

RESEARCH

Open Access

# Opportunistic wireless network coding with relay node selection

Jangseob Kim and Jungwoo Lee\*

## Abstract

Broadcasting nature of wireless communications makes it possible to apply opportunistic network coding (OPNC) by overhearing transmitted packets from a source to sink nodes. However, it is difficult to apply network coding to the topology of multiple relay and sink nodes. We propose to use relay node selection, which finds a proper node for network coding since the OPNC alone in the topology of multiple relays and sink nodes cannot guarantee network coding gain. The proposed system is a novel combination of wireless network coding and relay selection. In this paper, with the consideration of channel state and potential network coding gain, we propose several relay node selection techniques that have performance gain over the conventional OPNC and the conventional channel-based selection algorithm in terms of average system throughput.

## 1 Introduction

Channel coding concept is used to mitigate the influence of noise and interferences in the physical layer. In [1], it was also shown that we can get coding gain in higher layers. Compared to the routing and scheduling techniques that are devised to prevent bottlenecks of packets from different senders, Alswede et al. [2] showed a way of making use of this disadvantage and showed that the achievable rate can be increased by applying certain in-network processing at an intermediate node when packets are received at the node simultaneously. This type of in-network processing is called network coding. Routing can be treated as a special case of network coding which is a simple permutation. Network coding has received attention since it can enhance system throughput and reliability. For throughput, network coding technique can take advantages of bottleneck effect of data at the intermediate node in wireless communication to improve the system throughput [3]. Ghaderi et al. [4] have shown that there are reliability benefits by applying network coding technique in their system. Li et al. [5] show that the maximum achievable rate can be achieved by linearly combining input packets at an intermediate node. Random linear network coding [6] (RLNC) and opportunistic network coding [7] (OPNC) have been known as one of practical implementations. RLNC randomly chooses elements from

a finite field as the coefficients for a linear combination of packets.

OPNC performs bitwise XOR operation of packets that are selected by reception report. RLNC is suitable for the distributed system, and no reception report is needed since it contains all the information in the header to decode the received packets at the receiver node. However, as the number of hops or the number of participants increases, the length of the header also increases, which might degrade the throughput. Although OPNC needs extra report, the portion is not significant compared to the original information, and the implementation of coding and decoding is simple. As a practical implementation of OPNC, Katti et al. [7] introduced a scheme, COPE, that takes advantage of broadcasting nature of wireless communications.

COPE employs practical network coding technique for unicasts in wireless mesh networks to improve total throughput. They showed through experiments that with OPNC in the system, there exist significantly improvements in throughput of wireless networks with UDP traffic. Recently, Fang et al. [8] gave an analysis of COPE and argue that the key to COPEs success lies in the interaction between COPE and the MAC protocol. How MAC protocol deals with competing nodes in a given network plays an important role in performance improvement. In this paper, we consider the following two factors: one factor is the channel state information, which can affect the performance of a system, and the other factor is how to deal

\* Correspondence: junglee@snu.ac.kr  
School of Electrical Engineering, Seoul National University, Seoul, Korea

with multiple intermediate nodes, which can perform network coding simultaneously. This kind of networks, without certain decision methods at the intermediate nodes, cannot guarantee the throughput gain by using network coding in the system as in [7].

An uplink model that consists of multiple users, multiple relays, and a single base station (receiver) was used in [9], which proposed finite field network coding with superposition coding at the relay nodes. In [10] which uses the same uplink system model, they replaced the relay nodes with a set of user nodes. With the multi-user cooperative communication system, they proposed a diversity network coding scheme over finite fields. In [11], a down link model was considered, and it consists of a single transmitter base station, a single relay node, and multiple receiving user nodes. They proposed an instantaneously decodable binary network coding scheme and showed its improved transmission efficiency compared to the existing ARQ and network-coding-based schemes. Bletsas et al. [12] dealt with a cooperative communication system consisting of single source node, single sink node, and multiple relay nodes and introduced a distributed network path selection algorithm which performs opportunistic relaying by using an objective function based on the channel states at the relay nodes.

In this paper, we consider a system model that includes multiple relay nodes and multiple sink nodes. With this system model, we combine the opportunistic relaying with network coding and propose a relay selection measure which considers the channel state between the relays and the destination nodes. We compare the performance of proposed algorithms with conventional OPNC and opportunistic relaying in terms of throughput. The rest of this paper is organized as follows. The system model is described in the section of system model and scenario. In the section of proposed relay selection techniques for network-coded transmission, we propose several relay selection schemes for network-coded transmission. The performance of these schemes is compared with the conventional relay selection schemes. The results are verified by simulations in the section of simulation results, and we draw our conclusions in the last section.

## 2 System model and scenario

The system scenario and the system model are introduced in this section. A source node has packets that need to be delivered to different destinations. There are multiple relay nodes, some of which might have better channels to the destination than the channel between the source and the destination. After the source broadcasts the packets, some packets may not reach their destination nodes successfully, and it is needed to retransmit the missing packets. Since some destination nodes overhear the packets which are sent to other destination nodes,

network coding can be effective in this scenario. With network coding and relay selection, the best intermediate node for retransmission is selected.

### 2.1 Transmission from source to neighbor nodes

We have a source node  $S$ , a set of relay nodes  $R$ , and a set of sink nodes  $D$ . Assume that  $S$  has  $n$  packets to transmit to corresponding sink nodes (i.e.,  $S_a = \{\alpha_1 \dots \alpha_n\}$ ),  $R$  include  $l$  nodes ( $R = \{r_1 \dots r_l\}$ ), and  $D$  includes  $m$  elements ( $D = \{d_1 \dots d_m\}$ ). Each packet  $\alpha_i \in S_a$  has its own destination address to be delivered. We assume all nodes in  $R$  and  $D$  are within communication range from  $S$ . At first, the source node  $S$  broadcasts  $n$  packets to all the nodes in its range. Every neighbor node is assumed to be able to overhear data traffic of other nodes as in OPNC and stores all the overheard packets in its buffer. A relay node  $r_j$  receives a set of packets,  $\alpha_j$ , and a sink node  $d_i$  gets a set of packets,  $\beta_i$ . Both  $\alpha_j$ s and  $\beta_i$ s are subsets of the original  $n$  packets. ( $n \geq |\beta_i|, |\alpha_j|, \forall d_i \in D$  and  $\forall r_j \in R$ ).

After the source transmission is over, there may be packet loss at sink nodes due to a poor channel between source and those nodes. Hence, we need retransmissions for those missing packets. If the source retransmits data, the packet loss may occur again. If there exists a relay node ( $r_j$ ) with better channel response than the source node  $S$ , it may be better for  $r_j$  to retransmit the packet to the destination. It is assumed that the relay set  $R$  receives all the packets that the source sent. We then have

$$\cup\{\alpha_j | \forall r_j \in R\} = S_\alpha. \quad (1)$$

This means that the union of packets of all relay nodes is identical to the set of all the packets from the source  $S$ . The number of packets from the source ( $n$ ) should be less than the buffer size to prevent overflow. When  $n$  is larger than the buffer size, we can divide  $n$  packets into a number of groups as in a practical RLNC scheme [13].

Next, it will be shown that the probability to satisfy (1) is close to 1 in the high SNR regime. Let  $\eta_0$  be an event that at least one node in  $R$  correctly receives a packet from the source and  $\eta_0^c$  be the complement of  $\eta_0$ . The relationship between two events is

$$P(\eta_0) = 1 - P(\eta_0^c). \quad (2)$$

where  $P(\cdot)$  is the probability. Assume the channel response is independent, then  $P(\eta_0^c)$  means that no node in  $R$  receives a correct packet. We have

$$P(\eta_0^c) = \varepsilon_p^l \quad (3)$$

where  $\varepsilon_p$  is the packet error rate given by  $\varepsilon_p = 1 - (1 - \varepsilon_{s,M})^N$ ,  $l$  is the size of set  $R$ ,  $\varepsilon_{s,M}$  is the symbol error rate

of M-QAM, and  $N$  is the number of symbols in a packet. We can calculate the lower bound of  $P(\eta_0)$  using the upper bound of  $\varepsilon_{s,M}$  from [14]. We then have

$$\varepsilon_{s,M} \leq 4Q \left( \sqrt{\frac{3kE_{\text{bav}}}{(M-1)N_0}} \right) \quad (4)$$

where  $M$  is the modulation order of QAM,  $k = \log_2 M$ ,  $E_{\text{bav}}$  is the bit energy, and  $N_0$  is the noise variance. Note that the upper bound in (4) is for an AWGN channel. By plugging (3) and (4) into (2), we can calculate  $P(\eta_0)$ . If  $n$  packets are transmitted, the probability that there is at least one relay node which receives each packet is simply the  $n$ th power of (2) since the channel is independent. Let us denote the event that satisfies (1) by  $\eta_1$ .

$$P(\eta_1) = P(\eta_0)^n \quad (5)$$

Since

$$\frac{E_s}{N_0} = \frac{\|h\|^2 P_s B}{N_0 R_s k}, \quad (6)$$

where  $E_s$  is the symbol energy,  $h$  is the channel response,  $P_s$  is the symbol transmit power,  $B$  is the channel bandwidth, and  $R_s$  is the symbol rate. From (2) to (5),  $P(\eta_1)$  is lower-bounded by

$$P(\eta_1) \geq \left[ 1 - \left\{ 1 - \left( 1 - 4Q \left( \sqrt{\frac{3\|h\|^2 P_s B}{(M-1)N_0 R_s}} \right) \right)^N \right\}^n \right]. \quad (7)$$

Figure 1 shows the plots of Equation 7. We use  $B = 5$  MHz, 16-QAM,  $R_s = 2$  bps, and Rayleigh fading channel for  $h$ . The plots indicate that the probability of having at least one relay node with a correctly received packet approaches to 1 at high SNR. Therefore, it is enough for the relays instead of the source to retransmit data.

## 2.2 Retransmission procedures

### 2.2.1 Reception report from the destinations to the relays

Each of relay and destination nodes operates in opportunistic listening mode which stores every received packets for a given period regardless of the destination. The storing period is a system dependent variable (500 ms in [7]). After the source transmission, each destination  $d_i \in D$  creates a report packet and sequentially broadcasts it to all the relays. Since there are multiple sink nodes in  $D$ , each sink node uses a random access method such as CSMA/CA to avoid collision. The report packet is sent to the source and the relay nodes. The information in the report packet consists of the source node ID, the current node ID, multiple original sink node IDs of received packets, and pilot signal as shown in Figure 2.

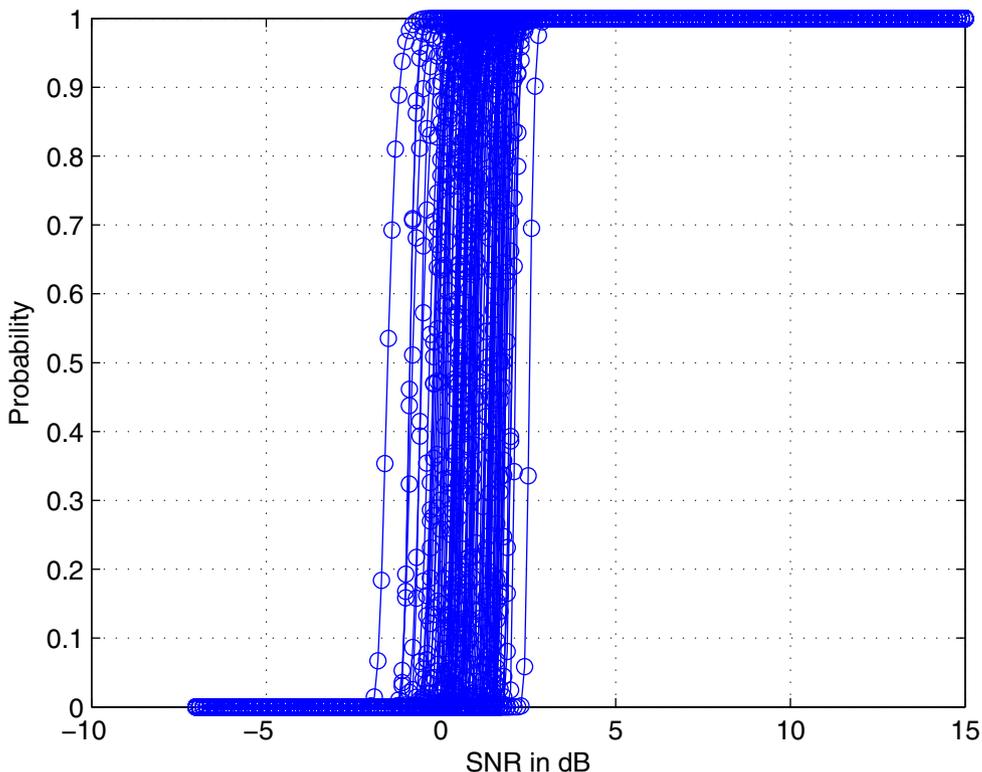
The portion of report packet is not significant compared to the information packet as indicated in [7]. Let us denote the number of packets at the source node by  $n$ , which is known to all the nodes. The report packet consists of a pilot, a source node ID, a current sink node ID, and the destination sink IDs of the  $|\beta_i|$  received (stored) packets. We assume that the IDs are represented by 64 bits as in the IPv4 format. Let us denote the pilot size by  $l_1$ , the packet size by  $l_2$ , and the number of network-coded packets by  $l_3$ . A packet then needs at least  $64 + n \log_2 n + l_1$  bits. Before the retransmission from a relay node to its destinations, the relay receives  $m$  report packets, where  $m$  is the number of sink nodes. The ratio ( $r$ ) of the overhead due to the report packets is

$$r = \frac{m(64 + n \log_2 n + l_1)}{l_2 l_3} \quad (8)$$

For example, if  $n = 10$ ,  $l_1 = 2$ ,  $l_2 = 1$  kB/packet,  $m = 10$ , and  $l_3 = 3$ , we have  $r = 4\%$ . Note that the feedback is performed at a packet level instead of a symbol level, and the overhead of the report packets is not too significant compared to the overall data traffic as can be seen in the example. The report packet transmitted from each sink node is overheard by each node in  $R$ . Based on the information in these report packets, each relay  $r_j \in R$  estimates the channel state to each destination and calculates the objective function which will be used for selecting the retransmitting node in a distributive manner.

### 2.2.2 Retransmission procedure from a relay node

After the packet report, each  $r_j$  has the knowledge of the packet set  $\beta_i$  of the destination  $d_i$  and estimates the corresponding channel response  $h_{ji}$  between  $r_j$  and  $d_i$  ( $1 \leq i \leq m$ ,  $1 \leq j \leq l$ ). Using that knowledge, each relay  $r_j$  checks its buffer for possible network coding. If there are more than 2 packets, it checks whether the packets can be network-coded or not. If affirmative, the relay node  $r_j$  creates a network-coded packet using the OPNC algorithm. If it is not possible to do network coding, the relay node simply retransmits only one packet without using network coding. If no relays get certain packets from the source, these packets are to be delivered directly from the source during the next source broadcasting phase. In the OPNC algorithm, the optimal network coding can be constructed based on how many packets  $r_j$  can mix to create a network-coded packet (i.e., how many destinations would receive packets). However, since the operation does not consider channel response between the relay node  $r_j$  and its destination node in  $D$ , the decoding failure may occur with high probability when the channel quality is poor. This failure increases the retransmission number and degrades



**Figure 1** The probability that there exists a relay node which receives a given packet correctly approaches to 1 at high SNR.

system performance such as throughput. To improve the throughput, we need to modify the selection rule by considering the channel state. We will define an objective function which depends on the number of packets that can be network-coded as well as the channel state, and the retransmission node will be chosen by this function.

Opportunistic relaying was introduced in [12], which proposed a distributed relay selection algorithm for a system which has multiple relays and single sink node. The basic idea is that each relay node sets up an internal timer which triggers transmission. This timer is a function of the channel responses of source-relay and relay-sink pairs, and it is given by

$$T_i = \frac{c}{h_i} \quad (9)$$

where  $T_i$  is the timer function of the relay  $R_i$ , and  $c$  is a constant. There is possibility of hidden node problem, which can be mitigated by adjusting the constant  $c$  in (9). Another method to reduce the hidden node effect is that we use the minimum channel response instead of harmonic mean value [12]. Hence,  $h_i$  is defined as a minimum of the channel responses of  $S - R_i$  and  $R_i - D$ , which is given by

$$h_i = \min(\|h_{SR_i}\|^2, \|h_{R_iD}\|^2). \quad (10)$$

When the timer has expired, the relay node is expected to broadcast a channel reservation message to neighboring relays to prevent other relays from transmission. The relay whose timer expired first broadcasts a channel reservation message to the neighbors. Contrast to [12], in our model, we do not need to consider



**Figure 2 Report packet structure.** Pilot is used to estimate channel state from a sender in  $D$  to a receiver in  $R$ . SrceID is the source identification, PSID is the current sink node identification, and OSID <sub>$i$</sub>  is the destination node identification of the  $i$ th stored packet in the buffer of the current sink node.  $\beta_i$  is the set of the packets overheard by the current node  $d_i$ .

the channel between the source and the relay node since only the relay performs retransmission. This will reduce processing delay in relay node selection. Based on this idea, we propose a new distributed relay node selection algorithm combined with OPNC for the topology of multiple relays and multiple sink nodes. If we use channel state as the only variable to choose a relay node as in the opportunistic relaying algorithm, the system performance may be poor.

Figure 3 shows an example of overall system scenario. There are a single source  $S$ , 2 relays, and 3 sink nodes. Each relay node has different amount of packets to be delivered to its destinations. After the source  $S$  broadcasts, the relays ( $R_1, R_2$ ) and the sink nodes ( $D_x, D_y, D_z$ ) overhear packets and stored them in their buffer. The packets  $a, b$ , and  $c$  are sent to  $D_x, D_y$ , and  $D_z$ , respectively.  $R_1$  sends 1 packet to  $D_x$  and  $R_2$  sends 2 network-coded packets to  $D_y$  and  $D_z$  in one time frame. Suppose  $h_1 = h_{1x}$  and  $h_2 = \min(h_{2y}, h_{2z})$ . We can then calculate the theoretical throughputs  $R_1$  and  $R_2$  for the two channels.

$$R_1 = \log_2 (1 + \|h_1\|^2 \rho) \tag{11}$$

$$R_2 \geq 2 \log_2 (1 + \|h_2\|^2 \rho) \tag{12}$$

where  $\rho$  is the transmit signal-to-noise ratio. The multiplication factor of 2 in (12) is due to the network coding at  $R_2$ , and the inequality is used because the larger

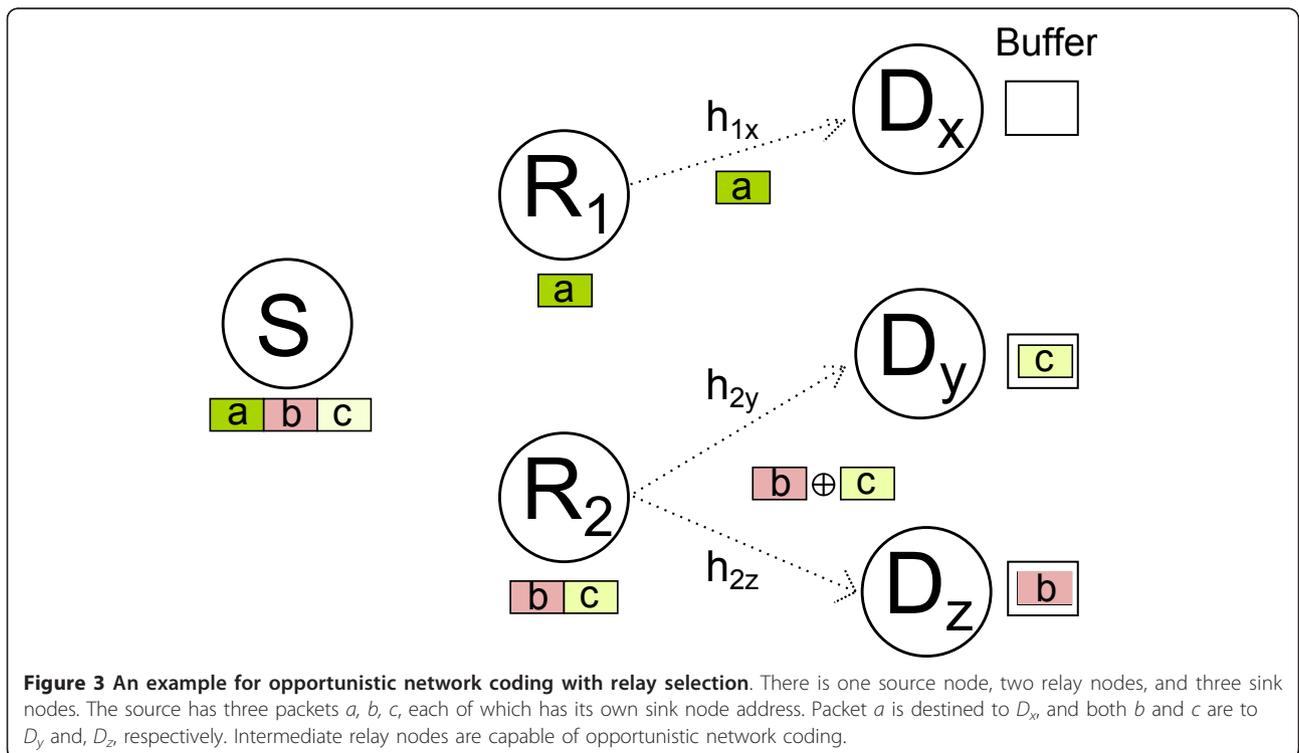
of the two channels has larger capacity than the capacity of the minimum channel. Suppose  $\|h_1\| > \|h_2\|$ , then opportunistic relaying algorithm will choose  $R_1$  to transmit packet  $a$  to  $D_x$ . However, if  $(1 + \|h_2\|^2 \rho)^2 > (1 + \|h_1\|^2 \rho)$ , it may be better to choose  $R_2$  for retransmission.

### 3 Proposed relay selection techniques for network-coded transmission

In this section, we propose relay selection techniques for network-coded transmission, which is based on a timer function. Let us denote the minimum channel response at the  $j$ th relay node by  $h_j$ , and the set of packets that can be network-coded by  $K_j$ . To improve throughput, we consider channel state information ( $h_j$ ) as well as the number of packets ( $\|K_j\|$ ) that each  $r_j$  can deliver simultaneously by network coding. We assume that the objective function at the relay node  $r_j$  is a function of  $h_j$  and  $\|K_j\|$ , which is denoted by  $f(h_j, \|K_j\|)$ . The objective function  $f$  is an increasing function of each variable. The minimum channel response ( $h_j$ ) from relay  $r_j$  to a sink node is a modified version of (10) since only the relay nodes can retransmit. We then have

$$h_j = \min_{\forall i, d_i \in D} \|h_{ji}\|. \tag{13}$$

The number of packets that the  $j$ th relay node  $r_j$  uses to create a network-coded packet is denoted by  $\|K_j\|$ . Both variables,  $h_j$  and  $\|K_j\|$ , may vary from one frame to



another. Though  $h_j$  can be defined differently from (13), Bletsas et. al. [12] empirically showed that it works well to use the minimum channel. Since the objective function is proportional to  $h_j$  and  $\|K_j\|$ , a relay  $r_j$  which has either larger channel response or larger number of packets that can be network-coded will have high probability of using the channel. We can then define the internal timer value at the relay node  $r_j$  as

$$T_j = \frac{c}{f(h_j, \|K_j\|)}. \quad (14)$$

We will use the timer value in (14) in choosing a proper relay node for retransmission. This means that a node with smaller internal timer value will transmit earlier than other relays, which is a kind of decentralized selection scheme. We compare 5 relay selection algorithms using different internal timer functions. First, set the objective function  $f$  as a modified version of opportunistic relaying algorithm of [12]. In this case, the function  $f$  at a certain relay node  $r_j$  depends only on the channel states between the relay and its corresponding sink nodes (13). Those sink nodes are the destinations of the packets that can be network-coded among all overheard packets in  $r_j$ . As mentioned before, we use only the channel between a relay node and a destination node unlike the original opportunistic relaying scheme of (10). Thus, the 1st kind of timer function for the modified opportunistic relaying algorithm is given by

$$T_j^A = \frac{c}{h_j} = \frac{c}{\min_{v_i, d_i \in D} \|h_{ji}\|}. \quad (15)$$

As in the method of OPNC in choosing the best network coding option to increase system throughput, we use only  $\|K_j\|$  as a variable of the objective function. In this case, the 2nd kind of timer function is inversely proportional to  $\|K_j\|$ , which is given by

$$T_j^B = \frac{c}{\|K_j\|}. \quad (16)$$

This means that the relay whose  $\|K_j\|$  is the largest would occupy the channel.

Let us now we introduce sum rate  $R_j^S$  which is given by

$$R_j^S = \sum_{q \in K_j} R_{jq} = \sum_{q \in K_j} \log_2(1 + \|h_{jq}\|^2 \rho). \quad (17)$$

Since  $R_j^S$  depends on both  $\|K_j\|$  and the channel response, we can use  $R_j^S$  as a variable in the objective function. In (18), we use both  $\|K_j\|$  and  $R_j^S$  in the 3rd kind of objective function. As  $R_j^S$  has a channel-related variable in it, the objective function considers the effect of channel and throughput simultaneously, and we have

$$T_j^C = \frac{c}{\|K_j\| \sum_{q \in K_j} \log_2(1 + \|h_{jq}\|^2 \rho)}. \quad (18)$$

In (18), we can replace the sum rate by the minimum channel. The 4th kind of timer function is given by

$$T_j^D = \frac{c}{\|K_j\| h_j} = \frac{c}{\|K_j\| (\min_{v_i, d_i \in D} \|h_{ji}\|)}. \quad (19)$$

The 5th kind of timer function is based on the minimum channel  $h_j$  and the sum rate  $R_j^S$  which is given by

$$T_j^E = \frac{c}{(\min_{v_i, d_i \in D} \|h_{ji}\|) \sum_{q \in K_j} \log_2(1 + \|h_{jq}\|^2 \rho)}. \quad (20)$$

As we mentioned,  $c$  is an empirical constant to control the collision among the relay nodes. Typically,  $c$  has a value of a few microseconds [12]. Each relay node  $r_j$  uses  $T_j$  as its internal timer value. A relay node whose internal timer expires first broadcasts a signal to neighbor relays to stop their transmission to reserve the channel, which is a first-come-first-serve policy. The sink nodes that successfully overhear the network-coded packets decode the packets using their own stored data and update their decoding results. After that, the sink nodes transmit report packets again. Until there are no more packets to be delivered from the relay nodes to the sink nodes, the procedure is repeated.

## 4 Simulation results

The simulation environment is summarized in Table 1. The channel from one node to another is modeled as independent Rayleigh fading channel. This is equivalent to the case where the relay nodes and the sink nodes are randomly distributed around the source node with equal distance. It is also assumed that the relay nodes and the sink nodes are within the communication range from the source node, and the feedback channel is error free. We use a large number of relay and sink nodes in the simulations to order to increase the possibility of network coding at the relay nodes.

### 4.1 Average transmission number

Five different timer function algorithms are compared in terms of average number of transmissions in Figure 4, where the number of relay nodes and the number of sink nodes are set to 50 and 50, respectively. The first

**Table 1 Simulation environment**

Channel model	Rayleigh fading channel
Packet size	1 KB/packet
Modulation order	16 QAM
Channel code	Convolutional code of rate $\frac{1}{2}$

two (A and B) algorithms are conventional ones, and the other 3 algorithms are proposed ones. At high SNR, all the curves converge to 1 (average number of transmissions), which is expected. Hence, we need to focus on the low SNR regime, where a larger number of retransmissions is needed. Algorithm A ( $T_j^A$ ) in Figure 4 is based on the opportunistic relaying algorithm of (15) which chooses a relay with the maximum channel amplitude  $h_j$  from the relay nodes in  $R$ . Algorithms B through E (with the timer function  $T_j^B$  through  $T_j^E$ ) in Figure 4 are based on the timer functions of (16) through (20). From these results, it is observed that the relay node selection algorithm using the timer function of (15) needs the largest average number of transmissions, while the algorithm of (20) requires the least. Algorithm B ( $T_j^B$ ) based on the modified OPNC

algorithm chooses a relay node with the largest number of packets that can be network-coded. Note that this algorithm does not consider the channel response. If there occurs deep fading on the path from the chosen relay node to its sink nodes, it is highly likely that the selected node would fail to deliver the information. That may increase the total number of transmissions. Algorithm C ( $T_j^C$ ) is based on the sum rate and the amount of packets to be network-coded ( $\|K_j\|$ ). By using these two variables in the objective function, the performance is improved over the two previous algorithms. Algorithm D ( $T_j^D$ ) is based on  $\|K_j\|$  and the  $h_j$  simultaneously, so Algorithm D can be thought of as a combination of Algorithms A and B. In Figure 4, it is observed that the performance of Algorithm D is better than previous three algorithms.

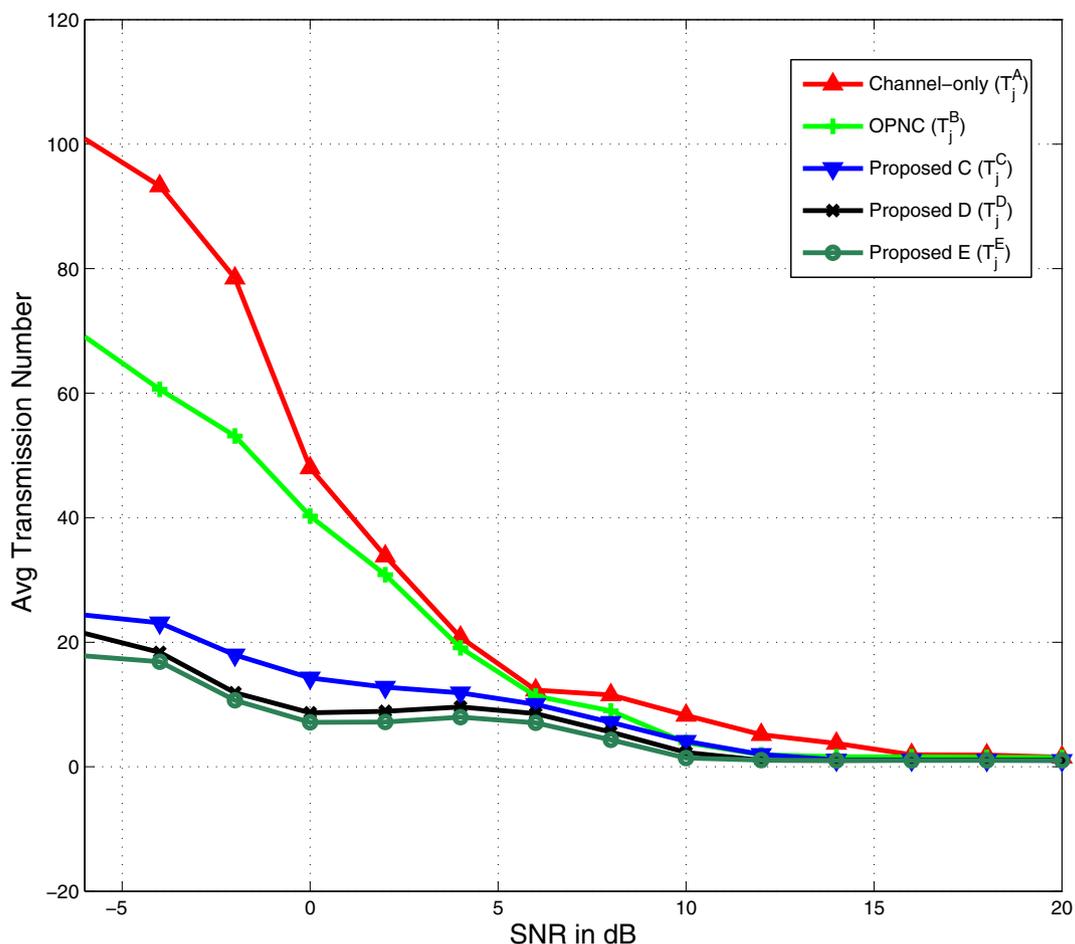


Figure 4 Comparison in terms of average number of transmission.

Algorithm E ( $T_j^E$ ) is based on the channel response and the sum rate with the timer function (20). Compared to the modified OPNC algorithm (Algorithm B), the sum rate measure lowers the possibility of choosing a node whose channel response  $h_j$  is low. Compared to the modified opportunistic relaying algorithm (Algorithm A), Algorithm E considers the sum rate as a measurement of throughput (related to  $\|K_j\|$ ) so that this algorithm balances the measures of  $\|K_j\|$  and  $h_j$ . Figure 4 shows that Algorithm E has the lowest average number of transmissions especially in the low SNR regime. At the high SNR regime, the transmission from the source to the sink nodes would succeed with high probability as mentioned before. In other words, there is not noticeable difference between different algorithms at the high SNR regime.

#### 4.2 System throughput

Figure 5 compares the average system throughput of Algorithms A through E, and the plot is normalized by

the total number of packets used in the simulation. The system throughput is defined by the total number of successfully delivered packets to the sink nodes per transmission. In the simulations, the number of relay nodes and the number of sink nodes are set to 50 and 50, respectively. It is observed in Figure 5 that Algorithm E performs the best in terms of throughput. It has throughput gain of 10-15% over the modified OPNC algorithm (Algorithm B) and 13-25% over Algorithm A in the SNR range between 20 and 25 dB. The average throughput difference is relatively large in the medium SNR range, but it gets negligible in the low and the high SNR regions. In the low SNR regime, the probability of error at a sink node increases. The error increases retransmission from the relay nodes. This phenomenon is believed to be almost independent of the type of the timer algorithm we use. This explains why there is little difference between the 5 algorithms in the low SNR regime. In the high SNR regime, most of the packets

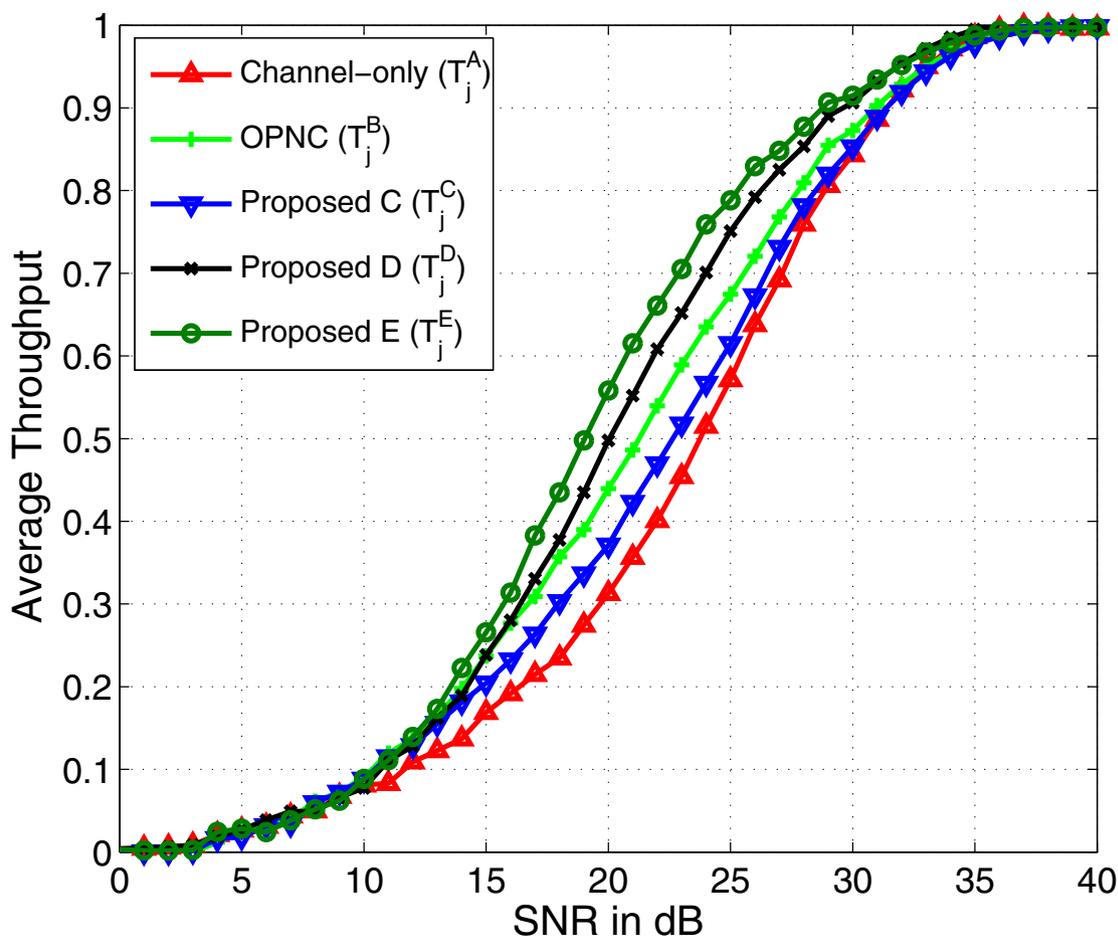


Figure 5 Comparison in terms of average system throughput.

tend to be decoded successfully at the sink nodes in the source transmission (broadcast) phase. It means that the contribution of the retransmission phase decreases as the SNR increases, which also explains why there is little difference between the 5 algorithms in the high SNR regime.

## 5 Conclusion

In this paper, we proposed a new opportunistic wireless network coding combined with distributed relay selection. By taking advantage of opportunistic listening capability of wireless networks, several feedback-based retransmission schemes are proposed. From the simulation results, it was shown that the algorithm based on the minimum channel gain and the sum rate has the best performance in terms of average number of transmissions and system throughput. It was also observed that the proposed relay selection scheme performs better than the conventional schemes especially in the medium SNR regime. It appears that the proposed approach is promising in that it is a practical wireless network coding scheme with improved throughput.

## Acknowledgements

This research was supported in part by Basic Science Research Programs (KRF-2008-314-D00287, 2010-0013397), Mid-career Researcher Program (2010-0027155) through the NRF funded by the MEST, Seoul R&BD Program (JP091007, 0423-20090051), the KETEP grant (2011T100100151), the INMAC, and BK21.

## Competing interests

The authors declare that they have no competing interests.

Received: 14 July 2011 Accepted: 6 December 2011

Published: 6 December 2011

## References

1. D Tuninetti, C Fragouli, in *Processing Along the Way: Forwarding vs Coding*, ISITA 2004, Parma, Italy, (Oct 2004).
2. R Ahlswede, N Cai, S-YR Li, RW Yeung, Network Information Flow. *IEEE Trans Inf Theory* **46**(4), 1204–1216 (2000). doi:10.1109/18.850663
3. M Charikar, in *A Argawal on the Advantage of Network Coding for Improving Network Throughput*, IEEE Information Theory Workshop San Antonio, (Oct 2004)
4. M Ghaderi, D Towsley, J Kurose, in *Reliability Gain of Network Coding in Lossy Wireless Networks*, IEEE INFOCOM 2008, (April 2008)
5. S-YR Li, RW Yeung, N Cai, Linear network coding. *IEEE Trans Inf Theory* **49**(2), 371–381 (2003)
6. T Ho, M Medard, J Shi, M Effros, DR Karger, in *On Randomized Network Coding*. Proceedings of Allerton Conference on Communication, Control, and Computing (2003)
7. S Katti, H Rahul, W Hu, D Katabi, M Medard, J Crowcroft, XORs in the air: practical wireless network coding. *IEEE/ACM Trans Networking* **16**(3), 497–510 (2008)
8. Z Fang, M Medard, in *On Analyzing and Improving COPE Performance*. Information Theory and Applications Workshop (Feb 2010)
9. M Xiao, M Skoglund, in *Design of Network Codes for Multiple-User Multiple-Relay Wireless Networks*, ISIT 2009 (Seoul, Korea, 2009)
10. M Xiao, M Skoglund, Multiple-user cooperative communications based on linear network coding. *IEEE Trans Commun.* **58**(12), 3345–3351 (2010)
11. L Lu, M Xiao, LK Rasmussen, in *Relay-Aided Broadcasting with Instantaneously Decodable Binary Network Codes*, ICCCN 2011 (Hawaii, USA, July, 2011)

12. A Bletsas, A Khisti, DP Reed, A Lippman, A simple cooperative diversity method based on network path selection. *IEEE J Sel Areas Commun.* **24**(3), 659–672 (2006)
13. PA Chou, Y Wu, K Jain, in *Practical Network Coding*. Proceedings of Allerton Conference on Communication, Control, and Computing, (Monticello, IL 2003)
14. JG Proakis, M Salehi, *Digital Communication*, 4th edn. (McGrawHill, 2001)

doi:10.1186/1687-1499-2011-196

**Cite this article as:** Kim and Lee: Opportunistic wireless network coding with relay node selection. *EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:196.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)