Unsaturated Throughput Analysis of Physical-Layer Network Coding Based on IEEE 802.11 Distributed Coordination Function

Shijun Lin, *Member, IEEE,* and Liqun Fu, *Member, IEEE*

***Abstract*—In this paper, we investigate the throughput performance of physical-layer network coding (PNC) under the IEEE 802.11 distributed coordination function (DCF). We consider the wireless network that two client groups communicate with each other across one relay node, and focus on the unsaturated network case. The difficulty in modeling the relay systems under the IEEE 802.11 DCF is that the minimum contention window sizes of the client nodes and the relay node may be different, which makes the traditional throughput analysis methods for the non-relay wireless networks inapplicable. Fortunately, we find that the relay system can be decomposed into four parts and respectively modeled. Analytical results show that the throughput gain of PNC scheme is heavily affected by the probability that a transmitted network-coding (NC) packet contains the information of two packets. The implication is that the throughput benefit of PNC is more significant for bidirectional isochronous traffic with rate requirements. We further derive an approximate closed-form solution of the optimal transmission probability of client nodes that maximizes the PNC network throughput. We validate our analytical model through extensive simulations and discuss the relationship between the PNC network throughput and other system parameters, such as the minimum contention window sizes of both the client nodes and the relay node.**

***Index Terms*—Physical-layer network coding, IEEE 802.11 DCF, unsaturated throughput analysis.**

I. Introduction

DUE to the broadcast nature of wireless media, wireless links operated on the same channel may cause interference to each other. The concept of physical-layer network coding (PNC), which makes good use of the interference, has great potential to improve the throughput of wireless networks [1], [2]. For example, PNC can improve the throughput of the simple two-way relay channel (TWRC) by 100% [1]. However, what is less understood is the throughput gain of PNC when it is applied to a general wireless network.

Manuscript received November 12, 2012; revised March 26 and June 19, 2013; accepted August 18, 2013. The associate editor coordinating the review of this paper and approving it for publication was A. Abouzeid.

This work was partially supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (No.20110121120019), the Natural Science Foundation of Fujian Province of China (No.2011J01372 and No.2012J01292), the AoE grant E-02/08 established under the University Grants Committee of the Hong Kong Special Administrative Region, China, and the Natural Science Foundation of China (No. 61201195).

S. Lin is with the Department of Communication Engineering, Xiamen University, Xiamen, Fujian, China (e-mail: linsj@xmu.edu.cn).

L. Fu is with the Institute of Network Coding, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong (e-mail: lqfu@inc.cuhk.edu.hk).

Digital Object Identifier 10.1109/TWC.2013.092513.121778

The carrier sense multiple access (CSMA) protocol is the most widely used distributed media access control (MAC) protocol in current wireless networks. In this paper, we focus on investigating the throughput performance of the relay system with PNC under the IEEE 802.11 distributed coordination function (DCF). In particular, we are interested in the unsaturated network case. The unsaturated network is more practical, as in a real wireless network, nodes do not always have packets waiting for being sent. However, both the modeling and the theoretical analysis of the unsaturated network are more complicated. Another difficulty in modeling the relay systems under the IEEE 802.11 DCF is that the minimum contention window sizes of the client nodes and the relay node may be different. This makes the throughput analysis methods for the traditional non-relay wireless networks inapplicable.

The main contributions of this paper are summarized as follows:

1) We derive the analytical unsaturated network throughput results of the PNC scheme, the traditional non-physical-layer network-coding (NC) scheme, and the non-network-coding (NNC) scheme under IEEE 802.11 DCF. What's more, we find that for all schemes, the relay system will be stable when the sending buffers of the client nodes and the relay node are not always non-empty, which can be achieved by controlling the packet generation probability and setting appropriate minimum contention window sizes of the client nodes and the relay node.
2) We show that the throughput gain of PNC scheme is heavily affected by the balance factor, α_{PNC}, which represents the probability that when one of the client nodes occupies the channel and sends a packet to its destination client node, the destination client node also has a packet to be sent to the source client node if the sending buffer of the destination client node is not empty. Compared with the non-physical-layer NC scheme and the NNC schemes, the throughput gain of PNC scheme becomes more significant as α_{PNC} increases. The implication of these results is that the throughput benefit of PNC is more significant for bidirectional isochronous traffic with rate requirements.
3) We optimize the network throughput of PNC scheme in terms of the transmission probability of client nodes k_c, and derive an approximate closed-form optimal solution of k_c. Furthermore, we discuss the relationship

between the PNC network throughput and other system parameters, such as the minimum contention window sizes of both the client nodes and the relay node. The results show that first, to achieve a better throughput performance, the minimum contention window size of the client nodes should be self-adaptive according to the number of client nodes in the system; second, the minimum contention window size of the relay node has little effect on the system throughput, and thus can be set as small as possible to make the system stable.

A. Related work

Network coding is a promising technique to improve the capacity of wired or wireless networks [3]. Originally, NC operations are applied at high-layer (not at physical layer), and three time slots are needed for exchanging two packets between two client nodes across a relay node. We refer to this scheme as high-layer network-coding (HNC) scheme. Considering that two packets encountering in the air can be seen as a natural way of network coding, two new techniques, namely, PNC [1], [2] and analog network coding (ANC) [4], [5] are proposed to further improve the wireless network throughput. In PNC and ANC schemes, only two time slots are needed when exchanging two packets between two client nodes across a relay node.

In the literature, most studies of PNC are focused on the physical layer, and within a simple TWRC network, i.e. [6], [7]. When the PNC scheme is applied to a general network, the transmission coordination of PNC scheme with MAC protocols should be considered. In [8] and [9], the authors proposed a distributed MAC protocol for PNC system which can be seen as an extension of IEEE 802.11 DCF. In [10], Argyriou proposed a MAC protocol for PNC scheme, which makes two independent packet transmissions interfere in a controlled and cooperative manner. Recently, Cocco et al. presented two new schemes to solve the collision problem in slotted ALOHA networks based on multi-user PNC scheme [11]. The coordination of MAC protocols with ANC scheme was investigated in [12]–[15]. However, to the best of our knowledge, none of the existing works analyzes the throughput performance of PNC scheme in a general network coordinated with MAC protocols.

The throughput performance analysis of IEEE 802.11 DCF is mostly on traditional NNC network [16]–[24]. In [16], Bianchi first considered a multistage exponential backoff window, and proposed a bi-dimensional discrete-time Markov model for the saturated single-hop network. And henceforth, Bianchi' model was widely used and extended. In [17], [18], Bianchi' model was extended to the throughput analysis of the unsaturated network by introducing an additional idle state. While in [19], [20] and [21], [22], Bianchi' model was extended by considering the retransmission limit and adding the back-off counter freezing probability, respectively. In [23], [24], the throughput performance of traditional IEEE 802.11 multi-hop networks was analyzed.

Recently, several works (e.g. [25]–[27]) discussed the performance of the HNC scheme coordinated with several distributed MAC protocols in a relay network with one relay node and two client groups. In [25], Umehara et al. analyzed the throughput and delay of the relay system employing HNC scheme and slotted ALOHA protocol under two representative unbalanced traffic cases. In [26], Umehara et al. extended the analysis to a general single-relay multi-user network, and discussed the achievable throughput region. And in [27], the performance of HNC scheme coordinated with an improved slotted ALOHA protocol was analyzed. In this paper, we propose analytical models for the throughput performance of PNC, HNC and NNC schemes coordinated with the widely used IEEE 802.11 DCF based on the relay network with one relay node and two client groups.

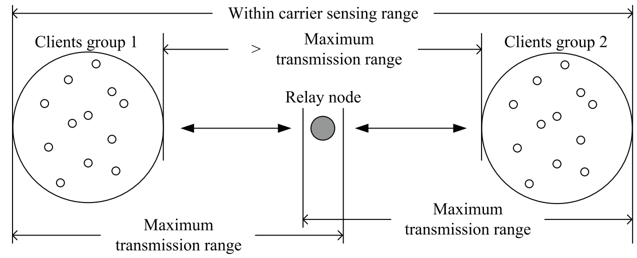


Fig. 1. The two-hop wireless relay network.

The rest of this paper is organized as follows. Section II presents the network model and discusses the transmission processes of the NNC, HNC and PNC schemes under the 802.11 DCF. In section III, we develop analytical models for the unsaturated network throughput of the NNC, HNC and PNC schemes. In section IV, we derive an approximate closed-form optimal solution for the transmission probability of client nodes k_c that maximizes the PNC network throughput. In section V, we carry out extensive simulations to validate our model, discuss the relationship between the PNC network throughput and system parameters, and compare the throughput of the NNC, HNC and PNC schemes. And section VI concludes this paper.

II. System Description

A. Network Model

In this paper, a two-hop wireless system is considered, as shown in Fig. 1. Two client groups each with $u/2$ client nodes communicate with each other across a relay node. All the nodes contend for the channel according to the IEEE 802.11 DCF and work in half-duplex mode. What's more, we only consider the bidirectional traffic across a relay node. That is, first, the relay node does not generate traffic; second, there is no traffic within one client group and all the traffic generated by the client nodes in one group will be transmitted to the client nodes in the other group. We further make the following assumptions: 1) The relay node is within the transmission range of all the client nodes, but the client nodes in different groups are out of the transmission range of each other. One possible network scenario is that several laptops in one room communicate with several other laptops in another room across a relay. 2) The client nodes in the same group are within the transmission range of each other, thus, when a client node transmits packets, the other client nodes in the same group can perform "opportunistic listening"[1] if necessary; 3) All the nodes are in the carrier-sensing range of each other, the channel condition is ideal, and there is no hidden node (HN). In fact, HN problem can be avoided when "RS (Re-Start) mode" is adopted, and the network has a sufficiently large carrier sensing range [28]. In this paper, we focus on the throughput analysis of HN free network. When HN exists, transmission failures will occur during the time that a node successfully occupies the channel. This will result in packet retransmissions in CSMA networks. Thus, the analysis of the network throughput could be more complicated. 4) There is perfect synchronization for PNC system to guarantee that the relay node can extract an NC packet from the superimposed electromagnetic (EM) waves. In practice, high-precision synchronization can be implemented using Locata synchronization technology [29]. Moreover, some technologies, such as Orthogonal Frequency Division Multiplexing (OFDM) [30], can be used so that the synchronization requirement can be relaxed.

B. The NNC, HNC and PNC Transmissions Coordinated with 802.11 DCF

In the 802.11 DCF, a node monitors the channel activity when there is a new packet to transmit. If the channel is idle for a distributed inter frame space (DIFS), the node sends a packet. Otherwise, the node waits until the channel is idle for a DIFS, and then after a random backoff time, the node sends a packet. What's more, a node will wait for a random backoff time after a successful transmission or a collision. In addition, in the 802.11 DCF, an exponential backoff scheme is adopted. That is, at the first transmission attempt of a packet, the backoff time is uniformly chosen in the range $(0, CW_{\min} - 1)$, where $CW_{\min}$ is the minimum size of the contention window. After each failed transmission, the size of the contention window is double until a maximum value of $2^{\mathrm{m}} \cdot CW_{\min}$, where m is called the maximum backoff stage.

For easy coordination with the PNC scheme, we consider the Request To Send (RTS)/Clear To Send (CTS) mechanism in this paper. In the RTS/CTS mechanism, the RTS and CTS packets contain the information of the length of the data packet to be transmitted, which can be read by the other listening nodes to update their network allocation vectors (NAV)[2] [16]. In the following, we describe the detailed RTS/CTS mechanisms based on NNC, HNC and PNC schemes. And in the description, we assume that the propagation delay between any two nodes is the same and equals δ.

In NNC scheme, when any node (a client node or the relay node) successfully occupies the channel, the transmission process is the same, as shown in Fig. 2a. First, when node 1 wants to send some data to node 2, it sends an RTS packet to node 2. Second, if node 2 can receive the RTS packet of node 1 correctly, it sends a CTS packet back to node 1 after a period of time called short inter frame space (SIFS). Third, when receiving the CTS packet, node 1 sends a data packet to node 2 after a SIFS delay. Fourth, when node 2 receives the data packet correctly, it sends back an acknowledgement (ACK) packet after a SIFS delay for confirmation.

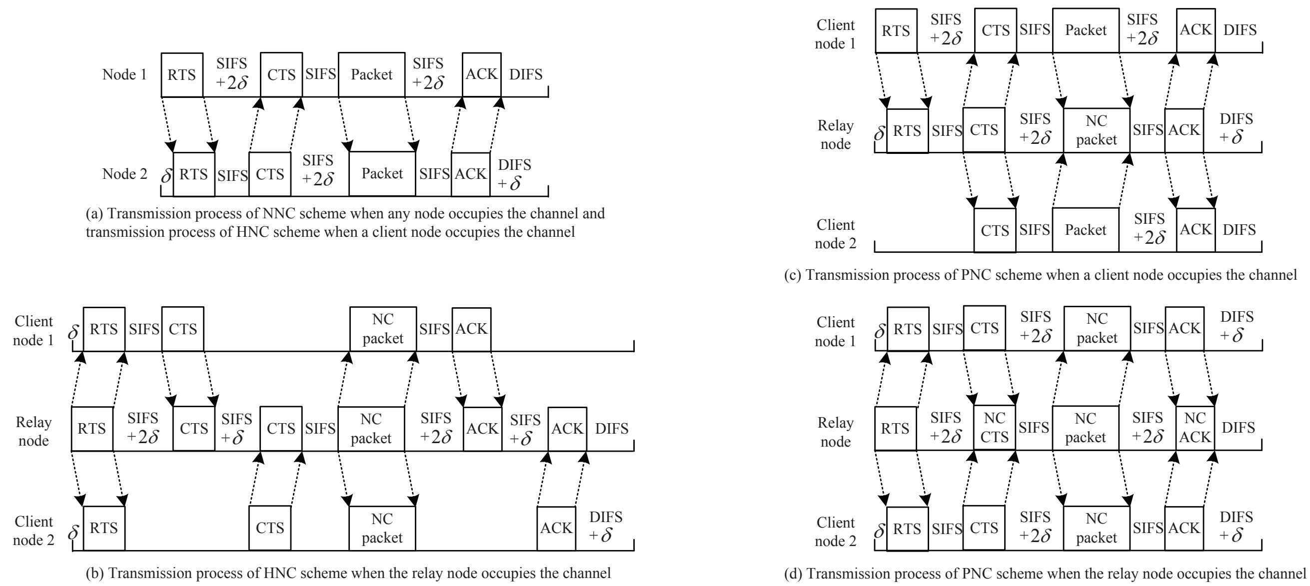
In HNC scheme, the transmission process when a client node successfully occupies the channel is the same as that in NNC scheme, as described above. And, when a client node transmits a packet, the other client nodes in the same group perform "opportunistic listening", which makes the client nodes can decode all the NC packets from the relay node if there are no significant reception errors. In the transmission process when the relay node occupies the channel, the CTS and ACK packets from different client nodes should be transmitted at different sub-time-slots to guarantee that the relay node can receive the CTS and ACK packets from different client nodes correctly, as shown in Fig. 2b. The transmission process is as follows: First, the relay node broadcasts an RTS packet to both client node 1 and client node 2. The RTS packet contains the transmission order of the CTS and ACK packets of client node 1 and client node 2. Second, after the RTS packet correctly arrives at both the client nodes, the two client nodes wait for a SIFS delay first, and then send two sequent CTS packets to the relay node with a SIFS time interval respectively according to the order given in the RTS packet. Third, after the CTS packets are correctly received, the relay node waits for a SIFS delay first and then sends an NC packet to both client node 1 and client node 2. Fourth, after the NC packet is correctly received by both the client nodes for a SIFS time, the two client nodes respectively send two sequent ACK packets to the relay node according to the order given in the RTS packet for confirmation.

In PNC scheme, the transmission processes of a client node and the relay node are different when they successfully occupy the channel. Fig. 2c shows the transmission process of the PNC scheme when a client node occupies the channel. When client node 1 wants to send a data packet to client node 2, it sends an RTS packet with the information of the client node 2 to the relay node. If the relay node can receive the RTS packet correctly, it broadcasts a CTS packet to both client node 1 and client node 2 after a SIFS delay. When correctly receiving the CTS packet, each client node waits for a SIFS delay first, and then sends a data packet to the relay node. In the case that client node 2 does not have a data packet waiting for being sent to the client node 1 when it receives the CTS packet from the relay node, client node 2 sends an empty packet. Under the assumption of perfect synchronization, the relay node can deduce an NC data packet from the received superimposed EM waves, and store it in its sending buffer. And, after a SIFS delay, the relay node broadcasts an ACK packet to both the client nodes for confirmation.

Fig. 2d shows the transmission process of the PNC scheme when the relay node occupies the channel. When there is an NC data packet in the sending buffer of the relay node waiting for being sent to client node 1 and client node 2, the relay node broadcasts an RTS packet to the two client nodes. If both the client nodes can receive the RTS packet correctly, each of them sends a CTS packet to the relay node after a SIFS delay. Under the assumption of perfect synchronization, the relay node can deduce an NC CTS packet. If the NC CTS packet is correct, the relay node waits for a SIFS delay and then broadcasts

[1] "Opportunistic listening" means that when a client node in one group sends a packet to the relay node, all the other nodes in the same group can also receive the packet.

[2] In 802.11 protocol, the NAV represents the number of microseconds the sending node intends to hold the channel busy.



(a) Transmission process of NNC scheme when any node occupies the channel and transmission process of HNC scheme when a client node occupies the channel

(b) Transmission process of HNC scheme when the relay node occupies the channel

(c) Transmission process of PNC scheme when a client node occupies the channel

(d) Transmission process of PNC scheme when the relay node occupies the channel

Fig. 2. Successful transmission processes of all schemes.

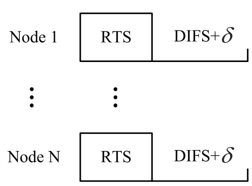


Fig. 3. Transmission process when collision occurs.

an NC data packet to both the client nodes. After correctly receiving the NC data packet, each client node sends back an ACK packet to the relay node after a SIFS delay. By checking the overlapped ACK packets, the relay node knows whether the NC data packet is correctly received or not.

When two or more nodes send RTS packets at the same time, collision occurs and no node can successfully occupy the channel. In this case, the transmission processes of all schemes are the same, which is illustrated in Fig. 3.

III. Throughput Analysis

In this section, we model the relay system based on the NNC, HNC and PNC schemes, and analyze their unsaturated network throughput performances. Since there are two kinds of nodes in the relay system, the whole system will be separated into four parts, and be respectively modeled. We make the following assumptions and approximations in the following discussions.

We assume that in each time slot, each client node generates a packet with a probability g, and the sending buffer of each client node is big enough and thus never full. And when g is given, we make the following approximations:

1) The collision probability that a node transmits is independent and constant, regardless of the number of failed transmissions [16], and that of each client node is the same. This approximation could be more accurate as $CW_{\min}$ and u become larger, since the collision probability means the probability of collision seen by a transmitted packet on the channel [16].
2) The probability that a node transmits when its sending buffer is not empty is constant, and that of each client node is the same. Indeed, the probability that a node transmits when its sending buffer is not empty would vary over time since the time slots with collisions are shorter than the time slots with successful transmissions. This approximation could be more accurate when the collision probability is smaller.
3) The probability that the sending buffer of a node is not empty is constant, and that of each client node is the same.

All the above approximations have been verified through simulations (see Section V-A). And based on these approximations, the relay system can be modeled by four separate processes: one is the changing process of the number of packets in a client node's sending buffer; one is the channel contention process of the client nodes based on the 802.11 DCF when their sending buffers are not empty; one is the changing process of the number of packets in the relay node's sending buffer; and the last one is the channel contention process of the relay node based on the 802.11 DCF when its sending buffer is not empty. In the following, we analyze the network unsaturated throughput of the three schemes by modeling the above processes.

A. The changing process of the number of packets in a client node's sending buffer

In this subsection, we focus on modeling the changing process of the number of packets in a client node's sending buffer. In NNC and HNC schemes, the number of packets in a client node's sending buffer will keep unchanged when one of the client nodes in the other group occupies the channel. However, in PNC scheme, the number of packets in a client node's sending buffer may change when one of the client

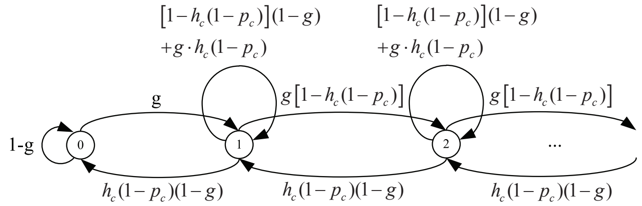


Fig. 4. The Markov chain of the changing process of $S_c(t)$ in NNC and HNC schemes.

nodes in the other group occupies the channel. Therefore, the changing process of the number of packets in a client node's sending buffer in PNC scheme should be separately modeled.

Let p_c, h_c and P_{NE_c} respectively denote the probability that the collision occurs when a client node transmits, the probability that a client node transmits under the condition that its sending buffer is not empty, and the probability that the sending buffer of a client node is not empty. Then, the probability that a client node transmits a packet successfully under the condition that its sending buffer is not empty equals $h_c(1-p_c)$.

Let $S_c(t)$ represent the number of packets in the sending buffer of a client node at discrete time t. In the system, the time is slotted, and t and $t+1$ respectively represent the beginnings of two adjacent time slots. Considering that in all schemes, $S_c(t+1)$ is determined by $S_c(t)$ and the number of packets that arrive at and depart from the sending buffer of the client node during the time $[t, t+1]$, $S_c(t)$ can be modeled with a Markov chain.

1. The changing process of $S_c(t)$ in NNC and HNC schemes:

In NNC and HNC schemes, $S_c(t)$ may change because of the following two events: 1) A new packet is generated by the client node in the current time slot, and the probability of this event equals g. 2) In the current time slot, the client node successfully occupies the channel and sends a packet when its sending buffer is not empty, and the probability of this event equals $h_c(1-p_c)$. Considering that the above two events may both happen during a time slot, the transition diagram of the Markov chain of the changing process of $S_c(t)$ in NNC and HNC schemes is shown in Fig. 4. The one-step transition probabilities are:

$$P\{S_c(t+1)=0\,|S_c(t)=0\}=1-g, \tag{1a}$$
$$P\{S_c(t+1)=1\,|S_c(t)=0\}=g, \tag{1b}$$
$$P\{S_c(t+1)=i\,|S_c(t)=i\}= \\ [1-h_c(1-p_c)]\,(1-g)+g\cdot h_c(1-p_c), i\geq 1, \tag{1c}$$
$$P\{S_c(t+1)=i+1\,|S_c(t)=i\}= \\ g\,[1-h_c(1-p_c)]\,, i\geq 1, \tag{1d}$$
$$P\{S_c(t+1)=i-1\,|S_c(t)=i\}= \\ h_c(1-p_c)(1-g), i\geq 1, \tag{1e}$$
$$P\{S_c(t+1)=j\,|S_c(t)=i\}=0,\ |i-j|\geq 2. \tag{1f}$$

The explanations of equation (1) are as follows:

1) Equation (1a) and (1b) represent that when the sending buffer of a client node is empty, only the event that a new packet is generated may happen.
2) Equation (1c) states that in NNC and HNC schemes, when the sending buffer of a client node is not empty, the number of packets in its sending buffer may keep unchanged under the following two cases: one is that no packet is generated and no packet is successfully transmitted, and the probability of this case equals $[1-h_c(1-p_c)]\,(1-g)$; the other case is that there is a new packet generated and a packet transmitted successfully, and the probability of this case equals $g\cdot h_c(1-p_c)$.
3) Equation (1d) accounts the fact that in NNC and HNC schemes, when the sending buffer of a client node is not empty, the number of packets in the sending buffer may increase by one if there is a packet generated but no packet is successfully transmitted.
4) Equation (1e) states that in NNC and HNC schemes, in the case that the sending buffer of a client node is not empty, the number of packets in the sending buffer may decrease by one if there is a packet successfully transmitted but no packet is generated.
5) Equation (1f) states the assumption that only one packet may be generated in each time slot and the fact that only one packet is transmitted in a successful transmission.

We next calculate the stationary probabilities.

Let $\pi_c(i)$ be the stationary probability of $S_c(t)$, which equals $\lim\limits_{t\to\infty} P\left\{S_c(t)=i\right\}$. According to the balance equations of steady states $\sum\limits_{i=0}^{+\infty}\pi_c(i)\cdot P\{S_c(t+1)=j\,|S_c(t)=i\}=\pi_c(j)$, $(j\geq 0)$, $\pi_c(i)$ in NNC and HNC schemes can be calculated by:

$$\begin{cases} \pi_c(1)=\frac{g}{h_c(1-p_c)(1-g)}\cdot\pi_c(0), \\ \pi_c(i)=\left[\frac{g-g\cdot h_c(1-p_c)}{h_c(1-p_c)(1-g)}\right]^{i-1}\cdot\pi_c(1), i\geq 2. \end{cases}$$

Then, if $\frac{g-g\cdot h_c(1-p_c)}{h_c(1-p_c)(1-g)}$ is less than one, the stationary probabilities exit. According to the equation $\sum\limits_{i=0}^{+\infty}\pi_c(i)=1$, $\pi_c(0)$ in NNC and HNC schemes equals:

$$\pi_c(0)=\frac{h_c(1-p_c)-g}{h_c(1-p_c)}. \tag{2}$$

Then, in NNC and HNC schemes, the probability that the sending buffer of a client node is not empty P_{NE_c} equals:

$$P_{NE\_c}=1-\pi_c(0)=\frac{g}{h_c(1-p_c)}. \tag{3}$$

In equation (3), g can be seen as the packet arrival probability, which is a given system parameter; while $h_c(1-p_c)$ can be seen as the packet departure probability when the sending buffer of the considered client node is not empty. Equation (3) shows that the changing process of the number of packets in a client node's sending buffer, and the channel contention process of the client nodes based on the 802.11 DCF are connected together. In fact, $h_c(1-p_c)$ also represents the system service probability when the sending buffer of a client node is not empty in NNC and HNC schemes. Thus, a bigger $h_c(1-p_c)$ will lead to a bigger system service rate and then a higher system throughput of the NNC and HNC schemes.

2. The changing process of $S_c(t)$ in PNC scheme:

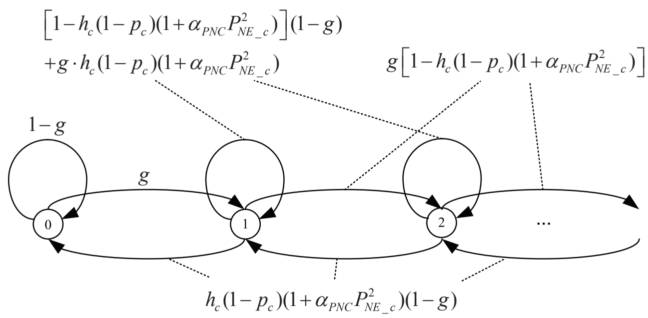


Fig. 5. The Markov chain of the changing process of $S_c(t)$ in PNC scheme.

Next we discuss the model of the changing process of $S_c(t)$ in PNC scheme. Since in PNC scheme, the number of packets in a client node's sending buffer may change when one of the client nodes in the other group occupies the channel, the model of the changing process of $S_c(t)$ in PNC scheme is different from the one in NNC and HNC schemes.

In PNC scheme, $S_c(t)$ of a client node n_c^x may change because of the following three events: 1) In the current time slot, there is a new packet generated by client node n_c^x, and the probability of this event equals g; 2) In the current time slot, client node n_c^x successfully occupies the channel when its sending buffer is not empty, and the probability of this event equals $h_c(1-p_c)$; 3) In the current time slot, when the sending buffer of client node n_c^x is not empty, one of the client nodes in the other group, i.e. client node n_c^y, successfully occupies the channel with client node n_c^x as the destination client node and client node n_c^x has a packet waiting for being sent to client node n_c^y. The third event only happens in PNC scheme, which makes the changing process of $S_c(t)$ in PNC scheme different from the one in the NNC and HNC schemes.

In the following, we calculate the probability of the third event. Since the probability that a client node occupies the channel successfully equals $P_{NE\_c}h_c(1-p_c)$, the probability that the channel is successfully occupied by a node in one client group equals $\frac{u}{2}\cdot P_{NE\_c}h_c(1-p_c)$. Assume that every client node in the other group has the same probability to be the destination client node. Then, the probability that client node n_c^x is the destination client node when the channel is occupied by one of the client nodes in the other group equals $\frac{1}{u/2}\cdot\frac{u}{2}P_{NE\_c}h_c(1-p_c)$. Therefore, when the sending buffer of n_c^x is not empty, the probability that the channel is occupied by one of the client nodes in the other group with n_c^x as the destination client node equals $P_{NE\_c}\cdot\frac{1}{u/2}\frac{u}{2}P_{NE\_c}h_c(1-p_c)$. Let $\alpha_{PNC}\in[0,1]$ be the balance factor in PNC scheme, which represents the probability that when the channel is successfully occupied by a client node, its destination client node has a packet waiting for being sent to the source client node under the condition that the sending buffer of the destination client node is not empty. Then, the probability of the third event equals $\alpha_{PNC}\cdot\left[P_{NE\_c}\frac{1}{u/2}\frac{u}{2}P_{NE\_c}h_c(1-p_c)\right]$, which can be simplified to $\alpha_{PNC}P^2_{NE\_c}h_c(1-p_c)$. Considering that the first and the second events, the first and the third events may both happen in the same time slot, but the second and the third events can not both happen in the same time slot because the channel can only be occupied by one node in a time slot, then, the transition diagram of the Markov chain of the changing process of $S_c(t)$ in PNC scheme is shown in Fig. 5, and the one-step transition probabilities are:

$$P\{S_c(t+1)=0\,|S_c(t)=0\}=1-g, \tag{4a}$$
$$P\{S_c(t+1)=1\,|S_c(t)=0\}=g, \tag{4b}$$
$$\begin{aligned}&P\{S_c(t+1)=i\,|S_c(t)=i\}=\\&\left[1-h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})\right](1-g)\\&+g\cdot h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c}), i\geq 1,\end{aligned} \tag{4c}$$
$$\begin{aligned}&P\{S_c(t+1)=i+1\,|S_c(t)=i\}=\\&g\left[1-h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})\right], i\geq 1,\end{aligned} \tag{4d}$$
$$\begin{aligned}&P\{S_c(t+1)=i-1\,|S_c(t)=i\}=\\&h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})(1-g), i\geq 1,\end{aligned} \tag{4e}$$
$$P\{S_c(t+1)=j\,|S_c(t)=i\}=0,\ |i-j|\geq 2. \tag{4f}$$

The explanations of equation (4) are as follows:

1) Equation (4a) and (4b) state that in PNC scheme, when the sending buffer of a client node is empty, only the first event may happen.
2) Equation (4c) accounts the fact that in PNC scheme, when the sending buffer of a client node is not empty, the number of packets in its sending buffer may keep unchanged under the following three cases: i) both the first and the second events happen; ii) both the first and the third events happen; iii) none of the events happens.
3) Equation (4d) states that in PNC scheme, in the case that the sending buffer of a client node is not empty, the number of packets in its sending buffer may increase by one when the first event happens but neither of the second and the third events happens.
4) Equation (4e) accounts the fact that in PNC scheme, in the case that the sending buffer of a client node is not empty, the number of packets in the sending buffer may decrease by one when the first event does not happen and one of the other two events happens.
5) Equation (4f) states the assumption that only one packet may be generated in each time slot and the fact that only one packet is transmitted in a successful transmission.

In the following, we calculate the stationary probability $\pi_c(i)$ in PNC scheme.

According to the balance equations of the steady states $\sum_{i=0}^{+\infty}\pi_c(i)\cdot P\{S_c(t+1)=j\,|S_c(t)=i\} = \pi_c(j), (j\geq 0)$, $\pi_c(i)$ in PNC scheme can be expressed by:

$$\begin{cases}\pi_c(1)=\frac{g}{h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})(1-g)}\cdot\pi_c(0),\\ \pi_c(i)=\left[\frac{g-g\cdot h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})}{h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})(1-g)}\right]^{i-1}\cdot\pi_c(1), i\geq 2.\end{cases}$$

If $\frac{g-g\cdot h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})}{h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})(1-g)}$ is less than one, $\pi_c(i)$ exits. Since $\sum_{i=0}^{+\infty}\pi_c(i)=1$, $\pi_c(0)$ equals:

$$\pi_c(0)=\frac{h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})-g}{h_c(1-p_c)(1+\alpha_{PNC}P^2_{NE\_c})}. \tag{5}$$

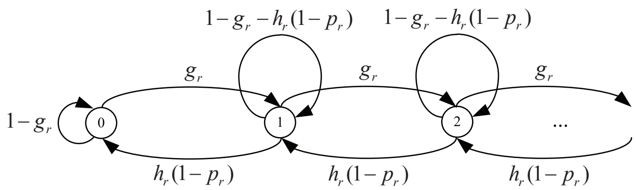


Fig. 6. The Markov chain of the changing process of $S_r(t)$.

Then, the probability that the sending buffer of a client node is not empty P_{NE_c} equals:

$$P_{NE\_c} = 1 - \pi_c(0) = \frac{g}{h_c(1-p_c)(1+\alpha_{PNC}P_{NE\_c}^2)}. \quad (6)$$

In equation (6), $h_c(1-p_c)(1+\alpha_{PNC}P_{NE\_c}^2)$ is the sum of the probabilities that the second and the third events happen, which indeed represents the total packet departure probability or total system service probability when there are packets in the sending buffer of the considered client node. Thus, a bigger $h_c(1-p_c)(1+\alpha_{PNC}P_{NE\_c}^2)$ will lead to a bigger system service rate and then a higher system throughput in PNC scheme. Furthermore, we have the following two observations: 1) The throughput gain of PNC scheme comes from the third event, which does not happen in the other two schemes. 2) The balance factor, α_{PNC}, has a significant impact on the throughput performance of PNC scheme since it directly determines the probability of the third event.

B. The changing process of the number of packets in the relay node's sending buffer

In NNC, HNC, and PNC schemes, the changing processes of the number of packets in the relay node's sending buffer are similar. The only difference is the packet arrival probability. Thus, in the following, we first propose a common model for all schemes, and then discuss the difference of the packet arrival probability in the three schemes.

Let p_r, h_r and P_{NE_r} respectively denote the probability that a transmitted packet of the relay node encounters a collision, the probability that the relay node transmits under the condition that its sending buffer is not empty, and the probability that the sending buffer of the relay node is not empty. Then, the probability that the relay node successfully occupies the channel when its sending buffer is not empty equals $h_r(1-p_r)$. Let g_r be the packet arrival probability of the relay node's sending buffer. Let $S_r(t)$ denote the number of packets in the sending buffer of the relay node at discrete time t. Since $S_r(t+1)$ can be determined by $S_r(t)$ and the number of packets that arrive at and depart from the sending buffer of the relay node during the time $[t, t+1]$, $S_r(t)$ can be modeled with a Markov chain.

In all schemes, $S_r(t)$ will change when one of the following two events occurs: 1) In the current time slot, there is a packet arriving at the sending buffer of the relay node, and the probability of this event equals g_r. 2) In the current time slot, there is a packet successfully transmitted by the relay node, and the probability of this event equals $h_r(1-p_r)$. Considering that the above two events cannot both happen in the same time slot, the transition diagram of the Markov chain of the changing process of $S_r(t)$ is shown in Fig. 6, and the one-step transition probabilities are:

$$P\{S_r(t+1)=0\,|S_r(t)=0\} = 1-g_r, \quad (7a)$$

$$P\{S_r(t+1)=i\,|S_r(t)=i\} = 1-g_r-h_r(1-p_r), i \geq 1, \quad (7b)$$

$$P\{S_r(t+1)=i+1\,|S_r(t)=i\} = g_r, i \geq 0, \quad (7c)$$

$$P\{S_r(t+1)=i-1\,|S_r(t)=i\} = h_r(1-p_r), i \geq 1, \quad (7d)$$

$$P\{S_r(t+1)=j\,|S_r(t)=i\} = 0, \, |i-j| \geq 2. \quad (7e)$$

The explanations of equation (7) are as follows:

1) Equation (7a) states that when the sending buffer of the relay node is empty, only the first event may occur.
2) Equation (7b) represents that in the case of the sending buffer of the relay node being not empty, $S_r(t)$ will keep unchanged when neither of the two events occurs.
3) Equation (7c) accounts the fact that no matter if the sending buffer of the relay node is empty or not, $S_r(t)$ may increase by one when the first event occurs.
4) Equation (7d) states that in the case that the sending buffer of the relay node is not empty, $S_r(t)$ may decrease by one when the second event occurs.
5) Equation (7e) accounts the fact that only one packet is transmitted when a node successfully occupies the channel.

We next calculate the stationary probabilities of $S_r(t)$.

Let $\pi_r(i)$ be the stationary probability of $S_r(t)$, which equals $\lim_{t\to\infty} P\{S_r(t)=i\}$. Then, according to the balance equations of the steady states $\sum_{i=0}^{+\infty} \pi_r(i) \cdot P\{S_r(t+1)=j\,|S_r(t)=i\} = \pi_r(j)$, (j $\geq$ 0), we have:

$$\pi_r(i) = \pi_r(0) \cdot \rho_r^i, (i \geq 0), \quad (8)$$

where ρ_r equals $\frac{g_r}{h_r(1-p_r)}$. When $\rho_r < 1$, $\pi_r(0)$ can then be calculated according to the equation $\sum_{i=0}^{+\infty} \pi_r(i) = 1$. That is, $\pi_r(0)$ equals $1-\rho_r$.

Then, for all schemes, the probability that the relay node's sending buffer is not empty equals:

$$P_{NE\_r} = 1 - \pi_r(0) = \frac{g_r}{h_r(1-p_r)}. \quad (9)$$

In equation (9), g_r is determined by the channel contention process of the client nodes, while $h_r(1-p_r)$ is determined by the channel contention process of the relay node. That is, the changing process of $S_r(t)$, the channel contention process of the client nodes, and the channel contention process of the relay node are connected together through equation (9).

In the following, we discuss the difference of g_r in the three schemes.

In NNC and PNC schemes, when one of the client nodes occupies the channel successfully, the relay node will receive a packet and directly store it in the sending buffer. Thus, the packet arrival probability of the relay node's sending buffer, g_r, equals the probability that one of the client nodes occupies the channel successfully. Therefore, in NNC and PNC schemes, g_r equals:

$$g_r = u \cdot P_{NE\_c} h_c(1-p_c). \quad (10)$$

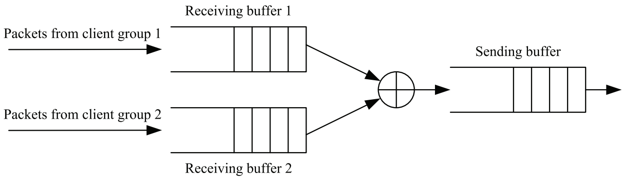


Fig. 7. The packets arrival flow at the relay node in HNC scheme.

However, in HNC scheme, since the network coding operation occurs at the high layer, the packets received from the client nodes are not directly stored in the sending buffer. The packets arrival flow at the relay node is shown in Fig. 7. The packets from the client group 1 and client group 2 are respectively stored in the receiving buffer 1 and receiving buffer 2. And, considering that the packet arrival probabilities of the receiving buffers are the same, the network coding operation is done only when both of the buffers are not empty. After the network coding operation, an NC packet is generated and stored in the sending buffer of the relay node. Thus, the packet arrival probability of the sending buffer of the relay node equals that of one of the receiving buffers in the relay node. That is, in HNC scheme, g_r equals:

$$g_r = \frac{u}{2} \cdot P_{NE\_c} h_c (1 - p_c). \tag{11}$$

C. The relationships among the variables

In this subsection, we discuss the channel contention process of the client nodes and the relay node, and the relationships among variables p_c, h_c, P_{NE_c}, p_r, h_r, and P_{NE_r}.

When the sending buffer of a node is not empty, no matter how many packets there are in the sending buffer, the node will contend for the channel in the same way according to the 802.11 DCF. Thus, according to Bianchi's model [16], the relationships between the probability that a node transmits under the condition that its sending buffer is not empty and the probability that a transmitted packet of a node encounters a collision are given as follows:

$$h_c = \frac{2(1-2p_c)}{(1-2p_c)(W_c+1) + p_c W_c \left[1 - (2p_c)^m\right]}, \tag{12}$$

$$h_r = \frac{2(1-2p_r)}{(1-2p_r)(W_r+1) + p_r W_r \left[1 - (2p_r)^m\right]}, \tag{13}$$

where W_c and W_r are respectively the minimum contention window size of a client node and the relay node, and m is the maximum backoff stage of any node.

What's more, since the collision probability that a node transmits is also the probability that at least one of the other nodes transmits in a slot time, in all schemes, we have:

$$p_c = 1 - (1 - P_{NE\_c} h_c)^{u-1} (1 - P_{NE\_r} h_r), \tag{14}$$

$$p_r = 1 - (1 - P_{NE\_c} h_c)^{u}. \tag{15}$$

Given g, all the variables can be solved by the above equations. In particular, in NNC scheme, the values of p_c, h_c, P_{NE_c}, p_r, h_r, P_{NE_r}, and g_r can be obtained by solving equations (3), (9), (10), (12), (13), (14), and (15); in HNC scheme, the values of p_c, h_c, P_{NE_c}, p_r, h_r, P_{NE_r}, and g_r can be obtained by solving equations (3), (9), (11), (12), (13), (14), and (15); in PNC scheme, the values of p_c, h_c, P_{NE_c}, p_r, h_r, P_{NE_r}, and g_r can be obtained by solving equations (6), (9), (10), (12), (13), (14), and (15).

The stability conditions of the relay system are as follows. In NNC and HNC schemes, the relay system is stable when $\frac{g - g \cdot h_c(1-p_c)}{h_c(1-p_c)(1-g)} < 1$ and $\frac{g_r}{h_r(1-p_r)} < 1$, which can be simplified to $g < h_c(1-p_c)$ (or $P_{NE\_c} < 1$) and $g_r < h_r(1-p_r)$ (or $P_{NE\_r} < 1$). While in PNC scheme, the relay system is stable when $\frac{g - g \cdot h_c(1-p_c)(1+\alpha_{PNC} P_{NE\_c}^2)}{h_c(1-p_c)(1+\alpha_{PNC} P_{NE\_c}^2)(1-g)} < 1$ and $\frac{g_r}{h_r(1-p_r)} < 1$, which can be simplified to $g < h_c(1-p_c)(1+\alpha_{PNC} P_{NE\_c}^2)$ (or $P_{NE\_c} < 1$) and $g_r < h_r(1-p_r)$ (or $P_{NE\_r} < 1$). That is, for all schemes, the relay system is stable when P_{NE_c} and P_{NE_r} are both less than one. The implications of these stable conditions are as follows: 1) The proposed model can only be used to analyze the unsaturated scenario, thus, the packet generation probability g should be small enough to ensure that the system is unsaturated. 2) In order to avoid the queue in the relay node being infinitely long, the relay node should have a large enough opportunity to access the channel, which can be achieved by setting W_c and W_r appropriately. In fact, since there are u client nodes and only one relay node, W_r should be set much smaller than W_c.

D. Throughput calculation

The network throughput, Q_c, can be defined as the ratio of the average time used by the client nodes to successfully transmit the payload information in a slot time to the average length of a slot time.

In the following, we discuss the calculation of Q_c. The probability that the channel is successfully occupied by one of the client nodes, P_{s_c}, and the probability that the channel is successfully occupied by the relay node, P_{s_r}, respectively equal:

$$P_{s\_c} = u \cdot P_{NE\_c} h_c (1 - p_c), \tag{16}$$

$$P_{s\_r} = P_{NE\_r} h_r (1 - p_r). \tag{17}$$

Furthermore, since the probability that no node transmits a packet equals $(1 - P_{NE\_c} h_c)^u (1 - P_{NE\_r} h_r)$, the probability that at least one node transmits a packet, P_{tr}, then equals:

$$P_{tr} = 1 - (1 - P_{NE\_c} h_c)^u (1 - P_{NE\_r} h_r). \tag{18}$$

After obtaining P_{s_c}, P_{s_r}, and P_{tr}, we are ready to calculate the average length of a slot time in the unsaturated case. In the unsaturated case, a node does not transmit in the following two cases: 1) the sending buffer of the node is empty; 2) the backoff counter of the node does not reach zero. In fact, for any node, the sending buffer being empty can be regarded as the value of the backoff counter being infinitely large. Thus, when the channel is idle in a slot time, we can postulate that the backoff counters of all nodes decrease by one and no backoff counter reaches zero. Then, the average length of a slot time can be calculated based on the fact that with probability $(1-P_{tr})$, the slot time is idle; with probability P_{s_c}, the slot time is successfully occupied by a client node; with probability P_{s_r}, the slot time is successfully occupied by the relay node; with probability $(P_{tr} - P_{s\_r} - P_{s\_c})$, a collision occurs. Let T_{AST} be the average length of a slot time in the unsaturated case. Then, T_{AST} can be calculated by:

$$\begin{aligned} T_{AST} = (1-P_{tr})\sigma + P_{s\_c}T_{s\_c} + P_{s\_r}T_{s\_r} \\ +(P_{tr}-P_{s\_r}-P_{s\_c})T_c, \end{aligned} \tag{19}$$

where T_{s_c}, T_{s_r}, and T_c are respectively the average time that a successful transmission of a client node experiences, the average time that a successful transmission of the relay node experiences, and the average time that a collision experiences. In different schemes, the values of T_{s_c}, T_{s_r}, and T_c may be different, which can be calculated according to the analysis in section II-B. The symbol σ is the basic unit of the value of backoff counter described in the 802.11 DCF.

Therefore, the network throughput Q_c can be expressed by:

$$Q_c = \frac{P_{s\_c}T_{L\_c}}{T_{AST}}, \tag{20}$$

where T_{L_c} is the average time used by a client node to transmit the total amount of the payload information of a packet.

T_{L_c} is different in different schemes. In NNC and HNC schemes, when a client node occupies the channel, the total amount of the payload information in the transmitted packet equals the payload size S_{PL}. Then, T_{L_c} in NNC and HNC schemes equals ${S_{PL}}/{R_{link}}$, where R_{link} is the physical link rate.

In PNC scheme, when a client node occupies the channel, the probability that the transmitted packet contains the information of two packets equals $P_{NE\_c}\alpha_{PNC}$, and the probability that the transmitted packet only contains the information of one packet equals $(1 - P_{NE\_c}\alpha_{PNC})$. Then, the average amount of payload information in the transmitted packet equals $[2S_{PL} \cdot P_{NE\_c}\alpha_{PNC} + S_{PL} \cdot (1 - P_{NE\_c}\alpha_{PNC})]$, which can be simplified to $S_{PL}(1 + P_{NE\_c}\alpha_{PNC})$. Therefore, T_{L_c} in PNC scheme equals $\frac{S_{PL}(1+P_{NE\_c}\alpha_{PNC})}{R_{link}}$.

IV. Throughput Maximization of the PNC Relay System

In this section, we focus on the throughput maximization of the PNC relay system. We consider the nearly saturated case, that is, the packet generation probability g is set to its maximum value which makes P_{NE_c} close to one. The reason is that the impact of system parameters on the throughput will be less significant when g is smaller. In particular, in the extreme case that g equals zero, the throughput will equal zero under any system parameters.

In the nearly saturated case, the throughput in PNC scheme, $Q_c|_{P_{NE\_c}\to 1}$, approximately equals:

$$Q_c|_{P_{NE\_c}\to 1} = \frac{{S_{PL}}/{R_{link}} \cdot P_{s\_c}(1+\alpha_{PNC})}{T_{AST}}. \tag{21}$$

Let k_c and k_r respectively denote the transmission probability of a client node and the relay node. Then, $k_c = P_{NE\_c}h_c$ and $k_r = P_{NE\_r}h_r$. Equations (16), (17) and (18) can be simplified.

$$P_{s\_c} = u \cdot k_c(1-k_c)^{u-1}(1-k_r), \tag{22}$$

$$P_{s\_r} = k_r(1-k_c)^u, \tag{23}$$

$$P_{tr} = 1-(1-k_c)^u(1-k_r). \tag{24}$$

According to equations (9), (10), (16) and (17), P_{s_c} equals P_{s_r} in PNC scheme. Then,

$$k_r = \frac{u \cdot k_c}{1-k_c+u\cdot k_c}. \tag{25}$$

Furthermore, according to equations (19), (21), (22), (23), (24) and (25), $Q_c|_{P_{NE\_c}\to 1}$ equals:

$$Q_c|_{P_{NE\_c}\to 1} = \frac{{S_{PL}}/{R_{link}}\cdot(1+\alpha_{PNC})}{\frac{T_c(1-k_c+u\cdot k_c)-(T_c-\sigma)(1-k_c+u\cdot k_c)(1-k_c)^u}{u\cdot k_c(1-k_c)^u}+(T_{s\_c}+T_{s\_r}-\sigma-T_c)}. \tag{26}$$

In equation (26), S_{PL}, R_{link}, α_{PNC}, u, T_{s_c}, T_{s_r}, T_c, and σ are all given system parameters, and only k_c is the control variable. That is, $Q_c|_{P_{NE\_c}\to 1}$ can be determined by k_c.

In the following, we discuss the maximization of $Q_c|_{P_{NE\_c}\to 1}$.

Let Y denote $\frac{u\cdot k_c(1-k_c)^u}{T_c(1-k_c+u\cdot k_c)-(T_c-\sigma)(1-k_c+u\cdot k_c)(1-k_c)^u}$. Then, according to equation (26), $Q_c|_{P_{NE\_c}\to 1}$ is maximized when Y is maximized.

Lemma 1: As k_c increases from 0 to 1, Y first increases to a maximum value and then decreases.

The proof of Lemma 1 is in the Appendix. Based on Lemma 1, the following theorem can be obtained.

Theorem 1: The optimal k_c that maximizes the network throughput approximately equals:

$$k_c = \frac{-(u+1)\sigma+\sqrt{(u+1)^2\sigma^2+4\sigma\left[T_c(u^2-u)+\frac{1}{2}u(u+1)(T_c-\sigma)\right]}}{2T_c(u^2-u)+u(u+1)(T_c-\sigma)}. \tag{27}$$

Proof: According to Lemma 1, there is only one $k_c \in [0,1]$ which makes Y reach its maximum value. The k_c which maximizes Y can be obtained by imposing the derivative of Y with respect to k_c equal 0. After some simplifications, we have:

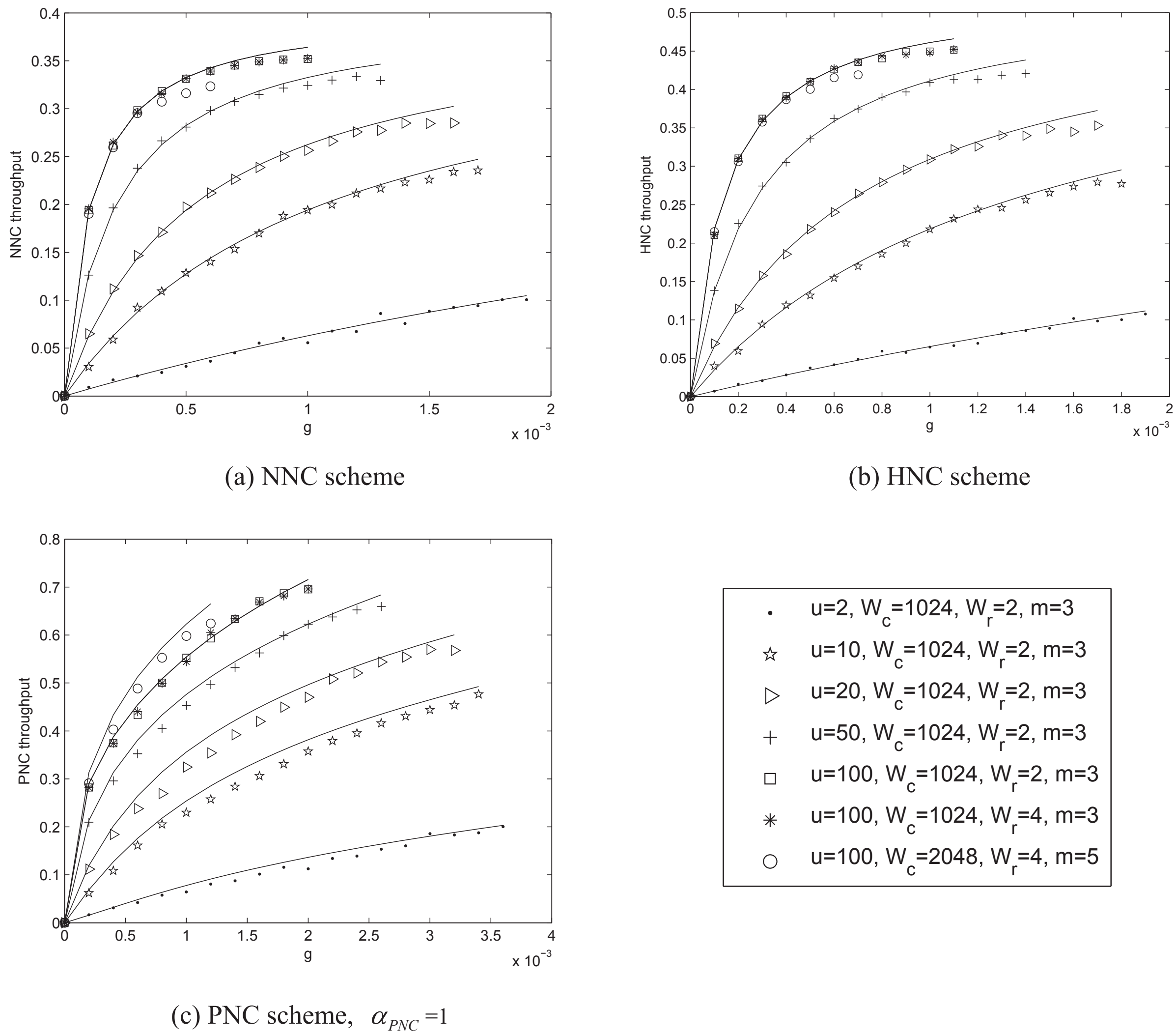
$$T_c\left[(u^2-u)k_c^2+(u+1)k_c-1\right]+(T_c-\sigma)(1-k_c)^{u+1}=0. \tag{28}$$

According to the Taylor series approximation, under the condition $k_c \ll 1$, we have:

$$(1-k_c)^{u+1} \approx 1-(u+1)k_c+\frac{(u+1)u}{2}k_c^2. \tag{29}$$

Then, according to equations (28) and (29), equation (27) can be obtained. ∎

Theorem 1 provides the optimal value of k_c that maximizes the nearly saturated throughput of the PNC relay system. Considering that P_{NE_c} is close to one in the nearly saturated case, the value of k_c is close to the value of h_c. Then, adjusting k_c is equivalent to adjusting h_c, which can be achieved by adjusting W_c in a practical system. Furthermore, the maximization process of the PNC relay system is general and can be applied to the relay system with other schemes (e.g., NNC and HNC).



(a) NNC scheme (b) HNC scheme

(c) PNC scheme, $\alpha_{PNC}=1$

Fig. 8. Throughput comparison: analysis v.s. simulation.

V. Simulations Results

In the simulation, the 802.11b protocol is adopted and the physical link rate equals 11Mb/s. Other parameters are shown in Table I.

TABLE I
Simulation parameters

Parameter	Value	Parameter	Value
σ	20 us	CTS	112 bits
SIFS	10 us	ACK	112 bits
DIFS	50 us	Packet size	8472 bits
Physical header	128 bits	Payload size	8184 bits
RTS	160 bits	Propagation delay	1 us

A. Model validation

To validate our analytical models of NNC, HNC and PNC schemes, we calculate the network throughputs of these schemes under different system parameters when the packet generation probability g varies from 0 to its maximum value which makes the network nearly saturated ($P_{NE\_c} = 0.99$), and compare the analytical throughputs with simulations, as shown in Fig. 8. We can see that our models are quite accurate since the analytical results (lines) and the simulation results (symbols) of the three schemes are very close. Another observation from Fig. 8 is that the accuracy of our model decreases when both u and g are bigger. The possible reasons are: When both u and g are bigger, the number of client nodes that have packets to be transmitted increases, which makes the collision probability increases. Since the time slots with collisions are shorter than the time slots with successful transmissions, the variation of h_c is bigger when both u and g are bigger. That is, the accuracy of the approximation that h_c is constant decreases when both u and g are bigger, which leads to a bigger deviation of network throughput.

B. Throughput optimization of the PNC relay system

Next we discuss the impacts of the system parameters on the throughput of the PNC system.

Fig. 9 shows the network throughput versus the minimum contention window size of the client nodes (W_c) under different client node numbers and packet generation probabilities

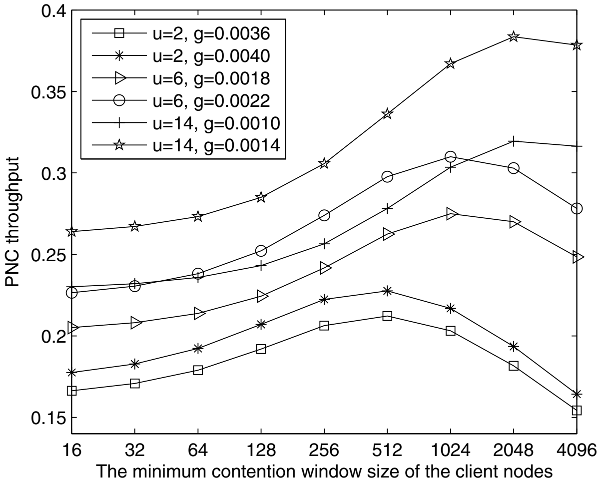


Fig. 9. The PNC throughput v.s. the minimum contention window size of the client nodes.

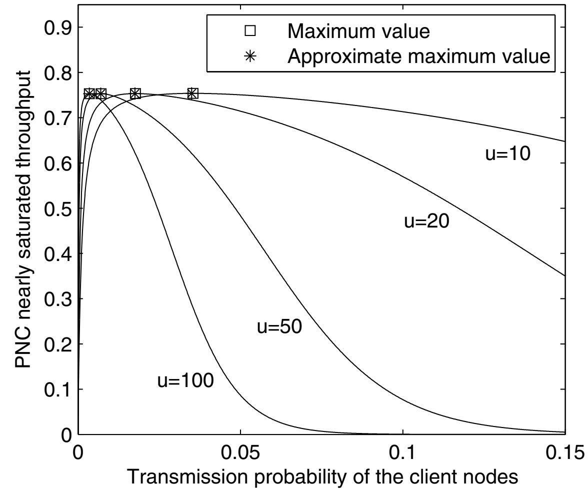


Fig. 11. The PNC nearly saturated throughput v.s. the transmission probability of the client nodes.

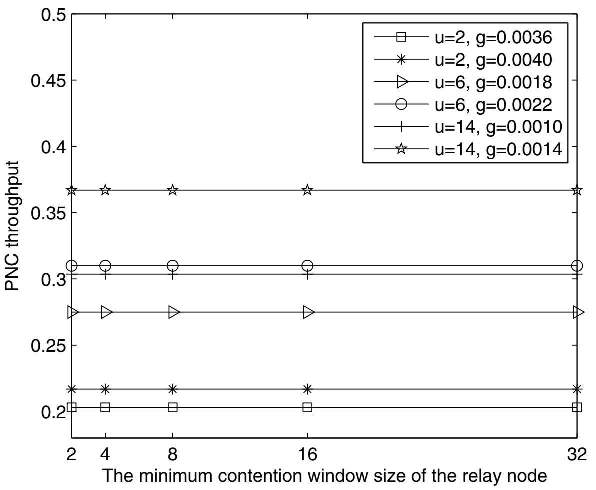


Fig. 10. The PNC throughput v.s. the minimum contention window size of the relay node.

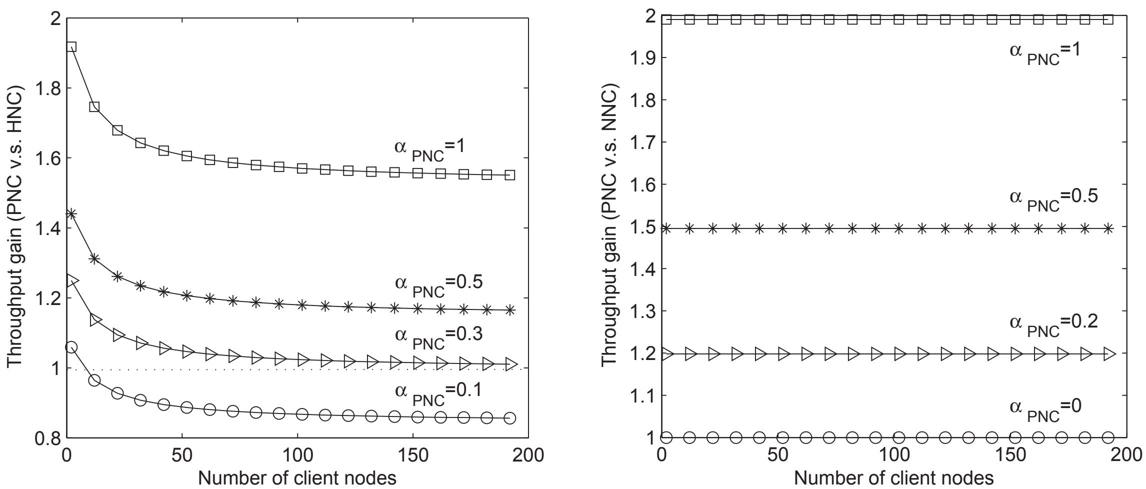
when α_{PNC}, W_r, and m respectively equal 1, 2, and 3. From Fig. 9, we can see that under any given u and g, the network throughput first increases and then decreases as W_c increases. The reason is that a small value of W_c will not only lead to a big transmission probability of a client node but also lead to a big collision probability. Then, when W_c is too small, the channel will be in the collision state most of the time; and when W_c is too big, the channel will be in the idle state most of the time. Therefore, the average time that the channel is used to successfully transmit packets is short when W_c is too small or too big, which leads to a small value of throughput. What's more, from Fig. 9, we can see that the throughput-optimal value of W_c is mainly determined by the number of client nodes in the system but almost independent of the packet generation probability. Thus, it is better to adjust W_c according to the number of client nodes.

Fig. 10 shows the network throughput versus the minimum contention window size of the relay node (W_r) under different client node numbers and packet generation probabilities when α_{PNC}, W_c, and m respectively equal 1, 1024, and 3. From Fig. 10, we can see that the network throughput is almost independent of W_r. The reasons are as follows. According to the analysis in section III-C, to guarantee that the relay system is stable, P_{NE_r} should be smaller than one. That is, the packet arriving probability of the relay node's sending buffer should be smaller than its packet departure probability. Thus, as long as W_r is small enough to make the system work in the steady-state, the total number of packets successfully transmitted by the relay node indeed equals the number of packets successfully transmitted by the client nodes. That is, W_r has little effect on the network throughput and can be set as small as possible to make the system stable.

Fig. 11 shows the PNC nearly saturated throughput versus the transmission probability of a client node k_c when α_{PNC} equals 1. We can see that for any client node number, the approximation result of the optimal k_c in Theorem 1 is quite close to the optimal value obtained from numerical solution. What's more, from Fig. 11, we also find that the maximum nearly saturated throughput is almost independent of the number of client nodes in the system, which is similar to the case in the traditional 802.11 non-relay network [16]. In addition, because of the overhead of 802.11 protocol, the maximum nearly saturated throughput is much smaller than the cut-set capacity bound in this network scenario, which equals one (the normalized link capacity) [30].

C. Throughput comparison of different schemes

To compare the network throughput of different schemes, we consider the nearly saturated case that P_{NE_c} equals 0.99. And in the comparison, W_c, W_r, and m are respectively set to 2048, 2, and 3. The results are shown in Fig. 12. We can see that compared with the HNC and the NNC schemes, the throughput gain of PNC scheme becomes more significant as α_{PNC} increases. For example, for the relay system with one hundred client nodes, the throughput gains of PNC scheme versus HNC and NNC schemes respectively reach about 118%



(a) Throughput gain of PNC scheme versus HNC scheme

(b) Throughput gain of PNC scheme versus NNC scheme

Fig. 12. The throughput gain of PNC scheme.

and 150% when α_{PNC} equals 0.5, and about 157% and 200% when α_{PNC} equals 1. The reason is that α_{PNC} directly determines the probability that an NC data packet contains the information of two data packets. In the extreme case that α_{PNC} equals zero, all the NC data packets in PNC scheme only contain the information of one data packet and the PNC scheme reduces to the NNC scheme. Another observation from Fig. 12 is that the throughput gain of PNC scheme versus HNC scheme decreases when the number of client nodes in the system increases, while the throughput gain of PNC scheme versus NNC scheme keeps unchanged when the number of client nodes in the system varies. The reasons are as follows. As shown in section II-B, the time used by the relay node to successfully transmit a data packet in HNC scheme is longer than the one in PNC scheme. Then, when the number of client nodes in the system is small, the collision rarely occurs, which makes the impact of longer transmission time in HNC scheme more severe. And since the successful transmission time is the same in the NNC and PNC schemes when any node occupies the channel, the throughput gain of PNC scheme versus NNC scheme will be independent of the number of client nodes.

VI. Conclusions

In this paper, we studied the throughput performance of PNC scheme coordinated with IEEE 802.11 DCF. In particular, we derived the analytical unsaturated network throughput results of the relay network with PNC scheme and the traditional HNC, NNC schemes. We found that the throughput benefit of PNC is more significant for bidirectional isochronous traffic with rate requirements. Furthermore, we derived an approximate closed-form solution of the optimal transmission probability of client nodes that maximizes the system throughput in PNC scheme. In addition, we discussed the relationship between the PNC throughput and the system parameters and showed some interesting results: first, to achieve a better network throughput, the minimum contention window size of the client nodes should be self-adaptive according to the number of client nodes; second, the minimum contention window size of the relay node has little effect on the system throughput and thus can be set as small as possible to make the system stable.

Appendix

Proof of Lemma 1.

Proof: The derivative of Y with respect to k_c, $Y'(k_c)$, equals:

$$Y'(k_c) = u(1-k_c)^{u-1} \cdot \frac{\left\{T_{c\_PNC}\left[(-u^2+u)k_c^2-(u+1)k_c+1\right] - (T_{c\_PNC}-\sigma)(1-k_c)^{u+1}\right\}}{\left[T_{c\_PNC}(1-k_c+u\cdot k_c)-(T_{c\_PNC}-\sigma)(1-k_c+u\cdot k_c)(1-k_c)^u\right]^2}. \tag{30}$$

Note that:

$$\begin{cases} 0 \le k_c < 1, \\ u \ge 2, \\ T_{c\_PNC} > 0, \\ \sigma > 0. \end{cases} \Rightarrow \begin{cases} 0 < (1-k_c)^u \le 1, \\ 0 < (1-k_c)^{u-1} \le 1, \\ 1 - k_c + u \cdot k_c > 1, \\ T_{c\_PNC} - (T_{c\_PNC} - \sigma) \\ \cdot (1-k_c)^u > 0. \end{cases} \tag{31}$$

Thus, $\frac{u(1-k_c)^{u-1}}{\left[T_{c\_PNC}(1-k_c+u\cdot k_c)-(T_{c\_PNC}-\sigma)(1-k_c+u\cdot k_c)(1-k_c)^u\right]^2} > 0$.

What's more, let U denote $T_{c\_PNC} \cdot \left[(-u^2+u)k_c^2 - (u+1)k_c + 1\right] - (T_{c\_PNC} - \sigma)(1-k_c)^{u+1}$. Then, the derivative of U with respect to k_c, $U'(k_c)$, equals:

$$\begin{aligned} U'(k_c) &= -2T_{c\_PNC}(u^2-u)k_c \\ &\quad -(u+1)\left[T_{c\_PNC} - (T_{c\_PNC} - \sigma)(1-k_c)^u\right]. \end{aligned} \tag{32}$$

According to equation (31), $U'(k_c)$ is negative. Since $U|_{k_c=0} = \sigma > 0$ and $U|_{k_c=1} = -u^2 T_{c\_PNC} < 0$, U decreases from a positive value to a negative value. Let $U|_{k_c=k_{c0}} = 0$, then,

$$\begin{cases} U > 0, 0 \le k_c < k_{c0}, \\ U < 0, k_{c0} < k_c < 1. \end{cases} \Rightarrow \begin{cases} Y'(k_c) > 0, 0 \le k_c < k_{c0}, \\ Y'(k_c) = 0, k_c = k_{c0}, \\ Y'(k_c) < 0, k_{c0} < k_c < 1. \end{cases} \tag{33}$$

Therefore, as k_c increases from 0 to 1, Y first increases to a maximum value and then decreases. ■

REFERENCES

[1] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical-layer network coding," in *Proc. 2006 ACM Mobicom*, pp. 358–365.

[2] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *Proc. 2006 IEEE ICC*, pp. 3885–3890.

[3] R. Ahlswede, S. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[4] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proc. 2006 IEEE Veh. Technol. Conf.*, pp. 588–593.

[5] S. Katti, S. S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proc. 2007 ACM SIGCOMM*, pp. 397–408.

[6] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, June 2009.

[7] S. Zhang and S. C. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 788–796, June 2009.

[8] S. Wang, Q. Song, X. Wang, and A. Jamalipour, "Distributed MAC protocol supporting physical-layer network coding," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 1023–1036, May 2013.

[9] H. Yomo and Y. Maeda, "Distributed MAC protocol for physical layer network coding," in *Proc. 2011 International Symp. Wireless Personal Multimedia Commun.*, pp. 1–5.

[10] A. Argyriou, "MAC protocol for wireless cooperative physical-layer network coding," in *Proc. 2012 IEEE Wireless Commun. Netw. Conf.*, pp. 1596–1601.

[11] G. Cocco, C. Ibars, D. Gunduz, D. R. Herrero, and O. Collision, "Resolution in slotted ALOHA with multi-user physical-layer network coding," in *Proc. 2011 IEEE Veh. Technol. Conf.*, pp. 1–4.

[12] A. Argyriou and A. Pandharipande, "Cooperative protocol for analog network coding in distributed wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3112–3119, Oct. 2010.

[13] M. Khabbazian, F. Kuhn, N. Lynch, M. Medard, and A. ParandehGheibi, "MAC design for analog network coding," MIT Comput. Science Artificial Intelligence Laboratory Technical Rep., Aug. 2010.

[14] F. Farhadi and F. Ashtiani, "Throughput enhancement of a random access WLAN by combination of digital and analog network coding," in *Proc. 2011 IEEE ICC*, pp. 1–5.

[15] A. Argyriou, "Coordinating interfering transmissions in cooperative wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3804–3812, 2011.

[16] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[17] F. Daneshgaran, M. Laddomada, F. Mesiti, and M. Mondin, "Unsaturated throughput analysis of IEEE 802.11 in presence of non ideal transmission channel and capture effects," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1276–1286, Apr. 2008.

[18] E. Ziouva and T. Antonakopoulos, "The IEEE 802.11 distributed coordination function in small-scale ad-hoc wireless LANs," *International J. Wireless Inf. Netw.*, vol. 10, no. 1, pp. 1–15, Jan. 2003.

[19] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transports protocol over IEEE 802.11 wireless LAN: analysis and enhancement," in *Proc. 2002 IEEE INFOCOM*, vol. 2, pp. 599–607.

[20] P. Chatzimisios, A. Boucouvalas, and V. Vitsas, "IEEE 802.11 packet delay-a finite retry limit analysis," in *Proc. 2003 IEEE GLOBECOM*, vol. 2, pp. 950–954.

[21] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1506–1515, July 2005.

[22] Emad and E. Eylem, "Single hop IEEE 802.11 DCF analysis revisited: accurate modeling of channel access delay and throughput for saturated and unsaturated traffic cases," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3256–3266, 2011.

[23] C. N. Ping and S. C. Liew, "Throughput analysis of IEEE 802.11 multi-hop ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 15, no. 2, pp. 309–322, 2007.

[24] N. Gupta and P. R. Kuman, "A performance analysis of the 802.11 wireless LAN medium access control," *Commun. Inf. Syst.*, vol. 3, no. 4, pp. 279–304, Sept. 2004.

[25] D. Umehara, T. Hirano, S. Denno, M. Morikura, and T. Sugiyama, "Wireless network coding in slotted ALOHA with two-hop unbalanced traffic," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 647–661, June 2009.

[26] D. Umehara, S. Denno, M. Morikura, and T. Sugiyama, "Performance analysis of slotted ALOHA and network coding for single-relay multi-user wireless networks," *Ad Hoc Netw.*, vol. 9, pp. 164–179, Mar. 2011.

[27] D. Umehara, S. Denno, M. Morikura, and T. Sugiyama, "Throughput analysis of two-hop wireless CSMA network coding," in *Proc. 2010 IEEE ICC*, pp. 1–6.

[28] L. B. Jiang and S. C. Liew, "Hidden-node removal and its application in cellular WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2641–2654, Sept. 2007.

[29] R. Mautz, "Indoor positioning technologies," thesis, ETH Zurich, Feb. 2012.

[30] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: tutorial, survey, and beyond," *Physical Commun.*, vol. 6, pp. 4–42, Mar. 2013.

Shijun Lin is currently an assistant professor with the Department of Communication Engineering at Xiamen University (Xiamen, China). He received the B.S. degree (honors) in Electrical Engineering from Xiamen University in 2005, and the Ph.D. Degree in Electrical Engineering from Tsinghua University (Beijing, China) in 2010. His current research interests are in the area of wireless communications and networking, with focus on physical-layer network coding, resource allocation, and distributed protocol design. Dr. Lin is a member of IEEE.

Liqun Fu (S '08-M '11) is currently a postdoctoral fellow with the Institute of Network Coding at The Chinese University of Hong Kong. She received the B.S. degree (honors) in Electrical Engineering from Xiamen University (Xiamen, China) in 2003, the M.S. degree in Electrical Engineering from Tsinghua University (Beijing, China) in 2006, and the Ph.D. Degree in Information Engineering from The Chinese University of Hong Kong in 2010. During 2009, she worked as a visiting research student in the Department of Electrical Engineering, Princeton University. Her current research interests are in the area of wireless communications and networking, with focus on wireless greening, physical-layer network coding, resource allocation, and distributed protocol design.

Dr. Fu has served as the Technical Program Member of leading conferences such as, IEEE Global Communications Conference (Globecom, 2010-2013), IEEE International Conference on Communications (ICC, 2013-2014), IEEE WCNC, etc. She has also served as a frequent technical reviewer for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, etc. She is a founding member of the IEEE Sub-Committee on Green Communications and Computing.