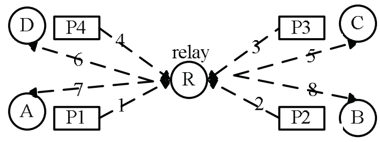
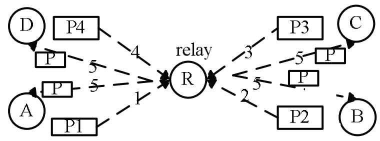
Increase Link Reliability Using Loop Coding in Highly Lossy Wireless Networks

Shuai Ding*, Zhen Chen+, Yixin Jiang*
*Department of Computer Science and Technology, Tsinghua University, Beijing, China
+Research Institute of Information Technology (RIIT), Tsinghua University, Beijing, China
{dingshuai, yxjiang}@csnet1.cs.tsinghua.edu.cn zhenchen@tsinghua.edu.cn

***Abstract*—In this paper, we propose LOCL, a loop coding layer which sits between IP and traditional MAC. Traditional network coding schemes mainly focus on reducing transmissions, while loop coding exploits network coding to reduce link loss rate. We analyze the performance of traditional network coding schemes and loop coding based on a rather general traffic pattern. We find that loop coding is more applicable in highly lossy wireless networks.**



(a) Traditional approach. Packet P1/P2/P3/P4 is intended to be sent to node C/D/A/B respectively. Totally 8 transmissions are required.



$P = P1 \oplus P2 \oplus P3 \oplus P4$

(b) Network coding. First each node sends its packet to R, then R broadcasts packet P. Because each node could overhear all packets except the one sent to it, it is able to get the original packet by XORing packet P with the packets it has overheard.

Fig. 1. Network coding

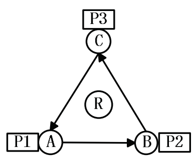


Fig. 2. Loop coding. Node A/B/C sends packet P1/P2/P3 to node B/C/A respectively. Each node could only hear and be heard by the relay.

I. Introduction

Network coding has drawn extensive attentions recently for its promising potential in improving throughput in wireless networks. The idea of network coding was first proposed by Ahlswede et al. [1], and initial research on it mainly focused on multicast traffic. Later on, network coding is applied in unicast traffic and practical systems such as COPE [2]. The application of network coding could be illustrated in Fig. 1, where each node is in the transmission range of all its neighbors except its diagonal counterpart (i.e., node C could overhear every node except A). In traditional approach, each node sends its packet to the relay which forwards its received packets to their intended next hops without modification. Hence, the traditional approach needs 8 transmissions. However, network coding exploits the broadcast nature of wireless medium and allows routers to mix their incoming packets, and it only requires 5 transmissions to finish the packets exchange process.

An intuitive observation of Fig. 1(b) will find that useful network coding requires high link quality and complicated cooperation among network users. Since wireless links have poor quality, for example, in Roofnet half of the operational links have a loss probability higher than 30% [3], it is a great challenge to use network coding in highly lossy wireless networks. Although future routers may have high link quality, there are scenarios where link quality is affected by environment. For instance, the skyscrapers connected by wireless networks may be so far from each other that link quality becomes low. Mathematical analysis of coding probability in lossy wireless networks can be found in Section II.

COPE [2] is the first practical network coding architecture for wireless mesh networks. COPE also shows that network coding can increase the throughput of wireless networks. Le et al. have proposed DCAR [5], which is a coding-aware routing mechanism selecting paths that create more coding opportunities. Han et al. have proposed CAMP [6], which is a coding-aware multi-path routing scheme. Zhang et al. have proposed optimized multi-path network coding [7]. Some papers have proposed routing metric which took coding opportunities into account [8]. Guo et al. have proposed general network coding conditions in multi-hop wireless networks [9]. Sengupta et al. have presented theoretical analysis for the improvements of coding-aware routing [10]. All of the above schemes could be labelled as the "nothing or all" category, which means upon receiving a coded packet, a node can either get the original packet, or get nothing.

In fact, the decoding of a coded packet is a gradual process, i.e., we can get the original packet sent to us after receiving more than one coded packets, which could be labelled as the "part or all" category. Loop coding is a "part or all" scheme, as illustrated in Fig. 2 [4]. From a traditional point of view,

978-1-4577-1139-8/12/$31.00 ©2012 IEEE

TABLE I
NOTATIONS AND DESCRIPTIONS

Notations	Descriptions
N_i	Node i
P_i	Packet i
$S(P_i)$	Packet P_i's sender
$R(P_i)$	Packet P_i's receiver
$O(N_j, P_i, k)$	Node N_j overhears P_i after k transmissions
$Re(P_i, k)$	Relay gets P_i after k transmissions
$Re(P_i)$	Relay gets P_i
$O(N_j, P_i)$	Node N_j overhears P_i

network coding could not take effect in the scenario shown in Fig. 2. However, if the relay node broadcasts $P_1 \oplus P_2$, $P_2 \oplus P_3$, and $P_1 \oplus P_3$, something unexpected will happen. For node A, it can get packet P_3 either by receiving $P_1 \oplus P_3$, or $P_1 \oplus P_2$ and $P_2 \oplus P_3$. Thus, node A can get packet P_3 at a higher probability without increasing transmissions. Detailed analysis of loop coding gain is presented in Section II.

We further exploit the idea of loop coding both theoretically and practically, and propose an efficient scheme to increase link reliability, called LOCL (Loop Coding Layer) which sits between IP and traditional MAC. The main contributions of this paper could be summarized as follows:

- We find the mathematical relation between link loss rate and coding probability, which shows that when loss rate increases, traditional network coding schemes such as COPE have low probability to succeed.
- We demonstrate the absolute existence of loop coding opportunity based on a rather general assumption. As long as each node has a packet to send, there exists a chance to use loop coding.
- We compare the performance of COPE (or COPE based schemes such as DCAR [5]) and loop coding using Generalized Stochastic Petri Nets [11]and provide theoretical analysis.
- We propose a practical architecture to use loop coding in real wireless networks, and further present the important techniques to implement it. Loop coding improves link reliability by extracting an original packet from multiple coded packets, which makes its implementation different from previous schemes.

The rest of this paper is organized as follows. In Section II, we analyze the challenges of using network coding in highly lossy wireless networks. In Section III, we present our LOCL scheme in details, followed by performance evaluation in Section IV. Finally we conclude the paper in Section V. The notations used in this paper are listed in Table I.

II. Challenges In Lossy Wireless Networks

As illustrated in Fig. 1(b), COPE requires high link quality and complicated cooperation among network participants. In this section, we present a general network traffic pattern and build the relation between coding probability and link loss rate for both COPE and loop coding.

A. COPE's Dilemma in Lossy Wireless Networks

In COPE, if the relay wants to mix the N incoming packets $P_1, P_2, ..., P_N$ into one coded packet, it must ensure that for each P_i, $R(P_i)$ has overheard all packets except P_i [2]. $R(P_i)$ can get a packet either by overhearing or because it is the sender of the packet. To analyse the relation between coding probability and link loss rate, we present the following assumption.

Assumption 1. Suppose there are N nodes around the relay node R. Every node wants to send a packet to a certain node that is out of its transmission range. Each node has one packet destined to it. All links have the same loss rate p. Transmissions between different pairs of nodes are independent. The retransmission limit for a single packet is β.

Theorem 1. Based on assumption 1, the probability that the relay can mix all the N incoming packets is α^{N^2-2N}, where $\alpha \triangleq 1 - \frac{p(1+p^\beta)}{1+p}$.

Proof: First we consider the probability that packet P_i can be received by the relay. Because the retransmission limit for any packet is β, we have $P(Re(P_i)) = \sum_{k=1}^{\beta} P(Re(P_i,k)) = \sum_{k=1}^{\beta} p^{k-1}(1-p)$, where k denotes how many times packet P_i has been transmitted. Each packet being coded must be received by all nodes except its receiver, thus we need to know the conditioned probability that P_i is successfully overheard by other nodes.

For any $N_j \neq R(P_i)$, we have

$$P(O(N_j, P_i)|Re(P_i))$$
$$= \frac{P(O(N_j,P_i),Re(P_i))}{P(Re(P_i))}$$
$$= \frac{\sum_{k=1}^{\beta} P(Re(P_i,k))P(O(N_j,P_i,k)|Re(P_i,k))}{\sum_{k=1}^{\beta} P(Re(P_i,k))}$$
$$= \frac{\sum_{k=1}^{\beta} p^{k-1}(1-p)(1-p^k)}{\sum_{k=1}^{\beta} p^{k-1}(1-p)} = 1 - \frac{p(1+p^\beta)}{1+p}$$

For simplicity, we define $\alpha \triangleq 1 - \frac{p(1+p^\beta)}{1+p}$. After the relay has got N packets, the probability it can mix the N incoming packets into one coded packet is $P_{mix-N} = \prod_{i=1}^{N} P(P_i \text{ is received by all nodes except } R(P_i)|Re(P_i))$ $= \prod_{i=1}^{N} \alpha^{N-2} = \alpha^{N^2-2N}$.

Theorem 1 shows that the probability to encode the incoming packets drops severely when N increases. That is because it is difficult to let a packet be overheard by all nodes except its receiver.

B. Efficiency of Loop Coding in Lossy Wireless Networks

The above analysis shows that COPE and COPE-based schemes are not suitable in highly lossy wireless networks for their critical requirement for traffic pattern and link quality. In this part we will show that a relay can make use of loop coding to increase link reliability without knowing much about its neighbors. The following analysis is based on assumption 1.

To find an opportunity for network coding, a relay needs to buffer the packets it has received for a while. If a relay forwards a packet once receiving it, the packet will certainly not be coded with other packets. Without a doubt, a relay can know the sender and receiver of each packet in its buffer. To define the idea of loop coding clearly, we first define traffic graph as below.

Definition 1 (Traffic graph). A relay can buffer the uncoded or original packets it has received. Traffic graph $G_T = (P, E)$ for the original packets it buffers is defined as follows: $P = \{P_i | P_i \text{ is an original packet}\}$ and $E = \{(P_i, P_j) | R(P_i) = S(P_j)\}$.

Theorem 2. Based on assumption 1 and definition 1, there must exist a circuit in G_T.

Proof: According to the assumption, each node has a packet to send, so there are N nodes and N edges in G_T. According to graph theory, for an undirected connected graph of N nodes, it must have one and only one circuit if it has N edges. If we ignore the directions of the edges in G_T, we get an undirected graph $G_T^{'}$. Any undirected graph consists of a series of connected subgraphs. Suppose $G_T^{'}$ consists of the following connected subgraphs: $G_{Tsub1}^{'}, G_{Tsub2}^{'}, ..., G_{Tsubm}^{'}$. Clearly, we have the following equations:

$$\left|E(G_{Tsub1}^{'})\right| + \left|E(G_{Tsub2}^{'})\right| + ... + \left|E(G_{Tsubm}^{'})\right| = N \quad (1)$$

$$\left|V(G_{Tsub1}^{'})\right| + \left|V(G_{Tsub2}^{'})\right| + ... + \left|V(G_{Tsubm}^{'})\right| = N \quad (2)$$

It is not difficult to prove that at least one subgraph, say $G_{Tsubk}^{'}$ satisfies $\left|V(G_{Tsubk}^{'})\right| \leqslant \left|E(G_{Tsubk}^{'})\right|$. Apparently $G_{Tsubk}^{'}$ is a connected graph with at least one circuit. Accordingly $G_T^{'}$ has a circuit. Furthermore, because each node has only one packet destined to it, G_T has a directed circuit.

Definition 2 (Loop coding). A relay buffers N original packets and its traffic graph is G_T. $G_{Tsub} = (P, E)$ is a subgraph of G_T such that $P = \{P_{k1}, ..., P_{km}\}$ and $E = \{(P_{k1}, P_{k2}), ..., (P_{km}, P_{k1})\}$. The relay can conduct loop coding by broadcasting $P_{k1} \bigoplus P_{k2}, P_{k2} \bigoplus P_{k3}, ..., P_{km} \bigoplus P_{k1}$.

Compared to COPE, if a relay wants to conduct loop coding it doesn't require neighbors to overhear original packets not destined to them. Thus loop coding is more suitable in highly lossy wireless networks.

Theorem 3. If a loop coding scheme has N original packets coded into N coded packets $(P_1 \bigoplus P_2, P_2 \bigoplus P_3, ..., P_N \bigoplus P_1)$, the loss rate of a single original packet P_i will be reduced to $p[1-(1-p)^{N-1}]$ (we refer to this as loop coding loss rate).

Proof: Without generality, suppose N_i is the sender of P_i. Take N_2 as an example, it can get P_1 if it receives $P_1 \bigoplus P_2$, or $P_2 \bigoplus P_3, ..., P_{N-1} \bigoplus P_N, P_N \bigoplus P_1$. P_i can be got by $R(P_i)$ if $R(P_i)$ receives $P_x \bigoplus P_i$ (because (P_i, P_x) is an edge in the loop) or $P_x \bigoplus P_{x+1}, ..., P_N \bigoplus P_1, ..., P_{i-1} \bigoplus P_i$. Therefore, the probability that P_i can't be got by $R(P_i)$ is $p[1-(1-p)^{N-1}]$. The relation between loop coding loss rate and loss rate is depicted in Fig. 3.

III. System Design

We have demonstrated the efficiency of loop coding in highly lossy wireless networks in previous sections. In this section, we will discuss the important techniques used in LOCL.

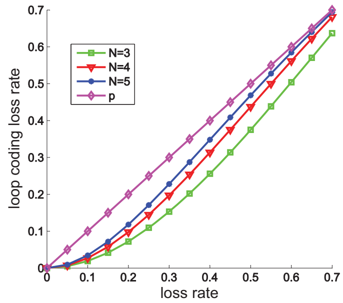


Fig. 3. Loop coding reduces link loss rate

(a) Coded Packets Transmission: The intuition of loop coding is to enhance link quality by creating more chances for packets to be decoded. In 802.11 unicast mode, a packet will be retransmitted until the transmitter gets an ACK or the packet has been retransmitted for a given number of times. If we choose the unicast mode to transmit coded packets, like COPE [2], the potential of loop coding will be wasted because each node could get its wanted packet through reliable unicast link, which achieves link reliability by retransmission. In LOCL, we use 802.11 broadcast mode to broadcast coded packets. Since there are no ACKs and no retransmissions in broadcast mode, link reliability is purely achieved by loop coding and LOCL's retransmission mechanism.

(b) Asynchronous ACK: Coded packets are transmitted in 802.11 broadcast mode, which has no ACKs. How does LOCL know whether the packets have been successfully received? Asynchronous ACK is the only method. After a packet is coded with another packet and broadcast to all neighbors, the transmitter sets up a timer for the packet. If no ACKs are received before the timer expires, the transmitter will retransmit the packet another time, perhaps coded with another packet or no packets.

(c) Loop Coding: Each node maintains a FIFO queue of original packets to be transmitted, which we call the forwarding queue. If a node has some packets that create a loop coding opportunity, it encodes the packets and stores them in the output pool for coded packets. Once a node has access to the wireless medium, it first checks whether there are packets waiting to be sent in the output pool of coded packets. If there are at least one packet in the pool, then the packet arrived earliest is sent; if there are no packets in the pool, then the packet at the head of forwarding queue is picked, and sent directly or coded with another packet.

When a neighbor has more than one original packets to be coded, the packet arrived earliest is selected to avoid packets reordering. As depicted in Fig. 3, loop coding loss rate increases severely as the number of coded packets increases. In practice, we'd better adopt the coding opportunity which has no greater than 4 packets to be coded.

Each node also possesses an input pool for overheard packets, which has one buffer for coded packets and one buffer for original packets. When a node receives a coded packet, it first searches the input pool to see whether the coded packet

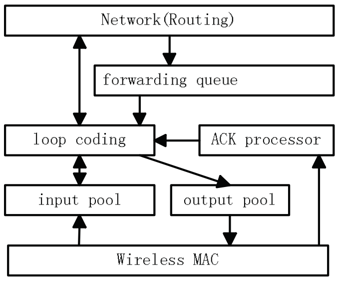


Fig. 4. The architecture of LOCL

could be decoded, if so, the extracted original packet is stored in the input pool or sent to IP layer and the coded packet is discarded; if not, the coded packet is stored in the input pool and waiting for an opportunity to be decoded. Each buffered coded packet has been set a timer, if it couldn't be decoded before the timer expires, it will be discarded.

After receiving an original packet, a node doesn't forward it at once. Instead, it caches the packet for a while. For each packet received, a node can find its sender and receiver. However, if the packet is sent to the node itself, it will be sent to IP layer at once. In this way, a traffic graph is built and loop coding opportunities could be found.

Currently the Internet packet size distribution is bimodal with its peaks at 40B and 1500B (40% and 20% of packets, respectively) [12]. If two packets are coded together, they should have the same size. Otherwise, the smaller packet should be padded with zeros. In LOCL, a node actually needs two kinds of buffers for each of its neighbors, one for small packets (In COPE, the default threshold is 100B), and one for large packets.

The architecture for our LOCL scheme is depicted in Fig. 4. The forwarding queue stores original packets from IP layer to be sent. Overheard packets are stored in input pool. The ACK processor handles asynchronous ACKs to ensure link reliability. The output pool stores coded packets to be sent. The loop coding module is the key component, which controls the encoding and decoding process of a node.

Due to lack of space, we do not provide detailed algorithm for how to find a loop and select the packets to be coded. The scheme we present is distributed and suitable in highly lossy wireless networks. Other important topics include packets reordering and the attack of evil nodes. Because of packets retransmission, the arrival sequence may be different from the forwarding sequence. Evil nodes could broadcast coded packets they create deliberately and make the whole network disabled. These topics are critical to use LOCL in real wireless mesh networks and out of the scope of this paper.

IV. Performance Evaluations

In this section, we compare the performance of COPE and loop coding using Generalized Stochastic Petri Nets (GSP-N) [11]. It is not realistic to analyze the two schemes based on a random topology and random traffic pattern. We refer to the topology and traffic pattern in Fig. 1, where nodes at the diagonal positions exchange packets continuously.

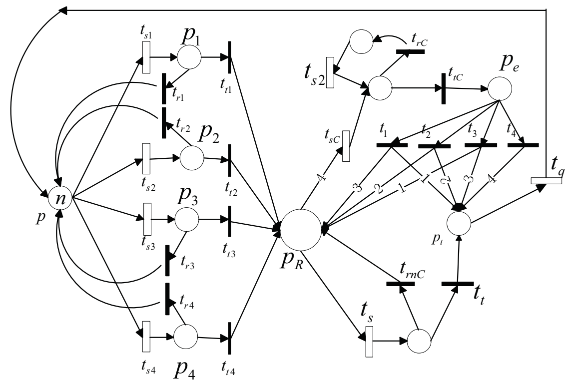


Fig. 5. GSPN for COPE

A formal definition of a Petri Net (PN) is as follows [11]:

$$\begin{cases} PN = (P, T, A, M^{'}) \\ P = \{p_1, p_2, ..., p_n\} \\ T = \{t_1, t_2, ..., t_m\} \\ A \subseteq \{P \times T\} \bigcup \{T \times P\} \\ M^{'} = \{m_1^{'}, m_2^{'}, ..., m_n^{'}\} \end{cases} \tag{3}$$

P and T are the sets of places and transitions respectively. A is the set of arcs between places and transitions. $M^{'}$ is initial marking, which defines the number of tokens of each place when a Petri Net is created. The Stochastic Petri Nets (SPN) are obtained by associating an exponentially distributed firing time for each transition in PN. A formal definition of SPN is :

$$SPN = (P, T, A, M^{'}, R) \tag{4}$$

where P, T, A, $M^{'}$ are as in 3 and $R = (r_1, r_2, ..., r_m)$ is the set of firing rates. GSPN divide transitions into two classes: immediate transitions and timed transitions. GSPN reduce state space by allowing immediate transitions to fire in no time. GSPN were successfully used in the performance analysis of systems whose characteristics include concurrency and synchronization.

The GSPN for COPE is illustrated in Fig. 5. In Fig. 5, p represents the packets waiting to be sent, p_1, p_2, p_3 and p_4 represent the four nodes respectively, p_R represents the relay. t_{s1} represents the sending rate of p_1. When p_1 sends a packet, the packet can either be got by p_R through t_{t1}, or be retransmitted by p_1 through t_{r1}. The situation is similar for p_2, p_3 and p_4. When the relay has more than 4 packets in its output queue, it can either broadcast the coded packet at a probability of α^{N^2-2N} through t_{sC}, or broadcast a non-coded packet through t_s. COPE uses the pseudo broadcast technique to transmit coded packets, which can ensure that the coded packet's receiver can get the packet. As a result, after a coded packet is got by its receiver, other nodes can either overhear the coded packet and get its original packet or remain ignorant of the coded packet.

TABLE II
SIMULATION PARAMETERS FOR COPE(p=0.6)

Parameters	Value
$t_{s1}, t_{s2}, t_{s3}, t_{s4}$	0.25
$t_{r1}, t_{r2}, t_{r3}, t_{r4}, t_{rnc}, t_{rc}$	6000
$t_{t1}, t_{t2}, t_{t3}, t_{t4}, t_{tc}, t_t$	4000
t_{sC}	0.0077
t_s	1
t_q	100000
n	10
β	3

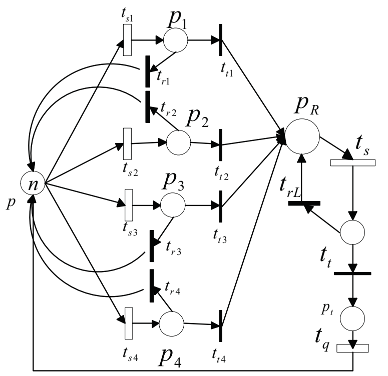


Fig. 6. GSPN for Loop Coding

In Fig. 5, place p_e represents the packets being successfully transmitted. t_q is used to form a closed Petri Net, since it's difficult to run computation for open Petri Nets. This should have little effect on the simulation result, because the rate of t_q is set far bigger than that of any other timed transition. t_4 represents that all of the four original packets are successfully decoded by their receivers, and t_3 indicates that three of the four original packets have been decoded. t_2 and t_1 have similar meaning. The value of some of the parameters is dependent on the link loss rate p. Due to lack of space, we don't provide detailed computation process for those parameters. The value of simulation parameters in Fig. 5 when p is 0.6 is listed in Table II.

The GSPN for loop coding is depicted in Fig. 6. As we demonstrated in Section II, coding opportunities always exist according to our assumption. Consequently, place p_R transmits packets at a lower loss rate. Similar to Fig. 5, t_{s1}, t_{s2}, t_{s3} and t_{s4} denote the sending rate of p_1, p_2, p_3 and p_4 respectively. The relay is represented by p_R. The value for the parameters can be got in the same way as in Fig. 5.

To enable places p_1, p_2, p_3, p_4 to send packets continuously, n should be relatively large and the rates of the transitions should be relatively small. We use SPNP [13] to analyze the models in Fig. 5 and Fig. 6. We mainly focus on the throughput of the two schemes and find that loop coding is more suitable when loss rate is between 0.35 and 0.6. The comparison of throughput for the two schemes is depicted in Fig. 7. Simulation results show that loop coding is more efficient in highly lossy wireless networks.

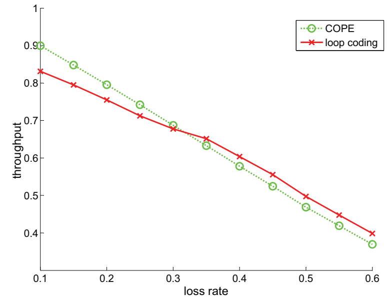


Fig. 7. Throughput of the two schemes all decrease as loss rate increases, while loop coding has higher throughput when loss rate is above 0.35

V. Conclusion

We propose LOCL, a loop coding layer that sits between IP and traditional MAC in wireless mesh networks. Coding schemes, such as COPE, exploit the maximal gain of a single transmission. Loop coding improves link reliability by "part or all" decoding pattern. We demonstrate the opportunities for loop coding based on a rather general assumption. Evaluation by GSPN for a common topology shows that loop coding overwhelms COPE when loss rate is between 0.35 and 0.6.

Acknowledgment

This work is supported by the National Grand Fundamental Research 973 Program of China (No. 2010CB328105, No. 2011CB302703, and No. 2009CB320504), the China NSFC (No. 60932003, No. 60970101, No. 60872055, No. 61020106002, and No. 60834004), and the NSFC A3 Program (No. 61161140320).

References

[1] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung. Network Information Flow. IEEE Trans. on Information Theory, 2000.

[2] S. Katti, H. Rahul, W. Hu, D. Katabi, M. M'edard and J. Crowcroft. XORs in the Air: Practical Wireless Network Coding. In Proc. ACM SIGCOMM'06.2006.

[3] S. Chachulski, M. Jennings, S. Katti, and D. Katabi. Trading Structure for Randomness in Wireless Opportunistic Routing. In Proc. ACM SIGCOMM, Kyoto, Japan, 2007.

[4] Q. Dong, J. Wu, W. Hu, and J. Crowcroft. Practical Network Coding in Wireless Networks. In Proc. ACM MobiCom'07, Sep. 9-14, 2007.

[5] J. Le, J.C.S. Lui and D.M. Chiu. DCAR: Distributed Coding-Aware Routing in Wireless Networks. In Proc. IEEE ICDCS, PP. 462-469, 2008.

[6] S. Han, Z. Zhong, H. Li, G. Chen, E. Chan, and A. K. Mok. Coding-Aware Multi-path Routing in Multi-Hop Wireless Networks. Performance, In Proc. IEEE INFOCOM, 2008.

[7] X. Zhang and B. Li. Optimized Multipath Network Coding in Lossy Wireless Networks. In Proc. IEEE ICDCS, 2008.

[8] Y. Wu, S.M. Das, and R. Chandra. Routing with a Markovian Metric to Promote Local Mixing. MSR Technical Report, Nov., 2006.

[9] B. Guo, H. Li, C. Zhou, and Y. Cheng. General Network Coding Conditions in Multi-hop Wireless Networks. In Proc. IEEE ICC, 2010.

[10] S. Sengupta, S. Rayanchu, S. Banerjee. An Analysis of Wireless Network Coding for Unicast Sessions: The Case for Coding-Aware Routing. In Proc. of IEEE INFOCOM, March 2007, pp. 1028-1036.

[11] M. Ajmone Marsan, G. Balbo, and G. Conte. A Class of Generalized Stochastic Petri Nets for the Performance Evaluation of Multiprocessor Systems. ACM Trans. Comput. Syst., vol. 2, pp. 93-122, May 1984.

[12] Internet Packet Size Distribution. http://www.postel.org/pipermail/end2end-interest/2005-October/005390.html

[13] SPNP. http://people.ee.duke.edu/~kst/software_packages.html