section, the proposed MNICRIF is modeled and extensively simulated. The simulations are performed in a realistic communication environment as discussed in Section 2. The incremental coding capability is obtained by employing RCPC codes presented in [6]. In particular, an 8/32-rate mother code with memory size $M = 4$ and puncturing period $P = 8$ is selected. The minimum and maximum numbers of coded bits that can be transmitted are $L_{\text{min}} = 2$ and $L_{\text{max}} = 24$, respectively. This translates to effective code rates of 8/10 and 8/32, respectively, as explained in Section 2. For practical convenience, the number of additional coded bits transmitted is always a multiple of two.

The nodes in the $n$-hop network depicted in Figure 1 are connected with each other through wireless links, and the channel between nodes is modeled as Rayleigh based on generalized Jakes’ model [22]. The typical pathloss exponent ($\alpha$) for mobile networks is in the range 3–5, and a value of 3.5 is chosen to model an urban environment.

For simplicity, the ACK/NACK feedback is assumed to be perfectly received, and a perfect implicit feedback channel is assumed. The nodes in the network are considered to be fixed and collinear across the $n$ hops. The intermediate nodes thereby divide the direct path between source and destination nodes into equal length segments. The nodes are assumed to be sufficiently separated such that the fading between each links is assumed to be identical but independently distributed (i.i.d.). Such a configuration serves to validate the analytical results and also to illustrate the benefits that can be realized by an optimal placement of nodes. A uniform power constraint as explained in Section 2 is also employed to comply with the total power constraint of the network. The code rates applicable for each relay stage ($P/(P + L_i)$) can be obtained from the corresponding redundant bits $L_i$’s received from the $i$th relay stage and is given in Table I.

### Table I. Redundant bits $L_i$’s received from $i$ relay stages.

<table>
<thead>
<tr>
<th>2 Hop</th>
<th>3 Hop</th>
<th>4 Hop</th>
<th>6 Hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_0$</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$L_1$</td>
<td>24</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>$L_2$</td>
<td>-</td>
<td>24</td>
<td>16</td>
</tr>
<tr>
<td>$L_3$</td>
<td>-</td>
<td>-</td>
<td>24</td>
</tr>
<tr>
<td>$L_4$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$L_5$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

5.1. BER Performance

The BER performance of the proposed MNICRIF scheme obtained through simulation exercise is first investigated and compared with the respective theoretical closed-form equations derived in the previous section. Figure 5 illustrates the BER performance where the straight lines represent the theoretical bounds and the dashed lines represent the simulated performance. From the figure, it can be inferred that the theoretical upper bound shows close correspondence with simulated performance at low BER’s (at high SNRs). It is interesting to note that multihop schemes exhibit similar slopes. Such behavior results from the temporal diversity that is inherent in ideally interleaved coding schemes. The intermediate relay nodes significantly shorten the propagation path, which effectively improves the BER performance of the network. Hence, the BER performance improves with the number of hops as shown in Figure 5. However, more number of hops degrades the throughput and delay performance. The number of hops should be selected on the basis of the QoS requirement of the system and on the SNR between the nodes comprising the network as described in the subsequent section.

![Figure 5. Theoretical (straight lines) and simulated (dashed lines) bit error rate performance comparison of multi-node incremental code redundancy with implicit feedback.](image-url)
5.2. Optimum number of hops

In multihop networks, the number of hops should be properly chosen to minimize the end-to-end delay and maximize the throughput [23]. As discussed in the former section, QoS requirement of the system determines the permissible number of hops. Figure 6 illustrates the optimum number of hops for a typical QoS requirement of received error rate to be less than $10^{-5}$ and $10^{-6}$. The optimum number of hops for a target QoS is computed from the average BER by following the steps illustrated in Section 4. The contribution of coded bits from each relay stage is given in Table I. As expected, the optimum number of hops decreases with an improvement in link quality between nodes (as SNR increases). Further, lesser number of hops is required to achieve a reduced QoS. It can be seen that for any target QoS, direct transmission is feasible only at high SNRs. This makes multihopping to be the best method of communication when the source and destination are separated by a large distance in the medium and low SNR range.

5.3. Effect of implicit feedback channel on throughput

An important measure of the performance of any communication system is throughput, which is defined as the ratio of decoded information bits to the total number of coded bits transmitted. For an incremental redundancy-based system, the total number of coded bits transmitted is dependent on the channel characteristics and the code rate employed by the system. The throughput of the proposed scheme is derived in Section 4. Figure 7 compares the performance of the proposed MNICRIF scheme with normal cooperative relaying systems. In normal cooperative relaying systems, the code rate between relay stages are predetermined as explained in Section 2, and relay takes over the transmissions once the permitted number of coded bits are received. However, in such cases, there can be erroneous relaying due to imperfect decoding at relay nodes. The proposed MNICRIF scheme utilizes the BC nature of the relay transmissions to determine the relay quality, which eliminates the need of explicit feedback slots. During erroneous relaying, the IFM node (the transmitter in the previous hop) takes over the transmissions. This ensures the reliability of information at the destination. From Figure 7, it can be observed that the proposed scheme provides a throughput improvement by mitigating erroneous relaying. Along with conserving feedback slots and reduced processing complexity at relay nodes, the higher throughput offered by the proposed MNICRIF makes it attractive for multihop systems.

5.4. Effect of implicit feedback with number of hops

Figure 8 illustrates the impact of implicit feedback channel on varying number of hops. It shows the percentage of

![Figure 8](image_url)
Figure 9. Percentage of initial and cooperative bits received at D of $\frac{8}{12}$, (b) initial rate of $\frac{8}{14}$, and (c) initial rate of $\frac{8}{16}$.
times the IFM node takes over erroneous relay transmissions with respect to the received SNR. At low SNRs or when the links are least reliable, the IFM node always takes over the transmissions due to imperfect decoding at the relay nodes. The quality of information at relay nodes improves with link quality (SNR), which reduces the number of times IFM node takes over. At high SNRs, the relay decodes reliably and forwards the correct information. The implicit feedback is a requisite at the low and medium SNR region where relay decoding can be erroneous due to an unreliable intermediate channel. Due to the collinear arrangement of equidistant nodes, the range of SNR values where the IFM nodes takes over the transmissions decreases with increasing number of hops.

5.5. Effect of initial code rate on system performance

The significance of the initial code rate on the performance of the system is illustrated in Figure 9. From the analysis on throughput, it is evident that the maximum performance of the system is achieved when D decodes successfully at the initial rate. The percentage of bits decoded at D for the proposed scheme from initial source and subsequent cooperative relay transmissions with varying SNRs are illustrated in Figure 9. A two-hop network is considered, and the performances with respect to varying source-relay distance are also illustrated. The graphs representing the percentage of decoded information bits received from initial source and subsequent relay transmissions cross over when the contribution from both transmissions remains the same (50%). It can be observed that the percentage of bits received from initial source transmissions increases with the initial code rate \( P = P_L + L_0 \). This result in less bits being decoded from subsequent relay transmissions and thus the crossover between contributions from both transmissions occur at lower SNRs.

The decoding quality at the relay reduces with increasing distance from the source, which results in more cooperative transmissions and lesser bits being decoded from the initial source transmission. It is observed that maximum and minimum percentages of decoded information bits are obtained from initial source and subsequent cooperative relay transmissions, respectively, when the relay is placed between the source and destination nodes.

6. CONCLUSIONS

In this paper, a spectrally efficient cooperative transmission scheme for multihop networks is proposed. The proposed scheme mitigates erroneous transmissions by exploiting an implicit feedback channel available due to the BC nature of relay transmissions. This implicit feedback validates the subsequent transmissions, which enhances the spectral efficiency and reliability of such networks. Further, the error detection mechanism needs to be employed only at the destination for such schemes, which reduces the complexity of signal processing at intermediate relay nodes. A performance upper bound for the proposed scheme employing convolutional code in a Rayleigh fading channel is derived in this paper. Extensive simulation studies are also performed to verify the analytical results obtained in the paper.

ACKNOWLEDGEMENTS

The authors would like to recognize the efforts of Dr. Surya Dharma Tio who was initially part of this work. His assistance in compiling some of the results and initial formulations were very helpful.

This work was supported by Agency for Science, Technology and Research (A*STAR), Singapore under A*STAR-NIH joint grant (SERC grant No: 102 149 0154).

REFERENCES


Efficient multihop transmission scheme for cooperative networks

A. James, A. S. Madhukumar and F. Adachi

1. The authors' biographies

Ashish James received his Bachelors degree from College of Engineering, Thiruvananthapuram, Kerala, India and his PhD degree from Nanyang Technological University, Singapore. He is currently a postdoctoral fellow with the School of Computer Engineering at Nanyang Technological University. His research interests include coding theory, cooperative communications, multiple access techniques, and femtocells.

A. S. Madhukumar received his B. Tech degree from College of Engineering, Trivandrum, India. M. Tech from Cochin University of Science and Technology, and PhD from Indian Institute of Technology, Chennai. He is currently an Associate Professor in the School of Computer Engineering, Nanyang Technological University, Singapore. He was involved in communications and signal processing research at Center for Development of Advanced Computing (formerly known as Electronics R&D Centre), Trivandrum, India and Institute for Infocomm Research (formerly known as Center for Wireless Communications), Singapore. His research interests are in the areas of signal processing algorithms, new modulation and multiple access schemes, reconfigurable radio systems, cooperative communication, and future wireless communication systems. He has published over 170 referred international conference and journal papers. Dr. Madhukumar is a senior member of IEEE.

Fumiyuki Adachi received the B.S. and Dr. Eng. degrees in Electrical Engineering from Tohoku University, Sendai, Japan, in 1973 and 1984, respectively. In April 1973, he joined NTT Laboratories and conducted various types of research related to digital cellular mobile communications. From July 1992 to December 1999, he was with NTT DoCoMo, Inc., where he led a research group on W-CDMA for 3G systems. Since January 2000, he has been with Tohoku University...
University, Sendai, Japan, where he is a Professor of Communications Engineering at the Graduate School of Engineering. His research interests are in gigabit wireless signal processing and networking including wireless access, equalization, transmit/receive antenna diversity, channel coding, and distributed MIMO signal processing. He is an IEEE Fellow and an IEICE Fellow. He was a recipient of the IEEE Vehicular Technology Society Avant Garde Award 2000, IEICE Achievement Award 2002, Thomson Scientific Research Front Award 2004, Ericsson Telecommunications Award 2008, Telecom System Technology Award 2010, Prime Minister Invention Prize 2010, and KDDI Foundation Research Award 2012.