

Energy-Efficient Buffer-Aided Relaying Systems With Opportunistic Spectrum Access

Kunlun Wang^{ID}, *Member, IEEE*, Qingqing Wu^{ID}, *Member, IEEE*, Wen Chen^{ID}, *Senior Member, IEEE*,
Yang Yang^{ID}, *Fellow, IEEE*, and Derrick Wing Kwan Ng^{ID}, *Senior Member, IEEE*

Abstract—In this paper, an energy-efficient cross-layer design framework is proposed for cooperative relaying networks, which takes into account the influence of spectrum utilization probability. Specifically, random arrival traffic is considered and an adaptive modulation and coding (AMC) scheme is adopted in the cooperative transmission system to improve the system performance. The average packet dropping rate of the relay-buffer is studied at first. With the packet dropping rate and stationary distribution of the system state, the closed-form expression of the delay is derived. Then the energy efficiency for relay-assisted transmission is investigated, which takes into account the queueing process of the relay and the source. In this context, an energy efficiency optimization problem is formulated to determine the optimum strategy of power and time allocation for the relay-assisted cooperative system. Finally, the energy efficient switching strategy between the relay assisted transmission and the direct transmission is obtained, where packet transmissions have different delay requirements. In addition, energy efficient transmission policy with AMC is obtained. Numerical results demonstrate the effectiveness of the proposed design improving the energy efficiency.

Index Terms—Energy efficiency and delay tradeoff, cross-layer design, opportunistic spectrum access, mode switching.

Manuscript received September 10, 2019; revised January 17, 2020; accepted February 19, 2020. Date of publication March 2, 2020; date of current version August 19, 2020. The work of Kunlun Wang was supported by the National Natural Science Foundation of China (NSFC) under Grant 61801463. The work of Wen Chen was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1801102, in part by the NSFC under Grant 61671294, in part by the National Huge Project under Grant 2018ZX03001009-002, and in part by STCSM Science and Technology Innovation Program under Grant 17510740700. The work of Derrick Wing Kwan Ng was supported in part by the UNSW Digital Grid Futures Institute, UNSW, Sydney, through a Cross-Disciplinary Fund Scheme and in part by the Australian Research Council's Discovery Project under Grant DP190101363. This article was presented in part at the 2017 International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, October 11–13, 2017. The associate editor coordinating the review of this article and approving it for publication was K. Adachi. (Corresponding author: Wen Chen.)

Kunlun Wang is with the School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China (e-mail: wangkl2@shanghaitech.edu.cn).

Qingqing Wu is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore (e-mail: elewuqq@nus.edu.sg).

Wen Chen is with the Shanghai Institute of Advanced Communications and Data Sciences, Department of Electronics Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: wenchen@sjtu.edu.cn).

Yang Yang is with the Shanghai Institute of Fog Computing Technology, SCA, ShanghaiTech University, Shanghai 201210, China (e-mail: yangyang@shanghaitech.edu.cn).

Derrick Wing Kwan Ng is with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: w.k.ng@unsw.edu.au).

Digital Object Identifier 10.1109/TGCN.2020.2977345

I. INTRODUCTION

DURING the past decades, wireless communication systems have been designed with the main focus on improving the network throughput or spectral efficiency [2]. However, with the rapid realization of Internet of Things (IoT), power consumptions at both base stations (BSs) and resource-limited devices such as user terminals are no longer sustainable, which is attributed to the fundamental bandwidth-power and delay-power tradeoffs [1], [3], [4]. As reported, the network energy consumption should be decreased by a factor of 1,000 in the upcoming fifth-generation (5G) networks while without compromising the quality of service (QoS) [5]. As a result, energy efficiency, which is measured by bits-per-joule, has emerged as a key figure of merit and become the most widely adopted design metric for green IoT systems. In fact, extensive research efforts have been made to improve the energy efficiency for various kinds of wireless communication applications, such as orthogonal frequency division multiplexing access (OFDMA) [6]–[9], multiple-input multiple-output (MIMO) [10]–[13], millimeter-wave (mmWave) communications [14], [15], cognitive radio [16], [17], small cell [18], [19], and wireless powered communications [20], [21], etc. However, most of the previous research works mainly focus on energy-efficient designs for physical layer transmission. In the future IoT networks, traffic delay requirements are diversified from the upper layers imposing challenges to physical layer design in resource allocation. Thus far, many studies have been carried out on cross-layer design to improve the energy efficiency of the whole system [1], [22]–[24], which is expected to be more appealing and practical for the next generation of wireless networks.

On the other hand, the gap between the designing of high data rate and spectrum scarcity motivates the advent of cognitive spectrum access, which has been regarded as a revolutionary technology to tackle such a challenge [25]. When the spectrum is detected to be unoccupied by the primary system, the secondary system could utilize the free bands temporarily to perform its own transmission. Then there exists a utilization probability for each transmission link of the secondary system. Furthermore, the utilization probability may influence the buffer-overflow and the delay performance of the packet for the secondary system, leading to different energy efficient strategies. In a cellular network, each secondary user has a chance to transmit data over the free bands. Due to the severe channel fading or shadowing, the selected secondary

user may terminate the transmission from the direct link to save energy. In this case, any other nearby secondary users with better channel conditions can act as a relay to assist the data transmission to improve the energy efficiency of the secondary system. Furthermore, since the secondary users may require different delay-aware applications over the cooperative network, providing heterogeneous QoS of delays for energy-efficient cooperative systems, has become critical and necessary.

A. Related Works

Relay nodes can help source node to transmit packets to destination node in cooperative systems [26], [27], which has been extensively studied recently. In addition, relay equipped with buffer can further improve system performance and introduce new degrees of freedom for system design [28]–[31]. In particular, various cooperative schemes have been proposed to exploit cooperative diversity for improving the physical performance [32]–[34], such as outage probability and signal-to-noise ratio (SNR). In practical systems, the data link layer performance metric, e.g., delay, also plays an important role in wireless networks. Since delay can influence the performance of the cooperative systems, it also has a significant impact on protocol and system designs, especially when the system is required to guarantee the performance of delay sensitive services and delay tolerant services simultaneously.

The notion of delay in cooperative wireless networks has been considered in recent literatures. In [35], the performance metrics defined by stable throughput region and delay are evaluated for a packet transmitted to the destination through either a direct link or relay links. The delay-optimal link selection for a two-hop three-node cooperative network with bursty packet arrivals is studied in [36]. Multi-hop cases are considered in [37]. Motivated by real-time applications having stringent delay constraints, optimal resource allocation for delay-limited cooperative communication in time varying wireless networks is recently discussed in [38].

Most of the existing works on cooperative communications, e.g., [35]–[38], are based on throughput optimization under delay constraint. However, all these works do not consider the energy efficient communication. In [24], delay constrained energy-efficient problem in cooperative wireless networks is studied. However, practical adaptive modulation and coding (AMC) is not considered. In [39], the authors investigate the AMC scheduling for a cooperative wireless system with multiple relays operating in a modified decode-and-forward protocol. However, adaptive scheduling strategy between the cooperative transmission and direct transmission is not considered. Although the scheduling strategy with direct transmission and relay-assisted transmission for cognitive radio network is studied in [40], the packet delay and energy efficiency are not considered. Furthermore, due to the multi-path channel fading, the noise, and the co-channel interference from primary system, it is difficult to guarantee the deterministic delay for communication services in cognitive radio networks. Thus, as a compromise scheme, it is a practical and ultimate goal

to provide statistical delay guarantee for cooperative cognitive radio networks.

B. Main Contributions

In this work, we consider the cross-layer energy-efficient design approach for cooperative IoT networks with opportunistic spectrum access, which aims at taking the physical layer transmission and delay-aware service into account, the influence of spectrum utilization probability can also be obtained. Generally speaking, the AMC and cooperative transmission strategy are part of the physical layer decisions, the packet retransmission is controlled by the data link layer, and the delay statistic of the traffic is provided by the application layer. In this research direction, there are some works on cross-layer designs considered in [39]–[41], where a similar system model is considered. In particular, the cross-layer design for cognitive relay systems is investigated in [40]. However, our work is fundamentally different from that of [39]–[41] in several significant aspects. First, our work aims to study the impact of delay-aware service on the design of energy-efficient scheduling, while [39], [40] focus more on the bit-error-rate (BER)-based scheduling. Second, our objective is to maximize the energy efficiency of the delay-aware service in cooperative IoT networks with opportunistic spectrum access. Based on such an objective, we derive an expression of the expected delay in a closed-form which captures the impacts of the queueing and transmission delay in both the source and relay buffers. To the best knowledge of the authors, this important aspect has not been reported in the literature [44], yet. Besides, we have also derived the packet dropping rate for the buffer of the relay node, which is not considered in [39], [40] either. Third, we investigate the performance of energy efficiency by utilizing the AMC transmission and the time allocation for the direct transmission and relay-assisted transmission. Last but not least, for maximizing the system energy efficiency, we derive an optimized threshold by taking spectrum access probability into account, which determines the selection between direct transmission and relay-assisted transmission. This contribution is novel and different from the previous works.

The challenge of this work is to derive the packet dropping rate at the relay node, which can influence the throughput and energy efficiency of the cooperative system. To overcome the challenge, we need to prove the existence of the stationary distribution of the cooperative system, and get the stationary distribution. Therefore, our contribution in this work can be summarized as follows: We obtain the closed-form expression of expected delay for the cooperative transmission system, which considers both the transmission and queueing delay in both the source and relay buffer. Then we obtain the closed-form expression for the system throughput by taking into account the packet drop caused by both the transmission error and the buffers overflow. Correspondingly, the average energy consumption of the cooperative system is also derived. Considering the opportunistic spectrum access, we reveal the intrinsic relationship between the energy-efficient scheduling strategy and different delay requirements. In particular, we

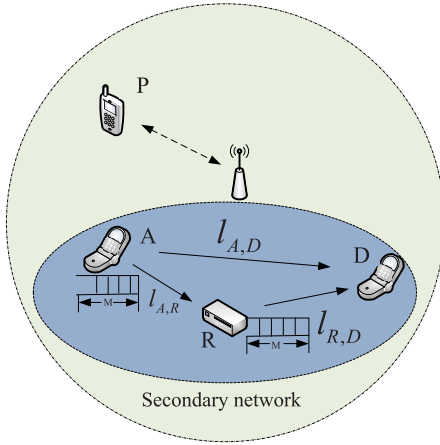


Fig. 1. System model.

show that the energy efficient scheduling strategy is to adaptively switch between direct transmission and relay-assisted transmission based on different delay requirements. At last, we obtain the energy-efficient transmission policies for the inter-node link transmission and direct link transmission.

C. Organization

The rest of the paper is organized as follows. Section II introduces the system model and problem formulation, including the network model, the queueing model and the traffic model, the channel model, and the medium access model. In Section III, we investigate the energy-efficiency for the relay assisted transmission and the direct transmission. In Section IV, we analyze the energy-efficient scheduling for the delay-aware service based on the theoretical results obtained in Section III. Section V presents numerical results to demonstrate the performance of the proposed cross-layer design. Finally, we conclude the paper in Section VI.

II. SYSTEM MODEL

A. Network Model

As shown in Fig. 1, we consider a cooperative secondary system which consists of three nodes that share the same spectrum with a legacy primary system (PS) node P . Nodes A , R , and D , are the source, relay, and destination of the secondary system, respectively. Specifically, the relay node adopts the decode-and-forward (DF) strategy to assist the end-to-end transmission. Due to the bursty service in the PS, the secondary system can access the spectrum of the PS in an opportunistic manner with a certain probability. We assume that source node A generates data traffic and communicates with node D . In the cooperative secondary network, the data packets can be transmitted via the direct link, $l_{A,D}$, or via the relay path composed of links $l_{A,R}$ and $l_{R,D}$ shown in Fig. 1. For each link, i.e., $l_{A,D}$, $l_{A,R}$, or $l_{R,D}$, the secondary system will terminate the transmission if one of the following situation occurs: 1) the PS occupies the spectrum during the transmission; 2) the packet error rate (PER) of the link $l_{A,R}$ exceeds a maximum tolerable level due to deep channel fading.

B. Traffic and Queueing Models

We assume that the incoming traffic of source A is random, which follows Poisson distribution. A wide range of multimedia traffic can be represented with an Poisson model, which is accurate and reasonable [42]. Specifically, the average packet arrival rate from the application layer is λ . Besides, both the source-buffer and the relay-buffer maintain the first-in-first-out (FIFO) queue and the packet streams have the size of L bits. Assume that the queueing model of the source node or the relay node is a single server M/G/1 queue [43].

For the source node A , the M/G/1 model assumes the single server with Markovian or memoryless arrivals at an average rate λ and a general service distribution. Due to the general service distribution, denote by χ_n the service rate at the physical link corresponding to channel state n . In addition, we assume that the set of service state is $\Psi = \{\chi_1, \dots, \chi_N\}$.

For simplicity, we assume that the buffers of the source and the relay have finite capacity in each of which, denoted as M . Due to the bursty arrival traffic, the queues of the nodes may be empty sometimes or full some other times. Therefore, for the latter case, the incoming packets may be dropped at both of the source node and the relay node because of the bursty traffic arrived and the finite buffer size. We assume that the packets dropped by the source and relay nodes are regarded as lost.

C. Channel Model

All inter-node wireless channels are assumed to have Nakagami- m fading with the same fading parameter m , and are block fading with additive white Gaussian noise (AWGN).

In order to get a well defined relation between the channel quality and received signal-to-noise ratio (SNR), the transmit signal power of the source node is constant at \bar{e} . The channel state information (CSI) is supposed to be perfectly known by the receiver with training, then the channel quality can be captured by the received SNR \mathcal{S} . Due to block fading, \mathcal{S} remains invariant within each transmission frame but can vary from frame to frame. It should be noted that the received SNR \mathcal{S} per frame of a fading channel follows a Gamma distribution, which has probability density function (PDF) as

$$f_{\bar{\mathcal{S}}}(\mathcal{S}) = \frac{m^m \bar{\mathcal{S}}^{m-1}}{\bar{\mathcal{S}}^m \Gamma(m)} \exp\left(-\frac{m\mathcal{S}}{\bar{\mathcal{S}}}\right), \quad (1)$$

where $\bar{\mathcal{S}} \triangleq E\{\mathcal{S}\}$ is the average received SNR and $\Gamma(m) \triangleq \int_0^\infty t^{m-1} e^{-t} dt$ is the Gamma function [44]. We assume that inter-node channels of $l_{A,R}$ and $l_{R,D}$ are independent identically distributed (i.i.d.) with the average SNR $\bar{\mathcal{S}}$ and $\bar{\mathcal{S}}_R$, respectively, and the average SNR of the direct channel $l_{A,D}$, denoted by $\bar{\mathcal{S}}_{S,D}$.

For packet transmission at each inter-node link, the adaptive modulation and coding (AMC) is considered. In particular, the total available number of AMC transmission modes is N . To select an AMC mode, the received SNR needs to fall into the corresponding interval. Then, we can divide the entire received SNR range into $N + 1$ nonoverlapping consecutive intervals, which can be denoted as $\{\mathcal{S}_n\}_{n=0}^{N+1}$, where the boundary points is given by $\mathcal{S}_0 = 0$ and $\mathcal{S}_{N+1} = +\infty$. Based

on the above description, the channel is in state n satisfying $\mathcal{S} \in [\mathcal{S}_n, \mathcal{S}_{n+1})$. Based on (1), we can obtain the probability that the channel being in state n as

$$P_r(n) = \int_{\mathcal{S}_n}^{\mathcal{S}_{n+1}} f_{\bar{\mathcal{S}}}(\mathcal{S}) d\mathcal{S}. \quad (2)$$

Similarly, the probabilities that the channel being in state n for the relay-to-destination link and the source-to-destination link can be obtained as $P_r^R(n)$ and $P_r^{S,D}(n)$.

Denote by b_n the transmission modulation when the channel is in state n . Specifically, if the receiver cannot correctly decode a packet, the transmitter will be notified to repeat transmitting the packet where the maximum number of retransmissions is $N_r^{\max} - 1$. In contrast, when the receiver correctly decodes a packet, it will feed back an acknowledgement (ACK) packet to the transmitter. However, in some cases for which the packet cannot be correctly decoded at the receiver even after $N_r^{\max} - 1$ times of retransmissions, then this packet will be dropped and a packet loss will be declared. Therefore, packet loss may happen in each link of the cooperative system. With the average packet error rate of transmissions, both the destination node in direct link and the relay node can fail to decode the received packet, an negative acknowledgement (NACK) packet will be transmitted to notify the source node. If the destination node successfully receives the packet, but the relay node fails to receive the packet, then the ACK packet will be transmitted, the packet will be transmitted through the direct link. On the other hand, if the relay node successfully receives the packet in the relay-assisted transmission, but the direct source-to-destination transmission fails, then the ACK packet will be transmitted, the packet will be transmitted through the relay-assisted link. In practice, the probability of packet loss is regained to no larger than a threshold P_{loss} to satisfy the quality-of-service (QoS). It is worth pointing out that the service rates of the source node and relay node depend not only on the transmission modulation b_n but also the maximum retransmission times $N_r^{\max} - 1$.

D. Medium Access Model

In general, the spectrum occupancy of the PS can be modeled as a continuous-time Markov chain with available (the link is idle) and unavailable (the link is busy) states [45]. In this work, we assume that the transmissions of the PS and secondary cooperative system are both continuous. Thus, the spectrum occupied times from the secondary cooperative system are independent and exponentially distributed with aggregated parameter $q_{i,j}^{-1}$ for the available state and $u_{i,j}^{-1}$ for the unavailable state on link $l_{i,j}$ ($i, j \in \{A, R, D\}$). Under this model, the stationary probability of the available and unavailable states are respectively given by [45] as

$$a_{i,j}^a = \frac{q_{i,j}}{q_{i,j} + u_{i,j}}, \quad (3)$$

and

$$a_{i,j}^u = \frac{u_{i,j}}{q_{i,j} + u_{i,j}}, \quad (4)$$

respectively.

E. Problem Statement

In the considered cooperative network, the system power consumption for transmitting one packet in the secondary network is denoted by $\mathcal{P}_{(T)}$, which indicates the total amount of energy consumed in one second. Accordingly, the system throughput is denoted by \mathcal{R} , which indicates the number of packets which are transmitted successfully without error in one second. In general, the energy efficiency is defined as the ratio of average system throughput over the total system power consumption, i.e.,

$$\eta_{ee} \triangleq \frac{\mathcal{R}}{\mathcal{P}_{(T)}} \quad \text{packets/joule}, \quad (5)$$

which indicates the number of successfully received packets per joule of energy consumed.

Since different delay requirements of the traffic and spectrum access probabilities may result in different energy efficiencies of the cooperative network, we should optimize the selection of transmission mode between direct transmission and relay-assisted transmission to maximize the system energy efficiency. While the maximum tolerable delay is \mathcal{D}_0 , we call the selection of transmission mode switching strategy φ . For the packet transmission, AMC schemes are used for both the source node A and relay node R . Then, we are interested in getting the rate adaption policies maximizing the energy efficiency by taking into account the maximum tolerable packet delay \mathcal{D}_0 . Specifically, the rate adaption policy is explicitly represented by the probability distribution of modulation b_n , $n = \{1, \dots, N\}$. Denote the rate adaptive policy as ω , which depends on the \mathcal{S}_n and $\bar{\mathcal{S}}$. Mathematically, the design of transmission mode switch strategy and AMC rate adaption policy can be formulated as

$$\max_{\{P_r(n)\}, \varphi} \left\{ \eta_{ee} | \mathcal{D}_{(T)} \leq \mathcal{D}_0, \varphi \in \{d, r\} \right\}, \quad (6)$$

where $\mathcal{D}_{(T)}$, d and r represent the average packet delay from the arrival in the source buffer until the destination correctly received, the direct transmission and the relay assisted transmission, respectively. Based on (2), the problem is equivalent to $\max_{\bar{\gamma}, \varphi} \{ \eta_{ee} | \mathcal{D}_{(T)} \leq \mathcal{D}_0, \varphi \in \{d, r\} \}$.

III. ENERGY EFFICIENCY ANALYSIS

In this section, we evaluate the delay and the energy efficiency of a packet transmission in the cooperative network with opportunistic spectrum access. When the direct transmission is not possible, i.e., the received SNR is smaller than a threshold, the source node selects the relay to assist the transmission to achieve better performance. We denote the total time duration of each transmission period as $T \triangleq T_1 + T_2$ and denote the ratio of the time allocation in the phase of source-to-relay over the whole period as $\alpha \triangleq T_1/T \in (0, 1)$, where T_1 and T_2 are the transmission time durations for source-to-relay link and relay-to-destination link, respectively.

A. Throughput Analysis

1) *Service Rate*: The source node selects transmission modulation based on AMC, which is also adopted by the relay

node. Intuitively, due to the long distance between the source and destination or limited power at the source, the packets transmission is assisted by a relay. At the destination, the receiver attempts to decode the packet from the source or the relay. Regarding the successfully receiving packets, an ACK packet will be transmitted to notify the source node at the end of the time slot. Once the source node receives the ACK packet, the corresponding packet is removed from the source buffer.

Assume each packet contains L bits. The transmission packet rate R_p can be obtained as

$$R_p = \frac{R_s b_n^{l_{A,D}}}{L}. \quad (7)$$

Now, with the packet error rate PER_n at channel state n and the probability that the channel being in state n , the average number of error transmitted packets can be derived as

$$\sum_{n=1}^N \frac{R_s b_n^{l_{A,D}}}{L} P_r^{S,D}(n) PER_n. \quad (8)$$

Similarly, the average total number of transmitted packets can be obtained as

$$\sum_{n=1}^N \frac{R_s b_n^{l_{A,D}}}{L} P_r^{S,D}(n). \quad (9)$$

With (8) and (9), we can derive the average packet error rate of the direct source to destination link as

$$P_{LD} = \frac{\sum_{n=1}^N \frac{R_s b_n^{l_{A,D}}}{L} P_r^{S,D}(n) PER_n}{\sum_{n=1}^N \frac{R_s b_n^{l_{A,D}}}{L} P_r^{S,D}(n)}. \quad (10)$$

Similarly, the average packet error rates of the source-to-relay link and the relay-to-destination link via the help of the relay can be derived as

$$\begin{aligned} P_{L1} &= \frac{\sum_{n=1}^N \frac{R_s b_n^{l_{A,R}}}{L} P_r^{S,D}(n) PER_n}{\sum_{n=1}^N \frac{R_s b_n^{l_{A,R}}}{L} P_r^{S,D}(n)}, \\ P_{L2} &= \frac{\sum_{n=1}^N \frac{R_s b_n^{l_{R,D}}}{L} P_r^R(n) PER_n}{\sum_{n=1}^N \frac{R_s b_n^{l_{R,D}}}{L} P_r^R(n)}, \end{aligned} \quad (11)$$

respectively. Note, the metric of average packet error rate is defined by the ratio of the average number of packets cannot be successfully received over the total average number of transmitted packets, and the packet error rate PER_n at channel state n is denoted as [39, eq. (5)]

$$PER_n = \begin{cases} \alpha_n \exp(-g_n \mathcal{S}), & \mathcal{S} \geq \mathcal{S}_{pn} \\ 1, & \mathcal{S} < \mathcal{S}_{pn} \end{cases} \quad (12)$$

and α_n , g_n , and \mathcal{S}_{pn} are mode dependent parameters.

Considering the single packet transmission with two phases, the destination node and the relay node may receive this packet. With the average packet error rate of transmissions, the probability that both the destination node and the relay node fail to decode the received packet in the first phase is given by

$$P_1 = P_{LD} P_{L1}. \quad (13)$$

On the other hand, if we assume that the relay node successfully receives the packet in the first phase, but the direct source-to-destination transmission fails in the first phase, then the packet cannot be successfully received at the destination with probability

$$P_2 = P_{LD}(1 - P_{L1})P_{L2}. \quad (14)$$

Based on (13) and (14), the overall probability that a single packet cannot be successfully received by the destination via the relay-assisted transmission can be obtained as follows

$$P_0 = P_1 + P_2 = P_{LD} \cdot P_L + P_{LD}(1 - P_{L1})P_{L2}. \quad (15)$$

Now, we study the packet's service rate via the following two lemmas.

Lemma 1: For the relay-assisted transmission at channel state n , the packet's service rate from the source is

$$\chi_n = \frac{\alpha(1 - P_1)}{f(\epsilon_n) + \bar{\epsilon}_n P_1 \left[1 - P_1^{N_r^{\max}} (1 + N_r^{\max}(1 - P_1)) \right]}, \quad (16)$$

where $f(\epsilon_n) = \epsilon_n(1 - P_1^{N_r^{\max}+1})(1 - P_1)$ and ϵ_n is shown in (17), at the bottom of the page, which is the transmission time of the packet.

Proof: The proof is given in Appendix A. ■

Lemma 2: For the relay-assisted transmission, the packet's service rate from the relay to the destination at channel state n is

$$\chi'_n = \frac{(1 - \alpha)(1 - P_{L2})}{f(\epsilon'_n) + \bar{\epsilon}'_n P_{L2} \left[1 - P_{L2}^{N_r^{\max}} (1 + N_r^{\max}(1 - P_{L2})) \right]}, \quad (18)$$

where $f(\epsilon'_n) = \epsilon'_n(1 - P_{L2}^{N_r^{\max}+1})(1 - P_{L2})$, $\epsilon'_n = \frac{L}{b_n^{l_{R,D}} R_s a_{R,D}^a}$ and $\bar{\epsilon}'_n = \frac{L}{b_n^{l_{R,D}} R_s a_{R,D}^a}$ represent the relay to the destination packet transmission time and average retransmission time for each retransmission, respectively.

Proof: Similar to the proof of Lemma 1. ■

Finally, for the source and the relay node, we obtain the sets of service rates as $\Psi = \{\chi_1, \dots, \chi_N\}$ and $\Psi' = \{\chi'_1, \dots, \chi'_N\}$, respectively.

$$\begin{aligned} \epsilon_n &= \tau_n^{l_{A,D}}(1 - P_{LD}) + \tau_n^{l_{A,R}, l_{R,D}} P_{LD} \\ &= \frac{L}{b_n R_s a_{A,D}^a} (1 - P_{LD}) + P_{LD} \left(\frac{L}{b_n R_s a_{A,R}^a} + \frac{L}{b R_s a_{R,D}^a} \right) \end{aligned} \quad (17)$$

2) *Packet Dropping Rate*: Due to the bursty service, the packet can be dropped from the buffer due to buffer overflow. Now, the packet dropping rate at the source node A and the relay node R can be written as P_d^A and P_d^R , respectively. Intuitively, when the number of packet arrivals is larger than the remaining space of the buffer, some packets will be dropped. At time t , assumed that the current service rate is χ'_t at relay node R , and the service rate is χ'_{t-1} at previous time. Then, the remaining space in the buffer of node R is sized by $r_t^R = M - (\chi'_{t-1} - \chi'_t)$. As a result, the buffer for the relay node can accommodate at most r_t^R arriving packets in the current time slot. Now, under the condition that the number of arriving packets a_t^R for relay node is larger than r_t^R , there are $a_t^R - r_t^R$ packets need to be dropped.

Theorem 1: Under the condition that the average transmission rate from the source to the relay $\bar{\chi} > M - (S_{t-1}^R - \chi'_t)$, the packet dropping rate at relay node R is $P_d^R = \frac{E\{\Delta^R\}}{\bar{\chi}}$, where

$$\begin{aligned} E\{\Delta^R\} &= \sum_{s' \in \mathbb{S}, \chi' \in \Psi'} \Delta^R \cdot \mathbb{B}(\bar{\chi} | S_{t-1}^R = s', \chi'_t = \chi') \\ &= \sum_{s' \in \mathbb{S}, \chi' \in \Psi'} \left[(\bar{\chi} - (M - (S_{t-1}^R - \chi'_t))) \times \pi'_{s', \chi'} \right] \end{aligned}$$

is the average number of dropped packets at relay node R and $\pi_{s', \chi'}$ is the stationary distribution of the buffer state and the service rate state for the relay to the destination transmission system.

Proof: The proof is given in Appendix B. ■

In order to derive the packet dropping rate P_d^A at the source buffer, the stationary distribution π of the buffer system should also be considered. At the time t , the number of arrival packets at source is a_t^A . Considering that the number of packets dropped at source node A is Δ_t^A , the packet dropped rate P_d^A can be computed as [46]

$$P_d^A \triangleq \lim_{T \rightarrow \infty} \frac{\sum_{t=1}^T \Delta_t^A}{\sum_{t=1}^T a_t^A} = \frac{E\{\Delta^A\}}{E\{a_t^A\}} = \frac{E\{\Delta^A\}}{\bar{\lambda}}. \quad (19)$$

Based on Theorem 1, the average number of dropped packets $E\{\Delta^A\}$ can be derived as in (20).

$$\begin{aligned} E\{\Delta^A\} &= \sum_{s \in \mathbb{S}, \chi \in \Psi} \Delta^A \cdot \mathbb{B}(a_t = a, S_{t-1}^A = s, \chi_t = \chi) \\ &= \sum_{s \in \mathbb{S}, \chi \in \Psi} \left[\left(a_t^A - \left(M - (S_{t-1}^A - \chi_t) \right) \right) \times \pi_{s, \chi} \right] \end{aligned} \quad (20)$$

With P_d^A and P_d^R available, we can get the average traffic rate and the system throughput. By considering the influences of both the source-queueing and the relay-queueing, we will analyse the packet delay in the next subsection.

3) *Network Throughput*: To analyze the performance of the network throughput, we need to consider how the packets are successfully transmitted to the destination from the cooperative network. During the packet transmission, the packet dropping rate from the source and relay queueing and packet violation from the channel with N_r^{\max} retransmissions are influencing the successfully transmission. This means that the system average throughput is related to not only the packet dropping rate P_d^A and P_d^R , but also the average overall packet error rate P_0 . Then, for an average packet arrival rate of $\bar{\lambda}$, we can obtain the system average throughput \mathcal{R} as

$$\mathcal{R} = \bar{\lambda} (1 - P_d^A) (1 - P_d^R) (1 - P_0^{N_r^{\max}}). \quad (21)$$

B. Delay Analysis

After studying the network throughput in the previous subsection, we now focus on the average packet delay for the relay-assisted network in details, which considers the spectrum access probability.

By considering the stationary distribution of the source-buffer system, the average queue length of source node A at channel state n can be obtained as

$$\bar{Q}_{(q,A)}^n = \sum_{s \in \mathbb{S}} \sum_{\chi = \chi_n} \pi_{(s, \chi)} \cdot s. \quad (22)$$

At the same time, by considering the stationary distribution $\pi_{(s', \chi')}$ of the relay-buffer system, we can get the average queue length of the relay node R at channel state n as

$$\bar{Q}_{(q,R)}^n = \sum_{s' \in \mathbb{S}, \chi' = \chi'_n} \pi_{(s', \chi')} \cdot s'. \quad (23)$$

With regard to the M/G/1 queue, it is known that the average waiting time of a packet consists of queueing time and service time. According to Little's formula [43], the queueing delay is $\bar{\mathcal{D}}_q = \frac{\bar{Q}_q}{r}$. Then, we can get the average packet delay $\bar{\mathcal{D}}_n$ for channel state n given by (24), shown at the bottom of the page, where $\bar{\chi} = \sum_{n=1}^N P_r(n) \chi_n$. In order to simplify the calculation, the same channel state can be considered for the inter-node channel. Then, we can obtain the average packet delay over the relay-assisted transmission system as

$$\mathcal{D}_{(T)} = \sum_{n=1}^N P_r(n) \bar{\mathcal{D}}_n. \quad (25)$$

Lemma 3: For the cooperative transmission systems, the average delay $\mathcal{D}_{(T)}$ is a decreasing function of the average received SNR \bar{S} .

Proof: Lemma 3 is intuitively correct. Notice that when increasing \bar{S} , the probabilities of choosing larger transmission rate for both the source node and the relay node are increasing. According to (24), $\bar{\mathcal{D}}_n$ is decreasing for $\forall n$. As a result, $\mathcal{D}_{(T)}$ is decreasing. Particularly, the delay can be formulated as $\mathcal{D}_{(T)} = \phi(\bar{S})$, and $\phi(\bar{S})$ is a decreasing function of \bar{S} . ■

$$\bar{\mathcal{D}}_n = \frac{\bar{Q}_{(q,A)}^n}{\bar{\lambda}} + \frac{\bar{Q}_{(q,R)}^n}{\bar{\chi}} + E\{\delta_n\} = \frac{\bar{Q}_{(q,A)}^n}{r} + \frac{\bar{Q}_{(q,R)}^n}{\bar{\chi}} + \frac{f(\epsilon_n) + \bar{\epsilon}_n P_1 [1 - P_1^{N_r^{\max}} (1 + N_r^{\max} (1 - P_1))]}{1 - P_1} \quad (24)$$

C. Power Consumption

Let \mathcal{P}_n denote the transmission power for the source node at channel state n , and the BER ϑ of the packet can be expressed as a function of the post-adaptation received SNR $\mathcal{S}\mathcal{P}_n/\bar{e}$ [47]

$$\vartheta \approx 0.2 \exp\left(-\frac{1.5}{2^{b_n}-1} \frac{\mathcal{P}_n}{\bar{e}} \mathcal{S}\right), \quad (26)$$

where b_n is the modulation size for AMC mode n .

Using (26), we obtain the transmission power \mathcal{P}_n as a function of the BER ϑ and \mathcal{S} as follows:

$$\mathcal{P}_n = \frac{\bar{e}(2^{b_n}-1)}{1.5\mathcal{S}} \ln \frac{0.2}{\vartheta}. \quad (27)$$

Thus, the average transmission powers of source node and relay node at channel state n can be respectively obtained as

$$\begin{aligned} \bar{\mathcal{P}}_n &= \int_{\mathcal{S}_n}^{\mathcal{S}_{n+1}} \frac{\bar{e}(2^{b_n}-1)}{1.5\mathcal{S}} \ln \frac{0.2}{\vartheta} f_{\bar{\mathcal{S}}}(\mathcal{S}) d\mathcal{S}, \\ \text{and } \bar{\mathcal{P}}_n^R &= \int_{\mathcal{S}_n}^{\mathcal{S}_{n+1}} \frac{\bar{e}(2^{b_n}-1)}{1.5\mathcal{S}} \ln \frac{0.2}{\vartheta} f_{\bar{\mathcal{S}}_R}(\mathcal{S}) d\mathcal{S}, \end{aligned} \quad (28)$$

where BER can be obtained by PER_n . Based on (2) and (28), the average transmit power \mathcal{P} of source node can be shown as

$$\mathcal{P} = \sum_{n=1}^N \bar{\mathcal{P}}_n P_r(n), \quad (29)$$

Similarly, we can obtain the average transmit power \mathcal{P}_R .

D. Energy Efficiency of Relay-Assisted Transmission

Note that the transmit power for the relay node is \mathcal{P}_R . Furthermore, we assume that the power consumption for the node at idle state is constant \mathcal{P}_0 . Then, the total energy consumption can be obtained by the following lemma.

Lemma 4: For the secondary relay-assisted transmission networks, the total energy consumption with opportunistic spectrum access for each transmission period is

$$\zeta = \left(a_{A,R}^a \mathcal{P} + a_{A,R}^u \mathcal{P}_0\right) T_1 + \left(a_{R,D}^a \mathcal{P}_R + a_{R,D}^u \mathcal{P}_0\right) T_2. \quad (30)$$

Proof: The proof is given in Appendix C. ■

Based on (21) and Lemma 4, the energy efficiency is given by

$$\begin{aligned} \eta_{ee} &= \frac{T\bar{\lambda}(1-P_d^A)(1-P_d^R)(1-P_0^{N_r^{\max}})}{\zeta} \\ &= \frac{T\bar{\lambda}(1-P_d^A)(1-P_d^R)(1-P_0^{N_r^{\max}})}{\left(a_{A,R}^a \mathcal{P} + a_{A,R}^u \mathcal{P}_0\right) T_1 + \left(a_{R,D}^a \mathcal{P}_R + a_{R,D}^u \mathcal{P}_0\right) T_2}. \end{aligned} \quad (31)$$

We aim to jointly optimize the average SNR $\bar{\mathcal{S}}$, $\bar{\mathcal{S}}_R$ and the time allocations for the cooperative transmission based on the closed-form expression of the energy efficiency. Note that

Algorithm 1 Energy Efficient Time Allocation Ratio and Power Allocation for the Proposed Relay Assisted Transmission

- 1: For $\alpha=0:1$
 - Step 1)** Initialize the energy efficient threshold value $\mathcal{S}_n = \mathcal{S}_n^*$ (see [41]), $\forall n = 1, 2, \dots, N$.
 - Step 2)** Repeat:
 - calculate η_{ee} using (31);
 - update $\bar{\mathcal{S}}$ and $\bar{\mathcal{S}}_R$ using the Golden Section search method;
 - Step 3)** Until η_{ee} converge.
- 2: End.
- 3: Energy efficient time allocation ratio: $\alpha^* = \arg \max \eta_{ee}$, and power allocation for the source node and relay node: $\mathcal{P}^* = \{\mathcal{P}\}_{\bar{\mathcal{S}}=\bar{\mathcal{S}}^{\text{opt}}, \bar{\mathcal{S}}_R=\bar{\mathcal{S}}_R^{\text{opt}}}$.

the time allocation ratio α may take any number in the range of (0, 1). The optimization problem can be specified as follows:

$$\begin{aligned} \max_{\alpha, \bar{\mathcal{S}}, \bar{\mathcal{S}}_R} \quad & \eta_{ee} \\ \text{s.t.} \quad & 0 < \alpha < 1, \\ & \bar{\mathcal{S}}_{\min} < \bar{\mathcal{S}}, \bar{\mathcal{S}}_R < \bar{\mathcal{S}}_{\max}. \end{aligned} \quad (32)$$

For any given time allocation ratio $\alpha \in (0, 1)$, we found that we are able to express the corresponding energy efficient transmit power \mathcal{P} and \mathcal{P}_R in terms of the time allocation ratio α with closed-form expressions, which are respectively denoted as $\mathcal{P}^*(\alpha)$ and $\mathcal{P}_R^*(\alpha)$. Since the variables α , $\bar{\mathcal{S}}$ and $\bar{\mathcal{S}}_R$ are coupled to the packet dropping rate and the transmit power, it's also too complex to solve (32). Then, we provide a sub-optimal solution for problem (32) in two stages. In the first stage, we fix α for some value in the interval (0, 1). To this end, a numerical search algorithm based on a Golden Section search method [48] can be utilized to get the energy efficient solution $\bar{\mathcal{S}}^{\text{opt}}$ and $\bar{\mathcal{S}}_R^{\text{opt}}$. In the second stage, we still apply numerical search of the single variable α over the interval (0, 1) to obtain the energy efficient time allocation ratio α^* that maximizes the energy efficiency η_{ee} . To summarize, we illustrate the numerical search based procedure in the following Algorithm 1.

E. Energy Efficiency of Direct Link Transmission

Based on (27), the average transmit power consumption of direct link at channel state n is calculated as

$$\bar{\mathcal{P}}_n^{S,D} = \int_{\mathcal{S}_n}^{\mathcal{S}_{n+1}} \frac{\bar{e}(2^{b_{S,D}}-1)}{1.5\mathcal{S}_{S,D}} \ln \frac{0.2}{\vartheta_{S,D}} f_{\bar{\mathcal{S}}_{S,D}}(\mathcal{S}) d\mathcal{S}, \quad (33)$$

where $b_{S,D}$, $\bar{\mathcal{S}}_{S,D}$ and $\vartheta_{S,D}$ denote the modulation size, the average SNR and the BER for the direct source-to-destination transmission, respectively. Then, the average transmit power for direct link transmission can be obtained as

$$\mathcal{P}_{S,D} = \sum_{n=1}^N \bar{\mathcal{P}}_n^{S,D} P_r^{S,D}(n). \quad (34)$$

According to (34) and lemma 4, the total energy consumption with direct transmission period T can be written as

$$\zeta = \left(a_{A,D}^a \mathcal{P}_{S,D} + a_{A,D}^u \mathcal{P}_0\right) T \quad (35)$$

where \mathcal{P}_0 is power consumption for the source-node being at idle state.

Based on the packet dropping rate P_d of the source node and (21), we can get the throughput of the direct transmission system as

$$\mathcal{R} = \bar{\lambda}(1 - P_d) \left(1 - (1 - \bar{P}_{LD})^{N_r^{\max}}\right). \quad (36)$$

Based on (36) and (35), the energy efficiency can be denoted as

$$\begin{aligned} \eta_{ee} &= \frac{T\bar{\lambda}(1 - P_d) \left(1 - (1 - \bar{P}_{LD})^{N_r^{\max}}\right)}{\zeta} \\ &= \frac{\bar{\lambda}(1 - P_d) \left(1 - (1 - \bar{P}_{LD})^{N_r^{\max}}\right)}{a_{A,D}^a \mathcal{P}_{S,D} + a_{A,D}^u \mathcal{P}_0}. \end{aligned} \quad (37)$$

Similarly, we aim to optimize the average SNR $\bar{\mathcal{S}}_{S,D}$ to achieve the energy-efficient direct link transmission based on the closed-form expression of the energy efficiency, which can also be solved with the Algorithm 1.

IV. ENERGY EFFICIENT SCHEDULING

In this section, we investigate the relationship between the transmission mode selection and packet delay demand, where the transmission mode includes the direct link transmission and relay-assisted cooperative transmission. Then we determine the energy efficient modulation scheduling strategy with the effects of the delay requirements.

For scheduling heterogeneous flows, the evaluation of end-to-end delays involves in avionics, multimedia (video and audio) and best-effort data virtual links [49]. Their performance depends upon scheduling policy, and different scheduling policies between direct transmission and relay-assisted transmission have different results of energy efficiencies. Therefore, we should determine the energy efficient scheduling policy for different flows having different delay requirements.

Theorem 2: In the opportunistic spectrum access cooperative network, to ensure energy efficient packet transmission, there exists a delay threshold $\mathcal{D}^* = \phi(\bar{\mathcal{S}}^*)$, which determines the selection between two transmission modes. If the access probabilities satisfy $a_{A,D}^a > \frac{a_{A,R}^a + a_{R,D}^a \cdot \alpha}{1 + \alpha}$, while the delay requirement satisfies $\mathcal{D}_0 < \mathcal{D}^*$, the packet is transmitted only over direct link $l_{A,D}$; While the delay requirement satisfies $\mathcal{D}_0 > \mathcal{D}^*$, the packet is transmitted over the cooperative links $l_{A,R}$ and $l_{R,D}$. On the other hand, if $a_{A,D}^a < \frac{a_{A,R}^a + a_{R,D}^a \cdot \alpha}{1 + \alpha}$, while the delay requirement satisfies $\mathcal{D}_0 < \mathcal{D}^*$, the packet is transmitted over the cooperative links $l_{A,R}$ and $l_{R,D}$; While the delay requirement satisfies $\mathcal{D}_0 > \mathcal{D}^*$, the packet is transmitted only over direct link $l_{A,D}$.

Proof: The proof is given in Appendix D. ■

Proposition 1 (Adopted From Theorem 2): The energy efficient modulation scheduling policy is $\omega^{opt} = \omega(\mathcal{S}_n^*, \bar{\mathcal{S}}^{opt}(\mathcal{D}_0))$ when $\mathcal{D}_0 > \mathcal{D}^*$. On the other hand, the energy efficient modulation scheduling policies are $\omega^{opt} = \omega(\mathcal{S}_n^*, \bar{\mathcal{S}}^{opt}(\mathcal{D}_0))$ and $\omega^{opt} = \omega(\mathcal{S}_n^*, \bar{\mathcal{S}}_R^{opt}(\mathcal{D}_0))$ for the source node and relay node when $\mathcal{D}_0 < \mathcal{D}^*$.

Proposition 1 is obtained by the direct transmission and relay-assisted transmission solutions, respectively, where $\bar{\mathcal{S}}^{opt}(\mathcal{D}_0)$ is the energy-efficient average SNR when the maximum tolerable delay is \mathcal{D}_0 . Based on Theorem 2 and Proposition 1, we can get the energy efficient solution of switching and modulation scheduling policies for different delay-aware services.

V. SIMULATION RESULTS

In this section, we validate the accuracy of the analytical results for delay-aware energy efficient scheduling over the cooperative wireless system by means of simulations. In all the simulations, the opportunistic spectrum access is considered.

A. System Parameters

The corresponding simulation parameters are shown as follows: the number of traffic states is $K = 2$, the channel states for the inter-node link and direct link is $N = 7$, and the buffer size is $M + 1 = 51$ for the source node and relay node; The packet size $L = 100$, the maximum retransmission times of packet is $N_r^{\max} = 6$. The maximum number of packet arrivals for source node is set as $A = 15$, with an average arrival rate, $\lambda_1 = 1$ packets/time-unit and $\lambda_2 = 2$ packets/time-unit, the symbol rate $R_s = 100\text{kHz}$; The Nakagami parameter $m = 1$, which corresponds to Rayleigh fading channel with no line of sight (LOS) component. For mathematical tractability, we assume $\bar{\mathcal{S}}_R = \bar{\mathcal{S}}$.

B. Performance Evaluation

In Fig. 2, the metric of energy efficiency based on (31) is investigated over different values of time allocation ratio α . For any time allocation ratio $\alpha \in (0, 1)$, we can get the energy efficient SNR as $\bar{\mathcal{S}}^{opt}(\alpha)$ by Algorithm 1, and the corresponding energy efficient power allocation can be calculated based on (27) and (34). In this figure, we plot the energy efficiency based on energy efficient partition method (EEP) as well as the minimum SNR required to achieve P_{target} (MSRE) method for comparison (see [41]). We can observe from this figure that the energy efficiency of EEP is better than that of MSRE, since we use the energy efficient threshold values for choosing the modulation. In the case of average SNR $\bar{\mathcal{S}} = 5$ dB, Fig. 2 shows that the optimal time allocation ratio is $\alpha^* = 0.53$. In the case of $\bar{\mathcal{S}} = 8$ dB, the optimal time allocation ratio is the same, since we assume that the average SNR for the source to the relay is the same with that of the relay to the destination.

Denote the spectrum access error as $\epsilon = \frac{|a_{A,R}^a - \hat{a}_{A,R}^a|}{a_{A,R}^a}$, where $\hat{a}_{A,R}^a$ is the estimated spectrum access probability. Fig. 3 shows the results of energy efficiency considering spectrum access probability and spectrum access error, which considers different spectrum access probabilities for the source-to-relay link. From this figure, it is observed that the energy efficiency performs worse with the spectrum access error than that with ideal spectrum access probability and the gap sizes of energy efficiencies first increase and then decrease as the spectrum access probability $a_{A,R}^a$ grows, which means the throughput increased faster than the power consumption at first. This is

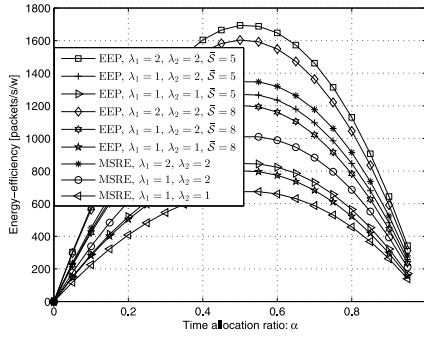


Fig. 2. The average energy efficiency versus the time allocation ratio α for relay-assisted transmission under the energy efficient partition method (EEP) and minimum SNR required to achieve P_{target} (MSRE) method, where $P_{\text{target}} = 0.001 \frac{1}{N_r^{\max} + 1}$.

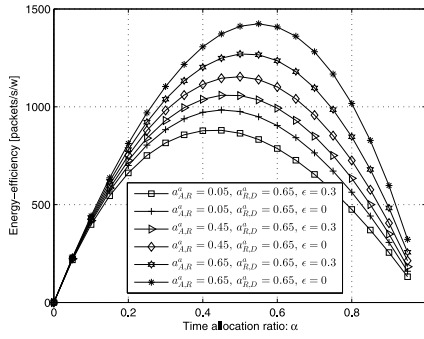


Fig. 3. The average energy efficiency versus the time allocation ratio α for relay assisted transmission considering the spectrum access probability.

because the idle power consumption increases as the spectrum access probability decreases. It could be concluded from this figure that, the throughput cannot increase as the time allocation ratio α is at large region, although the spectrum access probability increases. In fact, the packet dropping rate of relay-buffer increases as the spectrum access probability increases when the source-to-relay time allocation ratio is at large region, causing the decreasing throughput. Fig. 4 illustrates the influences of average SNR to the energy efficiencies of the direct transmission and relay-assisted transmission for different channel parameter m , where we assume that the direct link transmission and relay-assisted transmission have the same average SNR at the destination node. We observe that with the increasing \bar{S} , the energy efficiency of direct transmission gradually lose its superiority in energy efficiency. This corresponds to the analytical results in Theorem 2.

Fig. 5 shows the results of the total power consumption in relay-assisted transmission systems including transmit power and idle state power. From this figure we can observe that the total power consumption is increasing with increasing SNR, since the transmit powers for the source node and relay node are increasing with increasing SNR. We can also observe that the power is larger when the time allocation ratio is larger, which comes from the fact that the holding probability for the source to the relay transmission is larger than that of the relay to the destination transmission. At the same time, the gap of the power consumption is increasing when increasing SNR

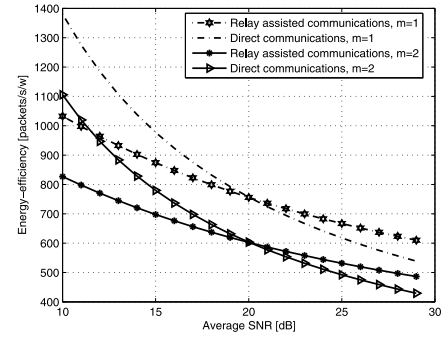


Fig. 4. The energy efficiencies of direct transmission and relay-assisted transmission, $\alpha = \alpha^*$.

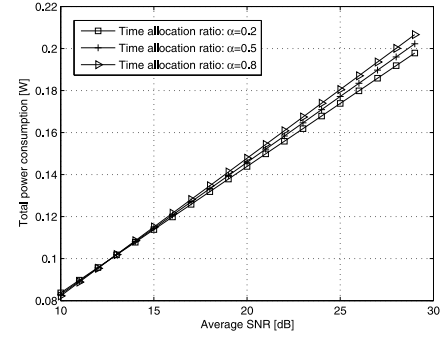


Fig. 5. The results of the average total power consumption versus average SNR for relay assisted transmission under Nakagami- m fading channel with $m = 1$.

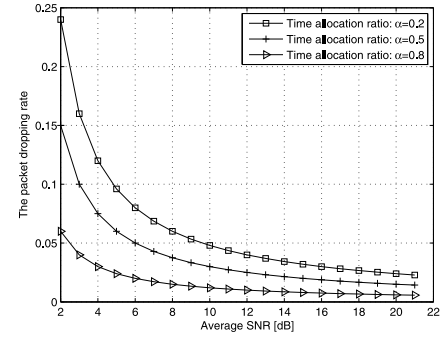


Fig. 6. The results of the packet dropping rate versus average SNR for relay assisted transmission with different time allocation ratio.

for different time allocation ratios, since larger α has smaller packet dropping rate for the relay-assisted transmission system at large SNR. Then the transmit rate is larger. Based on (30), we can get the result that the total power consumption is larger when α is larger.

In Fig. 6 and Fig. 7, we show the packet dropping rate of source node A, P_d^A , for the proposed relay assisted transmission. We can see in Fig. 6 that the packet dropping rate is decreasing when increasing the SNR, since the service rate for the source node is increasing when increasing the SNR. Moreover, we can observe that when the ratio of the source-to-relay time allocation over the relay-to-destination time allocation becomes larger, the packet dropping rate for source node A becomes smaller. When the source-to-relay

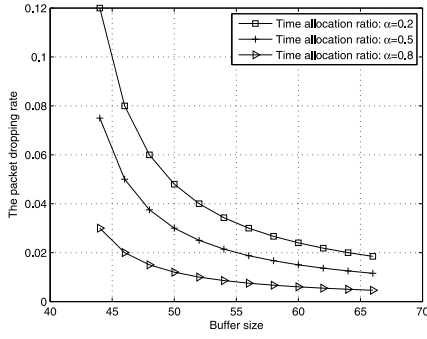


Fig. 7. The results of the packet dropping rate versus buffer size.

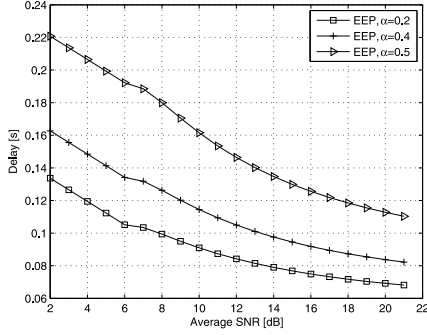


Fig. 8. The results of the delay for relay assisted transmission with different time allocation ratio.

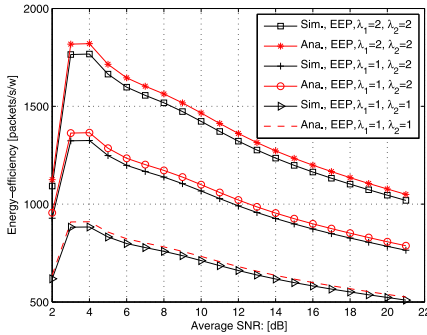


Fig. 9. The performance of energy efficiency versus average SNR for relay assisted transmission.

time allocation is increasing, the dropped number of packets for the source node during each transmission period is decreasing. Then the packet dropping rate becomes smaller. Similarly, it can be seen from Fig. 7 that the packet dropping rate is decreasing when increasing the buffer size. Due to smaller dropped number of packets, the larger source-to-relay time allocation ratio leads to smaller packet dropping rate. In Fig. 8, we show the delay performance for the relay-assisted transmission system when the average packet arrival rate $\bar{\lambda} = 2$. From this figure we can see that when the SNR is increasing, the delay is decreasing, which matches the results of Lemma 3. We can also see that the delay is smaller when the time allocation ratio is larger. This phenomenon can be understood as follows: when the packet arrival rate is larger, the packet dropping rate of the source node is in dominant place with small time allocation ratio. Then the delay can be smaller when increasing the source-to-relay time allocation.

In Fig. 9, the impact of the arrival traffic on the energy efficiency is studied for the relay-assisted transmission, when the time allocation ratio $\alpha = 0.6$. This figure shows the results of analysis and simulated values of energy efficiency for different average traffic rates. It can be observed that the variations of analytical and simulation results agree reasonably well. This figure shows that the energy efficiency is decreasing when increasing the SNR for given average arrival traffic $\bar{\lambda}$ in the regime of large SNR. On the other hand, the energy efficiency is increasing when increasing SNR in the regime of small SNR. In this case, we can get the energy efficient SNR by Algorithm 1. Thus, the energy efficient transmission policy ω^{opt} can be obtained, which also guarantees the delay requirement based on (24) and (25). For the case of different $\bar{\lambda}$, the energy efficiency is increasing when increasing the average arrival traffic rate $\bar{\lambda}$. This comes from the fact that the throughput is larger for larger arrival traffic rate.

VI. CONCLUSION

In this paper, we present a cross-layer framework in the cognitive cooperative network that determines the energy efficient transmission policy based on both the physical and the upper-layer information. We derive the closed-form expression of delay for the relay-assisted transmission system, which considers the queueing delay of the relay-buffer and source-buffer. With regard to the physical layer transmission, we propose a switching strategy between the direct transmission and the relay-assisted transmission in the cooperative transmission system under different delay requirements. Furthermore, for the relay-assisted transmission, we obtain the energy efficient transmission time allocation ratio for the inter-node link. At last, we get the transmission policies for the source node and relay node that achieve the energy efficient cooperative transmissions for different delay-aware services. Future work is in progress to consider more general case that arbitrary relay nodes are deployed to assist in conveying the source packet in the proposed framework.

APPENDIX A PROOF OF LEMMA 1

According to the packet error rate P_1 in the first phase, the packet service time δ_n in state n has the following probability mass function based on the M/G/1 queue model:

$$P\{\delta_n = \epsilon_n + \top \bar{\epsilon}_n\} = (1 - P_1)P_1^\top, \top = 0, 1, \dots, N_r^{\max}, \quad (38)$$

where ϵ_n , $\bar{\epsilon}_n$ and \top represent the packet transmission time when the channel is in state n , the average packet transmission time for retransmissions and the retransmission times, respectively. When the packet is transmitted from the source-to-destination link, the packet successfully received time at channel state n and average retransmission time for each retransmission are

$$\tau_n^{l_{A,D}} = \frac{L}{b_n R_s a_{A,D}^a}, \bar{\tau}_n^{l_{A,D}} = \sum_{n=1}^N \tau_n^{l_{A,D}} P_r^{S,D}(n), \quad (39)$$

respectively. Due to the independence of inter-node link, the channel state of relay-to-destination link would stay at any state of $\{1, 2, \dots, N\}$ when source-to-destination link is at state n . Then, we use the average rate of $\bar{b} = \sum_{n=1}^N b_n P_r(n)$ as the transmission modulation for relay-to-destination link. Similarly, the packet successfully received time from relay-assisted transmission at channel state n and average retransmission time for each retransmission are

$$\tau_n^{l_{A,R}, l_{R,D}} = \frac{L}{b_n R_s a_{A,R}^a} + \frac{L}{\bar{b} R_s a_{R,D}^a},$$

$$\text{and } \bar{\tau}_n^{l_{A,R}, l_{R,D}} = \sum_{n=1}^N \tau_n^{l_{A,R}, l_{R,D}} P_r(n), \quad (40)$$

respectively. In the sequel, the average transmission time of the packet at channel state n and average retransmission time for each retransmission are

$$\epsilon_n = \tau_n^{l_{A,D}} (1 - P_{LD}) + \tau_n^{l_{A,R}, l_{R,D}} P_{LD},$$

$$\text{and } \bar{\epsilon}_n = \bar{\tau}_n^{l_{A,D}} (1 - P_{LD}) + \bar{\tau}_n^{l_{A,R}, l_{R,D}} P_{LD}, \quad (41)$$

respectively, which can be denoted as (17).

Based on (38), the average service time at channel state n for the source transmission can be derived as

$$E\{\delta_n\} = \epsilon_n (1 - P_1) + \sum_{\tau=1}^{N_r^{\max}} (\epsilon_n + \tau \bar{\epsilon}_n) (1 - P_1) P_1^\tau$$

$$= \frac{f(\epsilon_n) + \bar{\epsilon}_n P_1 \left[1 - P_1^{N_r^{\max}} (1 + N_r^{\max} (1 - P_1)) \right]}{1 - P_1}, \quad (42)$$

where $f(\epsilon_n) = \epsilon_n (1 - P_1^{N_r^{\max}+1}) (1 - P_1)$.

Recall that the ratio of the time allocation in the phase of source-to-relay over the whole period is α . To see this, based on (17) and (42), we can derive the service rate χ_n at channel state n as

$$\chi_n = \frac{\alpha}{E\{\delta_n\}}$$

$$= \frac{\alpha (1 - P_1)}{f(\epsilon_n) + \bar{\epsilon}_n P_1 \left[1 - P_1^{N_r^{\max}} (1 + N_r^{\max} (1 - P_1)) \right]}. \quad (43)$$

APPENDIX B

PROOF OF THEOREM 1

Since the average transmission rate from the source to the relay is $\bar{\chi}$, the average packet arrival rate at the relay is $\bar{\chi}$, which can be derived as $\bar{\chi} = \sum_{n=1}^N \chi_n P_r(n)$ based on Lemma 1. Then, we can get the probability distribution of the instantaneous arrival rate a_t^R at the relay as

$$P(a_t^R = a) = \begin{cases} P_r(n), & \text{if } a \in \{\chi_1, \dots, \chi_N\}; \\ \frac{\lambda^a e^{-\lambda}}{a!}, & \text{otherwise.} \end{cases} \quad (44)$$

Assume that the queue state of the relay node is S_t^R at the start of the t -th time slot, and $S_t^R \in \mathbb{S} = \{s'_0 = 0, s'_1 = 1, \dots, s'_M = M\}$. While AMC transmission is used by the relay, it is known that $\chi'_t \in \Psi' = \{\chi'_1, \dots, \chi'_N\}$ can be denoted as the number of packets removed from the queue of the relay. Then, to obtain the resulting transition of the queue state at the relay, both buffer size and the service rate of relay node need to be considered, shown as

$$S_t^R = \min \left\{ M, \max \left\{ 0, S_{t-1}^R - \chi'_t \right\} + \bar{\chi} \right\}. \quad (45)$$

In order to express the system state transition, let (χ'_t, S_{t-1}^R) denote the joint service rate and the queue state, and let $\mathbb{B}_{(\chi'_x, s'_q), (\chi'_y, s'_l)}$ denote the transition probability from $(\chi'_t = \chi'_x, S_{t-1}^R = s'_q)$ to $(\chi'_{t+1} = \chi'_y, S_t^R = s'_l)$, where $(\chi'_x, s'_q) \in \Psi' \times \mathbb{S}$, and $(\chi'_y, s'_l) \in \Psi' \times \mathbb{S}$. Then, the state transition probability matrix of the relay transmission system can be organized in a block form as

$$\mathbf{T} = [\Theta_{\chi'_i, \chi'_j}], 1 \leq i, j \leq N, \quad (46)$$

where

$$\Theta_{\chi'_i, \chi'_j} = \begin{bmatrix} \Xi_{(\chi'_1), (\chi'_1)} & \Xi_{(\chi'_1), (\chi'_2)} & \cdots & \Xi_{(\chi'_1), (\chi'_N)} \\ \Xi_{(\chi'_2), (\chi'_1)} & \Xi_{(\chi'_2), (\chi'_2)} & \cdots & \Xi_{(\chi'_2), (\chi'_N)} \\ \vdots & \vdots & \ddots & \vdots \\ \Xi_{(\chi'_N), (\chi'_1)} & \Xi_{(\chi'_N), (\chi'_2)} & \cdots & \Xi_{(\chi'_N), (\chi'_N)} \end{bmatrix}, \quad (47)$$

the submatrix $\Xi_{(\chi'_x), (\chi'_y)}$ can be shown in (48), shown at the bottom of the page.

Finally, the system state transition probability $\mathbb{B}_{(\chi'_x, s'_q), (\chi'_y, s'_l)}$ can be obtained as (49), shown at the

$$\Xi_{(\chi'_x), (\chi'_y)} = \begin{bmatrix} \mathbb{B}_{(\chi'_x, s'_0), (\chi'_y, s'_0)} & \mathbb{B}_{(\chi'_x, s'_0), (\chi'_y, s'_1)} & \cdots & \mathbb{B}_{(\chi'_x, s'_0), (\chi'_y, s'_M)} \\ \mathbb{B}_{(\chi'_x, s'_1), (\chi'_y, s'_0)} & \mathbb{B}_{(\chi'_x, s'_1), (\chi'_y, s'_1)} & \cdots & \mathbb{B}_{(\chi'_x, s'_1), (\chi'_y, s'_M)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{B}_{(\chi'_x, s'_M), (\chi'_y, s'_0)} & \mathbb{B}_{(\chi'_x, s'_M), (\chi'_y, s'_1)} & \cdots & \mathbb{B}_{(\chi'_x, s'_M), (\chi'_y, s'_M)} \end{bmatrix} \quad (48)$$

$$\mathbb{B}_{(\chi'_x, s'_q), (\chi'_y, s'_l)} = \mathbb{B}(\chi'_{t+1} = \chi'_y, S_t^R = s'_l | \chi'_t = \chi'_x, S_{t-1}^R = s'_q)$$

$$= \mathbb{B}(\chi'_{t+1} = \chi'_y | \chi'_t = \chi'_x) \mathbb{B}(S_t^R = s'_l | S_{t-1}^R = s'_q, \chi'_t = \chi'_x)$$

$$= \mathbb{B}_{\chi'_x, \chi'_y} \mathbb{B}(S_t^R = s'_l | S_{t-1}^R = s'_q, \chi'_t = \chi'_x) \quad (49)$$

bottom of the previous page, where $B_{\chi'_x, \chi'_y}$ represents the channel state transition probability. Due to the fact that the channel state transition is independent of other states, the second equality in (49) can be obtained.

By considering the slow fading channel, the channel state transition only happens between adjacent states. Then, the transition probability from non adjacent states is zero, and the nonzero elements $B_{\chi'_x, \chi'_y}$ is described in [44]. In addition, with regard to the conditional probability of $B(S_t^R = s'_l | S_{t-1}^R = s'_q, \chi'_t = \chi'_x)$, it can be derived in (50), shown at the bottom of the page. In all, based on (49) and (50), the transition probability of the system state for the relay node can be obtained.

Notably particularly, by using the results from [50, Th. 4.1], we obtain the fact that the Markov chain $\{(S_t^R, \chi'_t), t \geq 0\}$ exists stationary distribution. From the definition of irreducible Markov chain, we know that the Markov chain only has one class. Under this circumstance, while any state $(\chi'_x, s'_q) \in \Psi' \times \mathbb{S}$ and $(\chi'_y, s'_l) \in \Psi' \times \mathbb{S}$, we only need to prove that (χ'_x, s'_q) can access (χ'_y, s'_l) .

In the following, we will prove that the Markov chain of relay system is irreducible. Based on the packet arrival probability, we can obtain the transition probability

$$B_{(\chi'_x, s'_q) | (\chi'_x, s'_l)} = P(a_t^R = s'_l - \max\{0, s'_q - \chi'_x\}). \quad (51)$$

Thus, state (χ'_x, s'_q) can access state (χ'_x, s'_l) , i.e.,

$$(\chi'_x, s'_q) \Rightarrow (\chi'_x, s'_l) \quad (52)$$

has nonzero probability. Note that the channel has finite states. Thus, there always exists transition path from the channel state x to the state y over the neighbour state. In this case, state χ'_x can go to state χ'_y , and

$$B_{(\chi'_x, s'_l) | (\chi'_y, s'_l)} = P(a_t^R = s'_l - \max\{0, s'_l - \chi'_x\}). \quad (53)$$

Therefore, state (χ'_x, s'_l) can access state (χ'_y, s'_l) , i.e.,

$$(\chi'_x, s'_l) \Rightarrow (\chi'_y, s'_l) \quad (54)$$

also has nonzero transition probability. With the results of (52) and (54), we prove that any state (χ'_x, s'_q) can access any state (χ'_y, s'_l) in the Markov chain of relay system. Then, we obtain that the system $\{(S_t^R, \chi'_t), t \geq 0\}$ is irreducible.

When implementing the conclusion of [50], the states from a finite irreducible Markov chain are recurrent. In all, the stationary distribution of the Markov process $\{(S_t^R, \chi'_t), t \geq 0\}$ exists. Then, we can get the stationary distribution by solving

$$\pi' = \pi' \Upsilon, \quad \sum_{s' \in \mathbb{S}, \chi' \in \Psi'} \pi'_{(s', \chi')} = 1. \quad (55)$$

while the number of arrival packets is larger than that of remaining space, i.e., $\bar{\chi} > M - (S_{t-1}^R - \chi'_t)$, the packets are dropped by the relay-buffer. By using the stationary

distribution, the average number of dropped packets can be expressed as

$$\begin{aligned} E\{\Delta^R\} &= \sum_{s' \in \mathbb{S}, \chi' \in \Psi'} \Delta^R \cdot B(\bar{\chi} | S_{t-1}^R = s', \chi'_t = \chi') \\ &= \sum_{s' \in \mathbb{S}, \chi' \in \Psi'} \left[(\bar{\chi} - (M - (S_{t-1}^R - \chi'_t))) \times \pi'_{s', \chi'} \right]. \end{aligned} \quad (56)$$

In all, the packet dropping rate at the relay node is $P_d^R = \frac{E\{\Delta^R\}}{\bar{\chi}}$.

APPENDIX C PROOF OF LEMMA 4

In Phase source node to relay node, source A transmits data to the relay R with power \mathcal{P} . Due to the opportunistic spectrum access, the average transmit power is $a_{A,R}^a \mathcal{P}$, the average circuit power consumption is $a_{A,R}^u \mathcal{P}_0$.

On the other hand, in Phase relay node to destination node, relay R transmits data to the destination D with power \mathcal{P}_R . Similarly, the average transmit power and circuit power of the relay node are $a_{R,D}^a \mathcal{P}_R$ and $a_{R,D}^u \mathcal{P}_0$, respectively. In all, the total energy consumption of the secondary relay-assisted transmission networks is given by

$$\zeta = (a_{A,R}^a \mathcal{P} + a_{A,R}^u \mathcal{P}_0) T_1 + (a_{R,D}^a \mathcal{P}_R + a_{R,D}^u \mathcal{P}_0) T_2. \quad (57)$$

APPENDIX D PROOF OF THEOREM 2

For the first case, suppose $\mathcal{D}_0 \rightarrow 0$, based on Lemma 3, the average SNR should satisfy $\bar{S}, \bar{S}_R \rightarrow \infty$, causing $P_d \rightarrow 0$ and $P_0 \rightarrow 0$. Thus, the average throughput $\mathcal{R} = \bar{\lambda}$. At the same time, the total power consumption is dominated by the transmit power. Consequently, based on Lemma 4, the total average energy consumption of source-to-destination direct transmission is $\zeta = a_{A,D}^a T \mathcal{P}_{S,D}$. In all, the energy efficiencies of the direct transmission and the relay-assisted transmission are

$$\begin{aligned} \lim_{\mathcal{D}_0 \rightarrow 0} \eta_{ee, direct} &= \frac{\bar{\lambda}}{a_{A,D}^a \mathcal{P}}, \text{ and} \\ \lim_{\mathcal{D}_0 \rightarrow 0} \eta_{ee, relay} &= \frac{T \bar{\lambda}}{a_{A,R}^a T_1 \mathcal{P} + a_{R,D}^a T_2 \mathcal{P}_R} \\ &= \frac{\bar{\lambda}}{a_{A,R}^a \cdot \alpha \mathcal{P} + a_{R,D}^a (1 - \alpha) \mathcal{P}}, \end{aligned} \quad (58)$$

respectively. From these two equations, we can obtain that if $a_{A,D}^a > \frac{a_{A,R}^a + a_{R,D}^a \cdot \alpha}{1 + \alpha}$, the energy efficiency of relay-assisted transmission is larger than that of the direct transmission. The packet should be transmitted with the cooperative links $l_{A,R}$ and $l_{R,D}$ to ensure energy efficient transmission.

$$B(S_t^R = s'_l | S_{t-1}^R = s'_q, \chi'_t = \chi'_x) = \begin{cases} P(a_t^R = s'_l - \max\{0, s'_q - \chi'_x\}), & \text{if } 0 \leq s'_l < M, \\ 1 - \sum_{0 \leq s'_l < M} B(S_t^R = s'_l | S_{t-1}^R = s'_q, \chi'_t = \chi'_x), & \text{if } s'_l = M \end{cases} \quad (50)$$

For the second case, as $\bar{\mathcal{S}}, \bar{\mathcal{S}}_R \rightarrow 0$, suppose $\mathcal{D}_0 \rightarrow \infty$, causing $P_{LD} \rightarrow 1$. Then, the idle state power dominates the total power consumption. Then the energy consumptions of direct transmission and relay-assisted transmission are

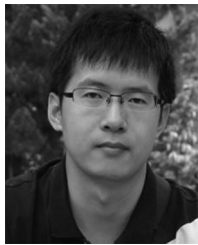
$$\zeta_{\text{direct}} = \left(a_{A,D}^u \mathcal{P}_0\right) T, \text{ and} \\ \zeta_{\text{relay}} = \left(a_{A,R}^u \mathcal{P}_0\right) T_1 + \left(a_{R,D}^u \mathcal{P}_0\right) T_2, \quad (59)$$

respectively. Note that spectrum access probabilities satisfy $a_{A,D}^u + a_{A,D}^a = 1$, $a_{A,R}^u + a_{A,R}^a = 1$ and $a_{R,D}^u + a_{R,D}^a = 1$. Similarly, we can also obtain that if $a_{A,D}^a > \frac{a_{A,R}^a + a_{R,D}^a \cdot \alpha}{1 + \alpha}$, the energy efficiency of the direct transmission is larger than that of the relay-assisted transmission. Therefore, the packet should be transmitted with the cooperative links $l_{A,R}$ and $l_{R,D}$.

REFERENCES

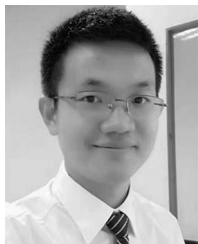
- [1] Y. Yang, K. Wang, W. Chen, M.-T. Zhou, and G. Mao, "Energy-efficient scheduling for buffer-aided relaying with opportunistic spectral access," in *Proc. IEEE WCSP*, Nanjing, China, Oct. 2017, pp. 1–6.
- [2] J. Ling, S. Kanugovi, S. Vasudevan, and A. K. Pramod, "Enhanced capacity and coverage by Wi-Fi LTE integration," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 165–171, Mar. 2015.
- [3] S. Zhang, Q. Wu, S. Xu, and G. Li, "Fundamental green tradeoffs: Progresses, challenges, and impacts on 5G networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 33–56, 1st Quart., 2016.
- [4] Y. Yang, J. Xu, G. Shi, and C.-X. Wang, *5G Wireless Systems: Simulation and Evaluation Techniques*, 1st ed. Cham, Switzerland: Springer Int., 2018.
- [5] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, "An overview of sustainable green 5G networks," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 72–80, Aug. 2017.
- [6] N.-T. Le, L.-N. Tran, Q.-D. Vu, and D. Jayalath, "Energy-efficient resource allocation for OFDMA heterogeneous networks," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7043–7057, Oct. 2019.
- [7] Q. Wu, W. Chen, M. Tao, J. Li, H. Tang, and J. Wu, "Resource allocation for joint transmitter and receiver energy efficiency maximization in downlink OFDMA systems," *IEEE Trans. Commun.*, vol. 63, no. 2, pp. 416–430, Feb. 2015.
- [8] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3412–3427, Jul. 2013.
- [9] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.
- [10] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Mar. 2018.
- [11] Z. Xu, C. Yang, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient configuration of spatial and frequency resources in MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 564–575, Feb. 2013.
- [12] Y. Rui, Q. T. Zhang, L. Deng, P. Cheng, and M. Li, "Mode selection and power optimization for energy efficiency in uplink virtual MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 926–936, May 2013.
- [13] X. Ge *et al.*, "Energy-efficiency optimization for MIMO-OFDM mobile multimedia communication systems with QoS constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, Jun. 2014.
- [14] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Mar. 2016.
- [15] Y. Niu *et al.*, "Energy-efficient scheduling for mmWave backhauling of small cells in heterogeneous cellular networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2674–2687, Jun. 2017.
- [16] J. Mao, G. Xie, J. Gao, and Y. Liu, "Energy efficiency optimization for OFDM-based cognitive radio systems: A waterfilling factor aided search method," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2366–2375, May 2013.
- [17] L. Fu, Y. J. A. Zhang, and J. Huang, "Energy efficient transmissions in MIMO cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2420–2431, Nov. 2013.
- [18] J. Zhang, L. Xiang, D. W. K. Ng, M. Jo, and M. Chen, "Energy efficiency evaluation of multi-tier cellular uplink transmission under maximum power constraint," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7092–7107, Nov. 2017.
- [19] Q. Wu, G. Y. Li, W. Chen, and D. W. K. Ng, "Energy-efficient small cell with spectrum-power trading," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3394–3408, Dec. 2016.
- [20] D. W. K. Ng, E. S. Lo, and R. Schober, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6352–6370, Dec. 2013.
- [21] Q. Wu, M. Tao, D. W. K. Ng, W. Chen, and R. Schober, "Energy-efficient resource allocation for wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2312–2327, Mar. 2016.
- [22] D. Qiao, M. C. Gursoy, and S. Velipasalar, "Energy efficiency in the low-SNR regime under queueing constraints and channel uncertainty," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2006–2017, Jul. 2011.
- [23] D. W. K. Ng, Y. Wu, and R. Schober, "Power efficient resource allocation for full-duplex radio distributed antenna networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2896–2911, Apr. 2016.
- [24] N. Nomikos, T. Charalambous, I. Krikidis, D. Skoutas, D. Vouyioukas, and M. Johansson, "A buffer-aided successive opportunistic relay selection scheme with power adaptation and inter-relay interference cancellation for cooperative diversity systems," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1623–1634, May 2015.
- [25] G. Yuan, R. C. Grammenos, Y. Yang, and W. Wang, "Performance analysis of selective opportunistic spectrum access with traffic prediction," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1949–1955, May 2010.
- [26] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMax and LTE-advanced mobile systems," *IEEE Commun. Mag.*, vol. 47, no. 10, pp. 100–105, Oct. 2009.
- [27] Z. Zhu, S. Jin, Y. Yang, H. Hu, and X. Luo, "Time reusing in D2D-enabled cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3185–3200, May 2018.
- [28] N. Zlatanov, A. Ikhlef, T. Islam, and R. Schober, "Buffer-aided cooperative communications: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 146–153, Apr. 2014.
- [29] N. Nomikos *et al.*, "A survey on buffer-aided relay selection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1073–1097, 2nd Quart., 2016.
- [30] N. Zlatanov, V. Jamali, and R. Schober, "Achievable rates for the fading half-duplex single relay selection network using buffer-aided relaying," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4494–4507, Aug. 2015.
- [31] D. Qiao, "Effective capacity of buffer-aided full-duplex relay systems with selection relaying," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 117–129, Jan. 2016.
- [32] C. Huang, J. Zhang, H. V. Poor, and S. Cui, "Delay-energy tradeoff in multicast scheduling for green cellular systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1235–1249, May 2016.
- [33] S. S. Ikki and M. H. Ahmed, "Performance analysis of cooperative diversity with incremental-best-relay technique over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2152–2161, Aug. 2011.
- [34] A. Nasri, R. Schober, and I. F. Blake, "Performance and optimization of amplify-and-forward cooperative diversity systems in generic noise and interference," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1132–1143, Apr. 2011.
- [35] B. Rong and E. Anthony, "Cooperative access in wireless networks: Stable throughput and delay," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5890–5907, Aug. 2012.
- [36] Y. Cui, V. K. N. Lau, and E. M. Yeh, "Delay-optimal scheduling for cooperative networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2011, pp. 963–967.
- [37] D. Xue and E. Ekici, "Cross-layer scheduling for cooperative multi-hop cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 534–543, Mar. 2013.
- [38] C. Lin, Y. Liu, and M. Tao, "Cross-layer optimization of two-way relaying for statistical QoS guarantees," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1583–1596, Aug. 2012.
- [39] N. Wang and T. A. Gulliver, "Cross-layer AMC scheduling for a cooperative wireless communication system over Nakagami- m fading channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2330–2341, Jun. 2012.

- [40] A. Rao, H. Ma, M.-S. Alouini, and Y. Chen, "Impact of primary user traffic on adaptive transmission for cognitive radio with partial relay selection," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1162–1172, Mar. 2013.
- [41] K. Wang, M. Tao, W. Chen, and Q. Guan, "Delay-aware energy-efficient communications over Nakagami- m fading channel with MMPP traffic," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 3008–3020, Aug. 2015.
- [42] B. L. Mark and Y. Ephraim, "Explicit causal recursive estimators for continuous-time bivariate Markov chains," *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2709–2718, May 2014.
- [43] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, 4th ed. Hoboken, NJ, USA: Wiley, 2008.
- [44] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.
- [45] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 785–796, Feb. 2008.
- [46] Q. Liu, S. Zhou, and G. B. Giannakis, "Queueing with adaptive modulation and coding over wireless links: Cross-layer analysis and design," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1142–1153, May 2005.
- [47] J. S. Harsini and F. Lahouti, "Adaptive transmission policy design for delay-sensitive and Bursty packet traffic over wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 776–786, Feb. 2009.
- [48] E. M. T. Hendrix and B. G.-Toth, *Introduction to Nonlinear and Global Optimization*. New York, NY, USA: Springer, 2010.
- [49] Y. Hua and X. Liu, "Scheduling heterogeneous flows with delay-aware deduplication for avionics applications," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 9, pp. 1790–1802, Sep. 2012.
- [50] S. M. Ross, *Introduction to Probability Models*, 9th ed. San Diego, CA, USA: Academic, 2007.



Kunlun Wang (Member, IEEE) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2016. From 2016 to 2017, he was with Huawei Technologies Company, Ltd., where he was involved in energy efficiency algorithm design. From 2017 to 2019, he was with the Key Laboratory of Wireless Sensor Network and Communication, SIMIT, Chinese Academy of Sciences, Shanghai. Since March 2019, he has been a Research Assistant Professor with the School of Information Science and Technology,

ShanghaiTech University. He is also with the Shanghai Institute of Fog Computing Technology. His current research interests include energy efficient communications, fog computing networks, resource allocation, and optimization algorithm.



Qingqing Wu (Member, IEEE) received the B.Eng. degree in electronic engineering from the South China University of Technology in 2012, and the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University (SJTU) in 2016. He is currently a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore. He has published more than 60 papers in IEEE top-tier journals and conferences, including ten ESI highly cited papers, two ESI hot papers, and one editor invited paper. His current research interests include intelligent reflecting surface, unmanned aerial vehicle (UAV) communications, and green 5G communications. He was a recipient of the Outstanding Ph.D. Thesis Funding in SJTU in 2016 and the Outstanding Ph.D. Thesis Award of China Institute of Communications in 2017. He received the IEEE WCSP Best Paper Award in 2015; the Exemplary Reviewer of IEEE COMMUNICATIONS LETTERS in 2016 and 2017, IEEE WIRELESS COMMUNICATIONS LETTERS in 2018, the IEEE TRANSACTIONS ON COMMUNICATIONS in 2017 and 2018, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS in 2017 and 2018. He serves as an Associate Editor for IEEE COMMUNICATIONS LETTERS and the IEEE OPEN JOURNAL OF COMMUNICATIONS SOCIETY, the Guest Editor for the IEEE OPEN JOURNAL ON VEHICULAR TECHNOLOGY on 6G Intelligent Communications, and the Leading Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on UAV Communications in 5G and Beyond Networks.



Wen Chen (Senior Member, IEEE) received the B.S. and M.S. degrees from Wuhan University, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro Communications, Tokyo, Japan, in 1999. He was a JSPS Fellow with the University of Electro Communications from 1999 to 2001. In 2001, he joined the University of Alberta, Canada, starting as a PIMS Postdoctoral Fellow and then as a Research Associate. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Institute for Signal Processing and Systems. He has published 99 papers in IEEE journals and more than 120 papers in IEEE conferences. His interests cover multiple access, coded cooperation, green heterogeneous networks, and vehicular communications. He is an editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE ACCESS, and the IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY. He is the Chair of the IEEE Vehicular Technology Society Shanghai Chapter. He is a Fellow of the Chinese Institute of Electronics and a Distinguished Lecturer of the IEEE Communications Society.



Yang Yang (Fellow, IEEE) received the B.Eng. and M.Eng. degrees in radio engineering from Southeast University, Nanjing, China, in 1996 and 1999, respectively, and the Ph.D. degree in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2002.

He has held faculty positions with the Chinese University of Hong Kong; Brunel University, U.K.; University College London, U.K.; and SIMIT, CAS, China. He is currently a Full Professor with ShanghaiTech University, China, serving as the

Executive Dean of School of Creativity and Art and the Co-Director of the Shanghai Institute of Fog Computing Technology. His current research interests include fog computing networks, service-oriented collaborative intelligence, wireless sensor networks, IoT applications, and advanced testbeds and experiments. He has published more than 200 papers and filed more than 80 technical patents in these research areas. He has been the Chair of the Steering Committee of the Asia-Pacific Conference on Communications since January 2019. He is the General Co-Chair of the IEEE DSP 2018 Conference and the TPC Vice-Chair of the IEEE ICC 2019 Conference.



Derrick Wing Kwan Ng (Senior Member, IEEE) received the bachelor's degree (First Class Hons.) and the M.Phil. degree in electronic engineering from the Hong Kong University of Science and Technology in 2006 and 2008, respectively, and the Ph.D. degree from the University of British Columbia in 2012. He was a Senior Postdoctoral Fellow with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg, Germany. He is currently working as a Senior Lecturer and a Scientia Fellow

with the University of New South Wales, Sydney, Australia. His research interests include convex and nonconvex optimization, physical layer security, wireless information and power transfer, and green (energy efficient) wireless communications. He received the Best Paper Awards at the IEEE TCGCC Best Journal Paper Award 2018, INISCOM 2018, the IEEE International Conference on Communications 2018, the IEEE International Conference on Computing, Networking and Communications 2016, the IEEE Wireless Communications and Networking Conference 2012, the IEEE Global Telecommunication Conference 2011, and the IEEE Third International Conference on Communications and Networking in China 2008. He was listed as a Highly Cited Researcher by Clarivate Analytics in 2018 and 2019. He served as an editorial assistant to the Editor-in-Chief for the IEEE TRANSACTIONS ON COMMUNICATIONS from January 2012 to December 2019. He is currently serving as an editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.